



Notes on Virtues

1. [Notes on notes on virtues](#)
2. [Notes on Honesty](#)
3. [Notes on Courage](#)
4. [Notes on Self Control](#)
5. [Notes on Respect-for-Others](#)
6. [Notes on Piety](#)
7. [Notes on Loyalty](#)
8. [Notes on Compassion](#)
9. [Notes on Wisdom](#)
10. [Notes on Temperance](#)
11. [Notes on Fitness](#)
12. [Notes on Sincerity and such](#)
13. [Notes on Justice \(as a virtue\)](#)
14. [Notes on Industriousness](#)
15. [Notes on Duty](#)
16. [Notes on Social Responsibility](#)
17. [Notes on Prudence](#)
18. [Notes on Know-how](#)
19. [Notes on Honor](#)
20. [Notes on Moderation, Balance, & Harmony](#)
21. [Notes on Patience & Forbearance](#)
22. [Notes on Care](#)
23. [Notes on Attention](#)
24. [Notes on Rationality](#)
25. [Notes on Amiability](#)
26. [Notes on Simplicity](#)
27. [Notes on Forgiveness](#)
28. [Notes on Integrity](#)
29. [Notes on Humility](#)
30. [Notes on Optimism, Hope, and Trust](#)
31. [Notes on Good Temper](#)
32. [Notes on Fairness](#)
33. [Notes on Endurance](#)
34. [Notes on Benevolence](#)
35. [Notes on Ambition](#)
36. [Notes on Perseverance](#)
37. [Notes on Kindness](#)
38. [Notes on Empathy](#)
39. [Notes on Frugality](#)
40. [Notes on Dignity](#)
41. [Notes on Courtesy](#)
42. [Notes on Chastity](#)
43. [Notes on Love](#)
44. [Notes on Resolve](#)
45. [Notes on Caution](#)
46. [Notes on Judgment and Righteous Anger](#)
47. [Notes on Gratitude](#)
48. [Notes on Shame](#)

Notes on notes on virtues

In the [Notes on Virtues sequence](#), I've been sharing my research into a variety of virtues. In this introductory post, I share my thoughts about why I've been on this case and what I'm trying to accomplish.

A one-paragraph summary: I think that becoming more skillful and well-rounded in the practice of the virtues is key to being a better, more satisfied, and more effective person. However, childhood training in the virtues is scattershot and haphazard, and remedial training (or self-help) as adults is also spotty. I'm trying to help fix this by assembling information about the various virtues, with a focus on ways to improve them in ourselves.

Why I think this is important

"The characteristic feature of all ethics is to consider human life as a game that can be won or lost and to teach man the means of winning." —Simone de Beauvoir^[1]

"The branch of philosophy on which we are at present engaged differs from the others in not being a subject of merely intellectual interest—I mean we are not concerned to know what goodness essentially *is*, but how we are to become good people, for this alone gives the study its practical value." —Aristotle^[2]

Life is complex. We are constantly confronted by a variety of challenges. To address those challenges well, we need to have learned a variety of basic life skills such that they are second-nature to us. "The virtues" are a set of such skills that apply to challenges common to typical human lives.

If you have a better command of the virtues, this helps you thrive as an individual and also improves your effect on those around you. Society at large benefits from a higher level of competence in the virtues of those in it. But our culture is not all that good at teaching or encouraging the virtues. Some virtues seem so lacking from the public sphere that I wince when I look at it.

Our institutions of formal childhood education are patchy at best in this regard. You'll get your reading, writing, and 'rithmetic, if you're lucky anyway, but will you get resourcefulness, resilience, restraint, responsibility, rectification, or reputability? Other institutions (scouting, religion, etc.) pick up some of the slack, but not nearly enough. Parents have little guidance on how to convey virtue education to their children effectively, and also have their own blind spots from their own spotty educations. There have been some gestures toward formal "character education" of children, which is probably a good sign. But my guess is that children are going to learn most from the example of their elders: if we don't value virtues enough to pursue them in our own lives, that will make more of an impression on the up-and-coming generation than any "do as we say, not as we do" education will.

A virtue gym?

For a few specific virtues or skills, there are adult education / training / exercise programs. If you want to be more fit, you can join a gym. If you want to be a better public speaker, you can join Toastmasters. If you want to sober up, you can attend Alcoholics Anonymous. But for most virtues, there's nothing like this, and that's a shame.

Two misconceptions that sometimes cause people to give up too early on developing virtues are these: 1) that virtues are talents that some people have and other people don't as a matter of predisposition, genetics, the grace of God, or what have you ("I'm just not a very influential / graceful / original person"), and 2) that having a virtue is not a matter of developing a habit but of having an opinion (e.g. I agree that creativity is good, and I try to respect the virtue of creativity that way, rather than by creating). It's more accurate to think of a virtue as a skill like any other. Like juggling, it might be hard at first, it might come easier to some people than others, but almost anyone can learn to do it if they're just willing to put in the persistent practice.

We are creatures of habit: We create ourselves by what we practice. If we adopt habits without giving them much consideration, we risk becoming what we never intended to be. If instead we deliberate carefully about what habits we want to cultivate, and then actually put in the work, we can become the sculptors of our own characters.

What if there were some institution like a "virtue gymnasium" through which you could work on virtues alongside others, learning at your own pace, and building a library of wisdom about how to go about it most productively? What if there were something like Toastmasters, or Alcoholics Anonymous, or the YMCA but for *all* of the virtues people need?

Ben Franklin's experiment

One day Benjamin Franklin "conceiv'd the bold and arduous project of arriving at moral perfection."^[3] He explains in his autobiography, "as I knew, or thought I knew, what was right and wrong, I did not see why I might not always do the one and avoid the other."

He quickly found that he had underestimated the task. "While my care was employ'd in guarding against one fault, I was often surprised by another; habit took the advantage of inattention; inclination was sometimes too strong for reason. I concluded, at length, that the mere speculative conviction that it was our interest to be completely virtuous, was not sufficient to prevent our slipping; and that the contrary habits must be broken, and good ones acquired and established, before we can have any dependence on a steady, uniform rectitude of conduct."

So he decided to be more methodical. He reviewed various lists of virtues in the literature he was familiar with, and then created his own list of a dozen virtues that he thought were particularly important. With the intention of making each of these virtues habitual, he struck on the idea of tackling them one-at-a-time, starting with ones he thought would help him more easily acquire the others. (Virtues have a way of building on each other. Some virtues, for example persistence, or curiosity, or honor, can make other virtues easier to acquire.)

He decided to do a daily accounting of each virtue he was practicing. He created a notebook with a table for each week. The table had one column for each day of the week, and one row for each of his virtues. Each time he failed to fulfill a particular

virtue on a certain day, he marked the table cell for that virtue/day with “a little black spot” (or more than one if he screwed up multiple times). The plan was that when he achieved a week in which he successfully kept the row for Temperance blank, he would move on to concentrating on Silence (attending to Temperance as well). When he managed to keep both of those rows clear for a week, he would move on to Order, and so on.

“I was surpris’d to find myself so much fuller of faults than I had imagined; but I had the satisfaction of seeing them diminish.” He carried his notebook with him for several years. “[T]ho’ I never arrived at the perfection I had been so ambitious of obtaining, but fell far short of it, yet I was, by the endeavour, a better and a happier man than I otherwise should have been...”

He hoped at one point to write a book, *The Art of Virtue*, which “would have shown the means and manner of obtaining virtue, which would have distinguished it from the mere exhortation to be good that does not instruct and indicate the means.”

He toyed with the idea of a political party that would not advocate for the benefit of a certain segment of the people, but for the good of the country and of mankind in general: the “United Party for Virtue.” This morphed into an idea for a fraternity: the “Society of the Free and Easy.” His plan was to initiate members by putting them through the same practice he had undergone with his notebook of weeks and virtues. He explained the name of the society this way:

The Society of the Free and Easy: free, as being, by the general practice and habit of the virtues, free from the dominion of vice; and particularly by the practice of industry and frugality, free from debt, which exposes a man to confinement, and a species of slavery to his creditors.

He got as far as getting two young men to sign up and begin the work, but then he got distracted with other things and abandoned it. “[T]ho’ I am still of opinion that it was a practicable scheme, and might have been very useful...”

The Society of the Free & Easy

Last year I set about trying to pull together something like Franklin’s Society of the Free and Easy (and borrowing his name). I worked with a group of friends and acquaintances to come up with what I think is [a pretty good framework for working on virtues in a peer-supported way](#). In a nutshell, the process is pretty simple:

1. Find a partner or form a small team.
2. Each of you choose a virtue to work on.
3. Take a close look at your virtue, and at any obstacles you feel when you try to practice it.
4. Work with your partner(s) to come up with exercises in which you will frequently, deliberately practice that virtue in ways that challenge your current level of fluency.
5. Check in with your partner(s) regularly to encourage each other and to keep each other accountable, and adjust your curriculum as you learn more about what works and what challenges you face.
6. When you feel you have integrated the virtue adequately into your character, start the process again with a new virtue.

Alas, after some initial promise, the group began to dwindle, and then the pandemic disrupted everything, and now as far as I know there are only two of us still working through the program on the regular. But in the course of researching, we dug up a lot of information about virtues in general and about particular virtues, and that's forming the basis for the posts I'm sharing here.

Notes on virtues

What I hope to do with these notes on virtues is to collect ideas that will be useful to people who want to improve in a certain virtue. This may include concrete advice about strengthening that virtue itself, and also some discussion about other virtues that are related in some way: maybe they're prerequisites, or harmonize in some way, or maybe there's some tension between them. I sometimes find it challenging to define the virtue precisely, or to distinguish it from another virtue—and sometimes the term for the virtue gets overloaded with a variety of meanings in common use—so I include discussion of those nuances too.

I'm aiming to be inclusive of a variety of useful perspectives, and of a variety of cultures, rather than to be definitive or dogmatic. It's a fuzzy subject matter to begin with. I'm feeling my way about, leaning on existing guides when I can find them (though I tend to find a lot more examples of people praising or advocating certain virtues than of people explaining them or giving practical advice on how to go about improving in them).

I take some inspiration from Aristotle, who, when he examined a set of virtues in his *Nicomachean Ethics*, started with virtue-concepts as already found in common language and folktale, rather than starting from a theoretical foundation and building ideal virtues from there. When it comes to dividing up a complex subject matter into manageable and coherent chunks, previous generations have already done a lot of the work for us and handed that down to us in the language and tropes we use. That we have found a word or trope useful is a good clue that there's some reasonably-helpful and worth-noticing regularity at the base of it. While this sort of understanding shouldn't be confused with the gospel truth of how reality is constituted, it seems wise to glean as much as we can from it before trying to systematize more deliberately.

One of the things I did was to investigate several virtue-based traditions (the Greek [cardinal virtues](#), the [traditional Christian virtues](#), the virtues of [Bushido](#), [Confucian virtues](#), the virtues of [Scouting](#), the [West Point virtues](#)), the virtues favored by some particular philosophers ([Aristotle](#), [Cicero](#), [Ben Franklin](#), [Ayn Rand](#), [Henry David Thoreau](#), [Shannon Vallor](#), [the Cynic philosophers](#), the developers of [care ethics](#), [William De Witt Hyde](#), [Eliezer Yudkowsky](#)), and the virtues identified as "[character strengths](#)" by psychologists operating in the positive psychology paradigm. This isn't comprehensive by any means, but it was revealing.

For one thing, there was a lot less consensus than I expected about which virtues are the important ones. This is somewhat complicated by problems of terminology. For example, what one philosopher will call self-control, another will call continence, another restraint, another discipline. Or, while Paul says that the greatest virtue is [love](#),^[4] he defines "love" in such a way^[5] that it incorporates patience, [kindness](#), *mudita*, modesty, humility, respect, good temper, forgiveness, righteousness, care, trust, hope, and perseverance. Or different cultures will partition virtue-space differently: *sisu* is only kind-of like perseverance; *mudita* is only kind-of like sympathy;

nying je is only kind-of like compassion. This can be challenging for works in translation, where the translator has chosen the closest equivalent English word, but a close reading reveals that the author meant something different from what we mean by that word.

I tried to correct for things like these. I consolidated various terms for similar virtues together, and created a spreadsheet where I could note which virtue-clusters had been promoted in which systems or by which philosophers. But of the hundreds of virtues I found, only six of them were on more than half of the lists:

- [courage/bravery](#)
- [honesty/trustworthiness/promise-keeping](#)
- [self-control/continence/restraint/discipline](#)
- [loyalty/fidelity](#)
- [compassion](#)
- [respect for others/thoughtfulness/consideration](#)

If you add those that were on exactly half of the lists, you also get [justice](#), [wisdom/philosophy](#), [sincerity/straightforwardness/earnestness/frankness](#), [industry/effort/enterprise/productiveness](#), [duty/responsibility/purposefulness](#), [piety/reverence](#), and [strength/toughness/vitality/health/fitness](#).

You may have heard that [patience](#) is a virtue, but it didn't make that cut. Neither did [humility](#), hope, perseverance, [courtesy](#), generosity, friendliness, creativity, [caution](#), cleanliness, mercy, forgiveness, wit, originality, calm, warmth, curiosity, hospitality, pride, or gratitude. Some lonely virtues like boldness, imagination, spontaneity, and playfulness appeared on only one list. Other skills that are often popularly admired—like being influential, having emotional intelligence, or being good in bed—weren't on any lists at all.

Some virtues are debatable. Selflessness, pride, altruism? The apostle Paul and Ayn Rand would disagree about what's the virtue and what's the vice. Virtues like [chastity](#), obedience, and patriotism give some of us the willies.

I'm aiming to be inclusive and to eventually give some attention also to these less-prominent and more controversial virtues.

Why am I going to this trouble?

My hope is that, whichever virtue you're hoping to improve, you'll be able to get a head start from the research and write-ups I've done.

I'm also motivated by self-improvement. I've been working to deliberately improve some of these virtues in my life, and I hope to make that an ongoing project, so putting together these virtue-dossiers helps me to lay the groundwork for this.

If we manage to reboot the Society of the Free & Easy in the post-pandemic time, these may help us hit the ground running.

I also have vague ideas about this being a worthwhile political project. I've come to distrust talk of elections and revolutions and institutional reforms. I think the longer, harder, more subtle project of helping people improve is a more reliable path to a better future than trying to impose wise policies on them from on high. If people

become braver, wiser, more just, and more honorable, public policy will follow their lead. If people become more cowardly, foolish, grasping, and disreputable, conniving politicians will lead them by the nose.

I'm sharing these at *LessWrong* in particular because I value the sort of insightful feedback people share here. Since I'm not an expert at any of the virtues I'm writing up, I'm slyly taking advantage of Cunningham's Law [6] to correct my misunderstandings about them.

Note

This is an ongoing research project. Whenever I dig up information about a particular virtue that I think can usefully add to what I've written, I will go back and revise the page associated with that virtue. Or if you leave a comment about something that I've gotten wrong or something important I've neglected to mention, I'll fix that. I think it's more important for these pages to be useful than for them to be a fixed record of what I thought when I first wrote them.

Tentative Sequence Outline

1. [Honesty](#) (truthfulness, veracity, trustworthiness, promise-keeping, probity, reputability)
2. [Courage](#) (bravery)
3. [Self Control](#) (continence, discipline)
4. [Respect for Others](#) (thoughtfulness, consideration)
5. [Piety](#) (reverence, faith, religion, spirituality, transcendence)
6. [Loyalty](#) (faithfulness, fidelity)
7. [Compassion](#)
8. [Wisdom](#) (philosophy)
9. [Temperance](#)
10. [Fitness](#) (health, fettle, hardiness, vitality, toughness)
11. [Sincerity](#) &c (straightforwardness, frankness, sincerity, earnestness, candor, *parrhēsia*)
12. [Justice](#) as a virtue (*epieikeia*)
13. [Industriousness](#) (assiduity, effort, enterprise, industry, productivity)
14. [Duty](#) (responsibility, purposefulness)
15. [Social Responsibility](#) (civility, community, citizenship, patriotism)
16. [Prudence](#) (practical wisdom, decision theory)
17. [Know-how](#) (practical knowledge, craft)
18. [Honor](#) (nobility)
19. [Moderation, Balance, & Harmony](#)
20. [Patience & Forbearance](#) (*kshanti*)
21. [Care](#)
22. [Attention](#) (awareness, discernment, mindfulness, presence, focus, concentration, alertness, observation, presence)
23. [Amiability](#) (friendliness, geniality, agreeableness, conviviality, affability, niceness, affection, warmth)
24. [Rationality](#) (reason, love of truth, clear-headedness, deduction/induction)
25. [Simplicity](#)
26. [Forgiveness](#) (mercy, clemency, *epieikeia*, placability, indulgence, leniency)
27. [Integrity](#) (unity, authenticity)

28. [Humility](#)
29. [Optimism, Hope, & Trust](#)
30. [Good Temper](#)
31. [Fairness](#) (impartiality)
32. [Endurance](#) (fortitude, hardness, toughness, stamina)
33. [Benevolence](#) (goodness, goodwill, *metta*)
34. [Ambition](#) (drive, aspiration, loftiness, initiative)
35. [Perseverance](#) (persistence, resilience, grit, fortitude, tenacity, *sisu*)
36. [Kindness](#) (including kindness to animals)
37. [Empathy](#) & Sympathy (*mudita*, pity)
38. [Frugality](#) (thrift, economy)
39. [Dignity](#) (self-respect, rectitude)
40. [Courtesy](#) (manners, politeness, consideration)
41. [Chastity](#)
42. [Love](#)
43. [Resolve](#) (determination, commitment, dedication, devotion) / Steadfastness (firmness, decisiveness, consistency)
44. [Caution](#) (carefulness, preparedness, foresight, deliberation, heedfulness, vigilance)
45. Cleanliness (orderliness, hygiene)
46. Gentleness (tenderness, mildness)
47. Propriety (decorum, etiquette, refinement, gravitas, sobriety, sportsmanship, chivalry)
48. Creativity (art, imagination, inventiveness, problem-solving)
49. Modesty
50. Altruism (philanthropy, generosity, charity, magnificence, magnanimity)
51. Awe (elevation, wonder, appreciation of beauty, sensitivity)
52. [Judgment & Righteous Anger](#)
53. [Gratitude](#) (appreciation, reciprocity)
54. Valor (heroism, defense-of-others, glory, chivalry)
55. Serenity (calm, equanimity, tranquility, composure, contentment, peacefulness, consolation)
56. [A Sense of Shame](#)
57. Purity (innocence)
58. Philomathy (studiousness, scholarship, seeking good advice)
59. Self-reliance (independent thinking, originality, independence, liberty)
60. Wit (humor, playfulness)
61. Helpfulness (service)
62. Filial Piety (and respect for elders)
63. Joy (cheer)
64. Flexibility (adaptability, agility, resourcefulness)
65. Cooperation (teamwork, coordination)
66. Taste (fashion sense, beauty)
67. Quiet (silence, stillness, listening)
68. Rhetoric (persuasion, conversational competence, expressiveness, influence)
69. Pride (self-worth, self-esteem)
70. Poise (confidence, grace, unflappability, authority, gravitas, refinement)
71. Perspective (prioritization)
72. Openness (candor)
73. Obedience (respect for authority, submission)
74. Mentoring (counsel, nurturance)
75. Intelligence (ingenuity, mental quickness)
76. Hospitality
77. Curiosity (exploration, inquisitiveness)

78. Connectedness (social intelligence)
79. Concern
80. Zest (exuberance, vigor, enthusiasm, wholeheartedness, zeal, *virya*)
81. Liberality (tolerance, understanding)
82. Leadership
83. Graciousness (accepting fault, accepting help)
84. Tact (discretion)
85. ...more as I discover them...

What am I missing?

1. ^

Simone de Beauvoir, *The Ethics of Ambiguity* (1947)

2. ^

Aristotle, *Nicomachean Ethics*, book II, chapter 2

3. ^

Benjamin Franklin, *Autobiography*, chapter 9 ("Plan for Attaining Moral Perfection")

4. ^

[1 Corinthians 13:13](#) "And now abide faith, hope, love, these three; but the greatest of these *is* love."

5. ^

[1 Corinthians 13:4-7](#) "Love suffers long *and* is kind; love does not envy; love does not parade itself, is not puffed up; does not behave rudely, does not seek its own, is not provoked, thinks no evil; does not rejoice in iniquity, but rejoices in the truth; bears all things, believes all things, hopes all things, endures all things."

6. ^

"The best way to get the right answer on the internet is not to ask a question; it's to post the wrong answer." ([Named after](#), but not formulated by, [Ward Cunningham](#).)

Notes on Honesty

This post examines the virtue of **honesty** (a.k.a. “truthfulness,” “veracity”). This could be a starting point for expanding [the LessWrong Wiki entry on Honesty](#), and I encourage you add comments/questions to help guide that effort. (See also: [Deception](#) and [Meta-Honesty](#))

Honestly ;-) this was a difficult post for me to put together. The topic is complex and it was hard to find the sweet spot between being too wordy and too superficial. People have written dense books to try to explain things I cheekily tried to summarize in a sentence or two.

Honesty and rationality

Much of LW concerns how we can better approach *knowing* the truth. Honesty concerns an aspect of how we *communicate* that truth. For this reason I think of it as a social virtue rather than an intellectual virtue. Sometimes, however, you hear expressions like “being honest with yourself” used to describe intellectual virtues.

Honesty requires at least a minimum exercise of the intellectual virtues. If you do not exercise epistemological due diligence before you communicate your understanding of the state of the world, you may be telling the truth as you see it, but you are failing to respect the virtue of honesty by not taking enough care to distinguish the false from the true.

For example, an acquaintance of mine is very into the woo. When she speaks to me of woo things that she thinks are important for me to learn about, I don’t get the impression that she’s *lying* to me, exactly, but she exercises such poor judgement about what to believe that I end up feeling as though I have been lied to and I tend to look upon her as being as much a dishonest person as a foolish one.

On the other hand, Scott Alexander warns against [“lie inflation”^{\[1\]}](#) in which we might accuse people of lying when they are merely honestly representing the poor results of their sloppy thinking. (His argument is something like this: if we expand our definition of dishonesty to cover mistaken sloppy thinking then we risk losing the ability to talk more precisely about deliberately deceptive dishonesty, and this lowers our defenses against just such liars.) See [“Maybe Lying Doesn’t Exist”](#) for more discussion of this point, and [“Rationalizing and Sitting Bolt Upright in Alarm”](#) for a plea that we invent some epithet as strong and undesirable as “liar” to apply to people who, though they are not being dishonest, are recklessly misreasoning.

Related virtues

There are some other virtues that are closely-related to honesty:

- If you are both honest and have cultivated a reputation for honesty, you have the virtue of “reputability” or “trustworthiness”.
- Being honest about your [commitments](#) to future action is sometimes called “accountability”, “reliability”, “[dependability](#)”, or “conscientiousness”.
- If you are not merely honest about what you are asked or what you say, but are also proactive about seeking out and expressing what the person you are communicating with needs to know, and/or are unreserved about which truths you express, you are exercising “candor” or “frankness” — or [parrhēsia](#) if you’re a Cynic.
- A subset of honesty involves meaning what you say about your own motives, desires, and opinions — “[sincerity](#)”. You can express sincerity in deeds as well as words, in which case you are being “earnest” (or sometimes “authentic”).

- If you don't beat around the bush or candy-coat your honesty or engage in euphemism, false modesty, or things of that sort, you are being "straightforward".

Defining honesty

Honesty, as a virtue, is the *habit* of being honest — being honest by default, characteristically. [But it can be surprisingly tricky to define in a water-tight way what being honest consists of.](#)^[2] Is it honest to express something false if you think it is true? to express something true if you think it is false? Is it honest to express something true that you know to be true if you also believe that by expressing it you will make the person to whom you express it believe something false?

The "wizard's oath" (never saying anything that is literally false) has been proposed as an attractive approximation for honesty. On closer inspection it "[is of limited practical utility, given the ubiquity of other kinds of deception.](#)" It is certainly not honest to deliberately create a false impression by means of selective or craftily-worded true statements. But the game of "[technically" telling the truth](#)" ("I never actually lied") while deceiving people is a popular one.

To make things yet more difficult, people sometimes say things that they know to be false but with the intention of helping someone understand the truth. Consider, for example, the oversimplifications of physics that we learn in high school or Physics 101 classes. These oversimplifications are falsehoods of a sort, and the professors know this, but it would be too complex for the student to get the whole truth all at once, so placeholder-lies are used in the service of truth.

If you think about this sort of thing deeply enough you may start to go down the [rabbit hole](#) of wondering how *any* communication or even any conceptualization can be a true one: how does language work, what is "meaning" and "representation", what is the relationship between the map and the territory, and so forth.

Bullshit

The philosopher Harry Frankfurt wrote the essay "[On Bullshit](#)" about a particular variety of dishonesty.^[3]

What is "bullshit"? Essentially it is *indifference* to the truth of one's communication: a sort of bluff in which one goes through the motions of conveying something factual without actually attending to what it takes to represent the truth. It differs from "lies" in part in that the liar *does* care about what the facts are (in order to misrepresent them advantageously). For this reason also, Frankfurt believes that bullshit is more dangerous to the enterprise of truth-seeking. It also seems to have become especially prominent lately.

Contexts designed to contain dishonesty safely

There are certain contexts in which saying things that aren't true is arguably unobjectionable. For example, if you are telling a joke, although you are describing a scenario in a way that superficially resembles how you would describe something that really happened, nobody with a mature understanding of conversation expects that you are expressing something that is literally true. If you are an actor delivering a soliloquy, it requires a willing suspension of disbelief — a sort of mutual agreement to temporarily play in a dishonest sandbox — for the audience to think that you are speaking sincerely of your own thoughts.

That particular convention, however, creates a grey area in which actually deceptive dishonesty thrives. For instance, consider an actor in a lab coat in [an advertisement](#) who tells you that Miraclon-X™ has transformed the lives of her patients. That actor is doing something superficially no more dishonest than what Laurence Olivier did when he complained about his uncle shacking up with his mom in *Hamlet*. But the advertisement has been crafted to give the impression that an actual medical doctor with expertise in a certain area has become enthusiastic about the efficacy of a certain drug. It is not meant to be seen as fiction, but is meant to leave a deceptive impression.

There are other ways in which we have carved out areas in which dishonesty is considered acceptable or even appropriate. “Fish stories” for example, in which the point of telling a true-ish story is to entertain or pass the time rather than to convey information — in such stories, exaggerations or even falsehoods that make the story more entertaining are often considered to be unobjectionable or even admirable.

There are [common verbal handshakes](#) (“How you doin’?” “Fine.”) that have the superficial appearance of a conveyance of information but that mostly have a different purpose and in which people rarely concern themselves with honesty.

If I am reading someone’s résumé, I consider them to have stepped over the line if they actually lie about having some degree, or having worked for some company, or something like that. But on the other hand, I expect that their résumé is them “putting their best foot forward.” I don’t expect them to be fully candid about their flaws, foibles, and failures but to selectively present attractive parts of their histories.

Another way we carve out space for dishonesty is to use verbal envelopes in which we put untrue things such that they are insulated from being taken as literally true: “Once upon a time...”, “Consider this counterfactual scenario...”, “Imagine for a moment that...”, “If we were to assume X...”, and so forth.

People also use irony and sarcasm, understatement and hyperbole, parody and caricature, modest proposals and other such rhetorical devices to say what is not literally true but what also is not designed to be interpreted as literal truth. These techniques expand our expressive and communicative repertoire but at the cost of playing fast and loose with truth-telling. I mention things like these (and this is not meant to be an exhaustive list) because sometimes they are overlooked by people who recommend a literal and thorough honesty.

Collateral damage of “harmless” dishonesty

In addition to these sorts of spaces in which tales can be told in a way that warns the recipients that they are likely to be a shade or two off from the truth, lies may also be *justified* in a more ad hoc way but one that preserves their status as actual dishonesty.

Most tempting to consequentialists are lies that are justified by their service to what appears to be a greater good: Giving a placebo to a patient in place of a missing pain reliever, but telling them it’s the real deal in the hopes of making the placebo work better. Telling the Gestapo that actually Anne Frank moved out of the attic some time ago and didn’t leave a forwarding address. Not telling your spouse about that hook-up you had on that business trip because you think the truth would be more harmful to your marriage than a lie would be. Recognizing when “does this make me look fat?” is not a request for an honest and literal evaluation of the question. [“Little white lies”](#) that protect someone’s feelings.

Sissela Bok argued, in her book *Lying: Moral Choice in Public and Private Life*, that when people justify their dishonesty, they tend to focus too narrowly — in their little white lies and fish stories as well as their big important utilitarian deceptions.^[4] If we restrict ourselves to looking at the immediate effects of the lie, and on only the harm done by the lie to the person or persons lied to, we’re missing some important terms in the calculation. Among the

other things we should consider are how the lie affects the character of the liar (eroding that person's habits of truth-telling, or reinforcing their habits of lying), how the lie might influence bystanders (by communicating something about the culture of truth-telling they belong to), and how the lie might erode future trust between the parties.

"Fish stories," for example, can reinforce habits of embellishment that spill over into other contexts where they aren't as harmless — ["we're not talking about grotesque major falsehoods — but the first words off my tongue sometimes shade reality, twist events just a little toward the way they should have happened..."](#) If we develop the habit when we open our mouths of seeking for words that make us interesting to hear or that meet people's expectations, without having even more regard for their veracity, it can be easy to slip into that habit when we're not intending to lie but aren't paying close attention. And it can take a lot of attention to hit the truth precisely; it is so much easier to just aim in its general direction.

"Falsehood is so easy, truth so difficult.... Examine your words well, and you will find that even when you have no motive to be false, it is a very hard thing to say the exact truth, even about your own immediate feelings — much harder than to say something fine about them which is not the exact truth." —George Eliot^[5]

Here's another example of how innocent-seeming lies may cause collateral damage: imagine that in your culture friends will always answer "you look beautiful in that" because that's what is expected of friends. You may as a result always be a little insecure about whether you look good or not because you cannot expect to get accurate feedback from your friends. This convention, rather than bolstering your self-esteem, instead leaves you in doubt. And if a friend *is* honest with you and tells you your clothing is not flattering today, the convention may reasonably lead you to doubt their friendship! What if you [want to have friends you don't need to second-guess](#) instead? So even in the case of a little white lie, there are broader consequences to consider.

Honest listening

Thoreau wrote: "It takes two to speak the truth—one to speak, and another to hear."^[6] In order for honest communication to take place, the speaker must speak honestly, and the listener must be prepared to hear honest speech. This means that some groundwork must have been established to indicate that the speech is meant to be honest, and not to be serving some other purpose. This might include trust between the parties, or a cultural expectation of honest communication.

The listener can help by signaling that honesty is what they expect and are prepared for. This can be done explicitly: "I'm not trolling for complements here. What did you think of my paper?" "Don't candy-coat it; how bad is the cancer?" You might [invoke "Crocker's Rules"](#) in which you explicitly request the speaker to prioritize honesty over other concerns like tact or respect for taboos. Certainly if you want people to be honest with you, you should not disincentivize them by punishing them for doing so.

The power of honesty

Some philosophers claim to have discovered unexpected powers in the virtue of Honesty. Tolstoy, for example, thought that honesty was revolutionary. "No feats of heroism are needed to achieve the greatest and most important changes in the existence of humanity.... it is only needful that each individual should say what he really feels or thinks, or at least that he should not say what he does not think." He was followed in this by Solzhenitsyn (["Live Not By Lies"](#)) and Václav Havel (["The Power of the Powerless"](#)):

In the post-totalitarian system... living within the truth has more than a mere existential dimension (returning humanity to its inherent nature), or a noetic dimension (revealing reality as it is), or a moral dimension (setting an example for others). It also has an unambiguous political dimension. If the main pillar of the system is living a lie, then it is not surprising that the fundamental threat to it is living the truth.^[7]

William Wollaston — a now mostly-forgotten philosopher who was important in the early enlightenment — defended a system in which all religion ultimately reduced to ethics and all ethics reduced to honesty (*The Religion of Nature Delineated*). “[E]very intelligent, active, and free being should so behave himself, as by no act to contradict truth; ...treat every thing as being what it is.”^[8] All else would follow from that.

When honesty is dangerous

It can be dangerous to tell the truth when the truth is unwelcome to someone who is hostile when angered. This can suppress truth-telling of certain sorts, particularly when the hostile are organized into powerful institutions. Truths that are blasphemous, unpatriotic, or in other ways challenge the sacrosanct or privileged can be difficult to utter safely. “Cancel culture” and the shifting and unpredictable taboos that may drive its piranhas into a frenzy can frighten people into being less candid or into saying things they do not believe to be true.

Strategies for truth-telling in such circumstances are varied. You might simply stay silent, picking your battles and being truthful elsewhere, accepting the passive falsehood of silence. You might use euphemism or other circumlocution to approach the truth at a safe distance without breaching the taboo. You might surround the truth so that its outline is clearly visible without stating it explicitly: engraving the first two lines of the syllogism in ten-foot-high letters and leaving the last line blank as an exercise for the reader.

Another strategy is *esoteric communication*, which means to tell the truth in a sort of encoded or between-the-lines way that will only be obvious to people who are not among those who will be enraged by it.

Kant

Kant made the duty of honesty a sort of showpiece of his impressive moral theory. I am no expert on Kant and will probably be oversimplifying dangerously here, so please take this with that caveat attached.

Kant asserted that lying was always wrong; that you always have a duty to be truthful. One of his arguments was that dishonesty could not be defended coherently. This is why: Say you believe that a person is justified in deceiving others because it can be useful to a person to deceive other people. But if everyone were deceitful whenever they thought fooling someone might be helpful, then no expectation of honesty would survive. Everyone would just assume that everyone else may be lying at any time and that therefore there is no good reason to believe them. But if that is the result, then dishonesty will lose any advantages it might have had — it will no longer be helpful to you — nobody can be usefully deceived by your lies because nobody puts any stock in them in the first place. So in this way the supposed justification for dishonesty undermines itself.

Kant’s recommendation of absolute honesty without exceptions is a tempting target to consequentialists who oppose the Kantian moral system. Imagine you have seen where someone is hiding from a crazed murderer who is intent on killing them — are you obligated to refrain from lying to the murderer who asks you where to find them? Kant’s answer to this objection (on first approximation: “yes”) was not very convincing, and has been subject to scorn and ridicule ever since ([perhaps unfairly](#)^[9]).

Philosophers sometimes suggest that contra Kant, there is a “universalizable maxim” that can encompass some dishonesty, if you just draw it up skillfully. Something like: “Be honest except in rare and extreme situations in which a [reasonable](#) person would conclude that with a high likelihood dishonesty would clearly result in a much better outcome...” for some to-be-determined values of reasonable, high, clear, and much better. (See Eliezer Yudkowsky’s discussion of [meta-honesty](#) for one possible formula.)

It occurs to me also that Kant’s objection lacks empirical support. Case in point: President Trump is notoriously dishonest. And rather than trying to project an air of reputability — trying to remain “technically true” and saving his lies for when they’re really necessary, as a more mundane politician would — he just starts making stuff up from word #1 and never bothers to touch down in the land of truth. Kant would predict that this would mean people would stop trying to evaluate his utterances for their meaning or to evaluate them against a standard of truth: that his lies would lose their effectiveness. But, four years into his presidency, the *New York Times* assigned two reporters to go line-by-line through one of his 87-minute campaign speeches to highlight the falsifiable statements that were also false ones. [\[10\]](#) “Trump said X. Is that really true?” remains a common headline and debates about the truthiness of whatever he said today continue to rattle the wires across the land.

Glomar response

The [“Glomar response”](#) (prototypically: “I can neither confirm nor deny X”) is a way of preserving reputability while concealing information that you do not want to be honest about: you refuse to answer rather than lying.

If you refuse to answer only on those occasions when you do not want to tell the truth, this can have the effect of leaking the very information you want to hide, as it may be obvious to your interrogator which answers you would want to conceal, and your interrogator can just fill in the Glomar blank with the more embarrassing of the answers. So for the Glomar response to work as designed, you must also deploy it on occasions when you would *not* ordinarily be bothered by speaking the honest truth. This means that you have to anticipate well ahead of time which class of questions might at some future time require an answer that you would prefer not to give truthfully, and begin Glomarizing them now. This is [difficult to pull off as a general practice](#).

Sometimes people will suggest a Glomar-style response as a possible solution to the murderer-seeking-the-hidden-victim thought experiment. When asked “Where is that no-good so-and-so I want to kill?” you can honestly answer “I feel no inclination to share information of that nature with you.”

Radical honesty

Description pilfered from the LessWrong wiki:

[Radical honesty](#) is a communication technique proposed by Brad Blanton in which discussion partners are not permitted to lie or deceive at all. Rather than being designed to enhance group epistemic rationality, radical honesty is designed to reduce stress and remove the layers of deceit that burden much of discourse.

The Radical Honesty technique includes having practitioners state their feelings bluntly and directly, even if it may be in a way typically considered impolite. Avoiding all “white lying” is said to lead to a more truthful relationship with themselves and others.

- See also: [“Radical Honesty”](#) by Eliezer Yudkowsky.

According to its proponents, radical honesty has many benefits: It allows you to develop a deeper connection with other people rather than only connecting on a falsification-to-falsification basis. By atrophying the lie-telling part of your brain, it allows you to see the world more accurately. It shows respect for the people you communicate with. It promotes an ethos of honesty in society. In a sort of Zen way, it takes you out of the stories you tell yourself and back into the real world. And it cures insomnia & sexual dysfunction, saves marriages, makes you rich, and other pop-psychology marketing points of that sort.

How do you go about being *radically* honest? Here's the program in a nutshell: Train yourself to accept honesty at all times — permit yourself to see the truth even when it hurts. Learn the ways you deceive yourself — cognitive biases, illusions, logical fallacies, bad habits, etc. — and retrain yourself. Observe yourself lying, ask yourself what occasions it, look more deeply and explicitly at what motivates you to lie and try to find a more honest way to deal with such situations that also meets your needs. Confess your dishonesty to others when you catch yourself ("I just said how much I'd love to come to your book club, but actually it doesn't sound like the sort of thing that would hold my interest and I wasn't being honest about my level of enthusiasm.") Start by saying what you notice, then how it makes you feel, and only much later what you think about it. Be a detective of cultural dishonesty: euphemisms, things that "aren't talked about," the elephant in the room, things we do to avoid discussing certain topics, myths we take for granted, political correctness, taboos, etc.

Honesty about future commitments

It is a valuable thing to be able to make credible promises about your future actions. It is an honest thing, having made such promises, to follow through on them. But such promises are problematic because the future is uncertain. You may die before you are able to fulfill your promise, or fate may interfere in some other way that makes it impossible for you to do as you said you would or that causes an unanticipated conflict between incompatible commitments.

Perhaps it is best to assume by convention that all such promises have a "but for unforeseen accidents of fate" clause attached to them. However it can be tempting, when fulfilling a promise turns out to be harder or more costly than anticipated, to claim that this nullifies our promise under such an implicit clause. If we give in to such a temptation, we will not be honoring the virtue of honesty, and so this should make us [cautious](#) when we make promises about our future acts.

Open-ended promises are especially bad this way: Oaths of undying loyalty, pledges to obey future orders without specifying the content of those orders, and so forth. Something like 45% of marriages in the United States end in divorce, and most of them begin with a pledge to stay married "until death do us part."

Oaths, pledges, and formal declarations of intent to honesty

Marriage vows are an example of a formal pledge, which, from its solemn and traditional format and context is supposed to give greater weight to the vows made. Such deliberate attention to phrasing and ritual context gives the pledger notice that they ought to carefully consider their vow and its implications. Other examples include oaths of office, pledges of allegiance, swearings-in of witnesses, and so forth. People may more informally try to emphasize how much they intend to honor their promises by means of pinky-swears, "may God strike me down if I'm lying", "I swear on the name of my mother," and other such formulas.

None of these rituals seems to be reliable in compelling truth-telling or promise-keeping. They also imply that the promises people make in the absence of such rituals are not to be taken as seriously, which strikes me as potentially erosive of any general tendency to truth-telling.

Some Christians for this reason eschew oath-giving, citing the advice in *James* 5:12 ("But above all things, my brethren, swear not, neither by heaven, neither by the earth, neither by any other oath: but let your yea be yea; and your nay, nay; lest ye fall into condemnation.")

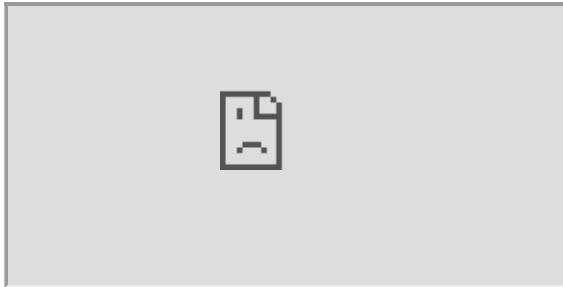
Contracts are another matter. They typically enforce promises and make them more reliable by providing for sanctions against promise-breaking. This then tends to leave the arena of honesty for the arena of material incentives.

Self-representation

In the *Nicomachean Ethics*, Aristotle discusses the virtue of honesty, but mostly in the context of self-representation:^[11] Do people exaggerate their skills and experiences and credentials, do they with false modesty self-deprecate, or do they straightforwardly share an honest assessment of themselves?

People have long used disguises, imposture, cosmetics, titles, credentials, and so forth to try to make an impression. But now that more of our social presence is digitally mediated, this has presented new possibilities — things like pseudonymous sock-puppetry, auto-suggested phrasing, catfishing, video filters, purchased “likes”, and so forth.

We can be more deliberate and selective about what we choose to share about ourselves on a social media outlet than we could be in a more free-form, face-to-face, real-world social context. If people are incentivized to filter how they project their lives in a certain way — showing happy and exceptional moments and suppressing sad and mundane ones, for example — this creates a misleading impression, and maybe a harmful one if people compare their own real lives to the filtered lives of those in their virtually-social circle.



The pressure to “brand” your internet presence threatens authenticity. If an increasing portion of what you are to other people is your mediated projection of yourself, and that projection represents your brand rather than your authentic person, do you run the risk of atrophying your authentic social self or making it subservient to your brand? Might you wake up one day to discover that your brand is well-loved but you are still lonely, for instance?

Social media culture is still fairly new and very much in flux. The choices we make today about how honestly to represent ourselves on-line will help set a standard that will have long-lasting influence.



Agnes Callard
@AgnesCallard

It takes time to learn that one can "just be a person" on social media--without "performing" so much--& until we all learn that, together, we're in sthg analogous to the "high school" phase--but I think eventually there will be a tipping point and we'll move into the adult phase.

10:18 AM · Nov 18, 2020 ·

Spin/Framing

Spin or Framing is the attempt to fit revealed or asserted facts into a rhetorical framework in such a way that they will lead people to desired conclusions or away from undesirable ones. Often this takes the form of seductively modeling the desired variety of motivated reasoning.

When this is called "spin" it usually implies purposeful dishonesty; when it is called "framing" its proponents sometimes claim it can be done in the service of clarity and honesty, or to defend against spin. (Impartial rationality, or our best approximations to it, might be considered as one variety of frame from within which we can draw conclusions.)

How to become more honest

One prong of becoming more honest has to do with overcoming ways of thinking that prevent one from approaching the truth in the first place. Learn to love truth and refuse to fear it. All of that is well-covered elsewhere on LW. The second prong is about *communicating* honestly, and that's what I'll concentrate on in this section.

One way to become more honest is to get in the habit of speaking truthfully even when it seems harmless to do otherwise. Tolstoy put it this way:

To tell the truth is the same as to be a good tailor, or to be a good farmer, or to write beautifully. To be good at any activity requires practice: no matter how hard you try, you cannot do naturally what you have not done repeatedly. In order to get accustomed to speaking the truth, you should tell only the truth, even in the smallest of things. [\[12\]](#)

(Tolstoy was an author of *fiction*, however, so he must have allowed for some safely-compartmentalized uses of untruth, unless perhaps he had a late-life change of heart about fiction.)

Another way to become more honest is to exert more effort and attention toward the details of your communication. Lazy inattention can lead to falsehood when we sleepwalk through something we're saying — spitting out phrases because the words sound right together because we've heard them that way before.

"Clichés, stock phrases, adherence to conventional, standardized codes of expression and conduct have the socially recognized function of protecting us against reality, that is,

against the claim on our thinking attention that all events and facts make by virtue of their existence." —Hannah Arendt^[13]

Although lies are often described as being delivered in fancy, honey-coated language, honesty may also be best served by sophisticated use of language: originality of metaphor, inventiveness in phrasing, a poet's attention to precise word-use, a rich vocabulary, and a critical eye for ambiguity.

Avoid the use of "[weasel words](#)" that give the appearance of saying something truthful while actually throwing up an obscuring cloud of syllables around an absence of truth.

Finally, in spite of our best efforts we may nonetheless give an unintentionally mistaken impression. It can be valuable to check in with those we are communicating with to ask them to verify that their understanding of what we are trying to say matches what we mean to communicate. Adding "checksums" of some sort to our communication may enhance its integrity.

1. ^

Scott Alexander, ["Against Lie Inflation"](#) *Slate Star Codex* 16 July 2019

2. ^

James Edwin Mahon, ["The Definition of Lying and Deception"](#) *Stanford Encyclopedia of Philosophy*

3. ^

Harry Frankfurt, "On Bullshit" *Raritan Quarterly Review*, Fall 1986

4. ^

Sissela Bok, *Lying: Moral Choice in Public and Private Life* (1978)

5. ^

George Eliot, *Adam Bede*, book II, chapter 17

6. ^

H.D. Thoreau, *A Week on the Concord and Merrimack Rivers* (1849)

7. ^

Václav Havel, "The Power of the Powerless" (1978)

8. ^

William Wollaston, [The Religion of Nature Delineated](#) (1722)

9. ^

Helga Varden, ["Kant and Lying to the Murderer at the Door... One More Time: Kant's Legal Philosophy and Lies to Murderers and Nazis"](#) *Journal of Social Philosophy* 18 November 2010

10. ^

Linda Qiu & Michael D. Shear, "[Rallies Are the Core of Trump's Campaign, and a Font of Lies and Misinformation](#)" *New York Times* 26 October 2020

11. ^

Aristotle, *Nicomachean Ethics* book IV, chapter 7.

12. ^

Leo Tolstoy, *A Calendar of Wisdom* (February 24), ~1910

13. ^

Hannah Arendt, "Thinking" *The Life of the Mind* (1971)

Notes on Courage

This post examines the virtue of **courage** and explores some avenues for how to improve it. This could be a starting point for expanding [the LessWrong Wiki entry on Courage](#), and I encourage you add comments/questions to help guide that effort.

Courage (sometimes “bravery” or the closely-related virtue of “valor”) is one of the most frequently-mentioned virtues in virtue-oriented traditions. It was one of [the four cardinal virtues](#) of ancient Greece, for example.

Courage is also often recommended as something that undergirds other virtues. C.S. Lewis wrote, for instance, that “Courage is not simply one of the virtues but the form of every virtue at the testing point, which means at the point of highest reality.”^[1] And Maya Angelou said that “Courage is the most important of all the virtues because without courage, you can’t practice any other virtue consistently.”^[2]

Fear

Courage has to do with our response to fear. This response has at least three components:

1. One concerns the way we judge how threatening a situation is — how easily spooked we are (emotional) and how sensible our risk assessment is (cognitive).
2. Another is how we act when we are immediately confronted with a frightening scenario — how well we think and perform while afraid.
3. The third is how we respond to the possibility of being in a fearful scenario at some future time (sometimes “fear” in this anticipatory context is called “anxiety,” “worry,” or “dread”) — whether our risk-aversion is well-honed or whether we are overly risk averse because we “fear fear itself.”

Fear is an unpleasant good in the same sort of way that pain and nausea are: Such things are no fun, but they are useful. Fear (when it is operating properly) informs you that you have managed to put yourself in a situation in which you run the risk of harm, and the unpleasantness of the sensation of fear prompts you to be averse to doing it again. Fear also can prepare you for an immediate, protective fight-or-flight response.

(Although we are averse to fear, we sometimes also perversely seek it out. In a similar way perhaps to how some people crave the pain of ghost chilies or spankings; some people crave the fright of horror movies and roller-coasters. Is this perhaps a way of helping to regulate our fear response through practice or inoculation?)

The visceral fear response is adaptive and it’s no surprise that we see it in other animals and that it seems to be to some extent a “deep,” subconscious part of our mental make-up. This can also make our fears difficult to work with on a conscious, rational level, as the experiences of people with phobias, panic disorders, anxiety disorders, or post-traumatic stress show.

Is there One Courage or Many?

In a lot of my reading from virtue-based traditions, courage is exemplified by the bravery of the warrior in battle. Aristotle, for example, started there and then generalized this to courage in the face of other deliberate human-caused threats, but he was reluctant to go further and say that the courage of someone who behaves bravely when threatened with disease or impoverishment was quite in the same ballpark.^[3] Nowadays we're more likely to recognize a variety of fears as being things we need courage to confront: fear of rejection, fear of mortality, fear of humiliation, fear of standing out, and so forth. We may speak of the "intellectual courage" it takes to resist the temptation to sweep an inconvenient truth under the rug, or the "moral courage" it takes to stand up for what you know is right in the face of social disapproval.^[4]

But it may be that when you stretch the word courage to cover so much territory, you are no longer describing a single virtue. When I was putting together this post I saw [this tweet from Zach Weinersmith](#) (of [SMBc comics](#) fame) who has been researching the history of space exploration: "For the space book, I am reading about people in extreme environments. Interesting thing: bravery is not cross contextual. You can be a brave mountaineer and still not brave at social situations."

Counterfeit Courage

In addition to the more common failure of cowardice, our response to fear can also fail in the opposite direction. There are brain disorders that can disable the ability to feel fear viscerally,^[5] thus throwing you back on mere conscious evaluation. Alcohol use is notorious for inducing temporary YOLO-recklessness and failure to recognize and respond appropriately to danger. Aristotle for this reason put the courageous "golden mean" at a mid-point between the vice of over-sensitivity to fear (cowardice) and the vice of under-sensitivity (rashness).^[3]

People without real courage will often try to counterfeit courage in social situations ("braggadocio"), as Shakespeare so vividly put it:

How many cowards, whose hearts are all as false
as stairs of sand, wear yet upon their chins
the beards of Hercules and frowning Mars,
who inward searched, have livers white as milk;
and these assume but valour's excrement
to render them redoubted.^[6]

Another form of counterfeit courage is exhibited by someone who is forced to choose between fearful things — a soldier who seems brave in battle only because he fears being shot for desertion or being disgraced in his community, for instance. Sometimes people suggest "hacks" for changing behaviors that seem to rely on this sort of thing (e.g. set up an artificial scenario in which if you fail to do frightening thing X, \$100 will be donated in your name to something you would be horrified to be associated with).

There is also a way of resolving fear that mostly side-steps the issue of cowardice or courage: that is, to make the fearful situation less fearful. One way to do this is to increase your competence. So for example if you have a fear of public speaking, you might participate in [Toastmasters](#), which is designed to create a non-threatening environment in which to practice a variety of public speaking skills. As your abilities

improve, so does your confidence, and what was fear-inducing no longer is. This in a way is another form of counterfeit courage (Aristotle said, for example, that in a storm, sailors were not exhibiting more courage than their frightened passengers, but merely a better handle on the situation).^[3] On the other hand, it is a way of meeting a frightening situation head-on and proving your mastery over it, which strikes me as something that could be a helpful way of bolstering courage.

Becoming Courageous

Zach Weinersmith, in that tweet above, cited the book *Extreme: Why some people thrive at the limits* by Emma Barrett and Paul Martin. Barrett & Martin conclude that “We all have a greater capacity to be brave than we sometimes appreciate” and identify three elements of the fear response — “physiological, cognitive, and behavioral” — each of which comes with handles we can learn to manipulate in order to take more conscious control over how we respond to fear and thereby develop more courage:

1. If you are aware and observant of your physiological response to fear, you can (once the initial shock passes, perhaps) take conscious steps to regulate it rather than just reacting to it or letting it take the reins. This implicates the additional virtues of [mindfulness](#) and emotional intelligence.
2. If you assess risk more [rationally](#), you will save your anxiety for situations that deserve it.
3. And with deliberation and practice, you can adjust how you respond while in fearful or anxiety-provoking scenarios.

Another suggestion, and again this comes from Aristotle, is to try to look on courage as a valuable end in itself and not just as something instrumental. In other words, rather than just saying “I wish I were more courageous, for then I could do scary things like X, Y, & Z, which I value” say also “and furthermore I would exhibit courage, which I *also* value.” This may improve the motivation you have for being courageous, and increase the pleasure you feel from your courageous acts (and therefore the reward you receive).

In Christopher Peterson's and Martin E.P. Seligman's *Character Strengths and Virtues*, they review the literature on courage and conclude:

Bravery can be promoted by practice (moral habit), by example (modeling), and by developing certain attributes of the individual (self-confidence) or group (cohesion).^[7]

They also summarized the not particularly well-tested, but intuitively appealing pop-psychology approaches to improving courage (e.g. *Awaken the Giant Within*) in this way:

This set of ideas... [builds] on a physiological, habitual, and attitudinal approach to cultivating bravery. Physiologically, people are encouraged to find a sense of courageousness within their body, and to use classical conditioning to associate some movement with the bodily sensation of power. Habitually, people are encouraged to become aware of their language and thought patterns and to break the ones that are especially limiting. Attitudinally, people are encouraged to engage in imagination and visualization exercises that help support a valorous disposition and help them with emotion regulation.^[7]

Other virtues may come to the assistance of courage. For instance if you have more [optimism](#), you may be more brave because the positive potential consequences of your bravery are more salient than they would be otherwise. If you have better [endurance](#), that may help you put up with fear, or may give you more confidence that you can get through the worst of whatever fearful thing you are up against. If you have more [loyalty](#), [honor](#), or [duty](#), such things may add to the value of your courageousness or the costs of your cowardice, and so may lead indirectly to bravery. Better [self-control](#) may help you regulate your response to fear so that it does not immediately carry you away.

Additional Resources

If you're fond of audio/visual learning, there are a couple of nice short videos out there: [How to stop feeling scared all the time](#) from *School of Life*, which concerns how to short-circuit excess anxiety, and [How to stop being a coward](#) from *Academy of Ideas*, which is a bit more on the philosophical side.

The Getselfhelp.co.uk site has [some worksheets and suggestions](#) you can use if anxiety is causing you to avoid situations that would be beneficial to you. Skills You Need has [a page on courage](#).

1. ^

C.S. Lewis, *The Screwtape Letters* (1942), letter XXIX

2. ^

Maya Angelou, *Meeting Dr. Du Bois* (audio interview by Krista Tippett, 2014)

3. ^

Aristotle, *Nicomachean Ethics*, book III, chapter 6

4. ^

See, for example, Rushworth M. Kidder, [Moral Courage](#) (2006)

5. ^

See, for example, Marissa Fessenden, ["This Woman Can't Feel Fear"](#) *Smithsonian*, 21 January 2015

6. ^

William Shakespeare, *The Merchant of Venice* (Bassanio speaking, Act 3, Scene 2)

7. ^

Christopher Peterson & Martin E.P. Seligman, *Character Strengths and Virtues* (2004), pp. 221, 226

Notes on Self Control

This post examines the virtue of **self control**. It is meant mostly as a summary of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is self control?

In another post, I covered the virtue of [temperance](#), which involves having a well-chosen and well-regulated set of desires. Self control is related but different. When an intemperate person gives into an unwise desire, they do so because they don't see anything wrong with it; when a person without self control does so, they do so in spite of knowing it's wrong. Another way of putting this is that intemperance is a problem in how you exercise the will, while lack of self control is a failure or malfunction of the will. Aristotle suggested this analogy: The intemperate person is like a city with bad laws; the person without self control is like a city that has good laws on the books but that doesn't enforce them.^[1]

Also, while temperance mostly has to do with our *desires*, self-control also comes into play in how we regulate our responses to other things, such as anger. (Lack of self control in the face of fear falls under the domain of [courage](#). Sometimes people split off response to anger under a distinct virtue like "[good temper](#).")

Often, particularly in older writing on the subject, you will see the word "continence" used for this virtue, but nowadays that word has become so linked to advertising for adult diapers that it's less common. "Willpower" is another common synonym. "[Akrasia](#)" (or "incontinence") is sometimes used for the lack of self-control.

How do we lose self control, and how can we strengthen it?

How is it that you can know the right thing to do, [resolve](#) to do that thing, know that you will regret not doing that thing, and yet *still* screw up at the last minute by choosing something else?

Part of the problem seems to be that it is easy to resolve to resist temptation when the tempting thing is at a distance and the tempting impulse is mostly theoretical. As the tempting thing becomes nearer and the tempting impulse more vivid, the earlier resolve is not strong enough to hold the fort. This may mean that the insights about [near and far mode thinking](#) will be important in understanding self-control: If you make your resolutions in far mode, but ultimately make your choices in near mode, they may get out of sync. This may also suggest that lack of self control is a sort of cognitive bias concerning [time discounting](#).

This suggests also that a distorted self image may also be the trouble: If you think of yourself as the sort of person who can easily overcome temptation (e.g. to finish off a

pint of Ben & Jerry's) when in fact you are not, and you fail to update your self image (due to shame, vanity, etc.) when you inhale that Cherry Garcia *in toto*, you will not take steps to improve your self control or to reduce the temptation in your environment. Better self-awareness and [humility](#) may come to the assistance of self-control in such a case.

Aside from incontinent lack of self-control, there is also a failure mode where the pendulum swings too far the other way: pig-headed stubbornness, in which you stick with your resolutions even when the underlying facts change or when it turns out your resolutions were faulty. A fetish for being "decisive" can lead you to stick with bad decisions when a wiser person would have been more flexible.

In addition to prompting unwise action, lack of self-control can also lead to unwise *inaction*, for example [procrastination](#) in which you commit to some wise course of action but then dawdle along doing something else instead, tempted by mere inertia or laziness.

Lack of self-control might be harder to fix than intemperance. With the intemperate person, you have the hope of persuading them that their desires are poorly chosen; with the incontinent person, they've already been persuaded but it doesn't seem to help matters any. On the other hand, the intemperate person usually has no regrets about their unwise course of action, while the incontinent person does, so at least the incontinent person has a motive to get better.

Aristotle, who considered self control extensively in [book VII of his Nicomachean Ethics](#), compared it with the virtue of [endurance](#).^[2] Self-control is resisting the temptation of things that seem immediately appealing; endurance is resisting the dissuasion of things that seem immediately uncomfortable. It can be difficult to distinguish them in some cases: is the regretful alcoholic reaching for the bottle because they cannot resist the temptation of a pleasing drink, or because they cannot endure the discomfort of withdrawal?

[Christians](#) have documented the struggle with akrasia in terms of a battle between the spirit which is inclined to God and the flesh which is mired in sin. Jesus, as his crucifixion approached, felt himself recoil from his chosen task, and noticed "the spirit indeed is willing, but the flesh is weak."^[3] Paul, in his letter to the Romans, complained, "I do not understand my own actions. For I do not do what I want, but I do the very thing I hate.... I know that the good does not dwell within me, that is, in my flesh. For the desire to do the good lies close at hand, but not the ability."^[4] And St. Augustine amusingly wrote in his *Confessions*, "As a youth I prayed, 'Give me chastity and continence, but not right away.' "^[5]

Some [philosophers](#) have insisted that incontinence does not really exist distinct from intemperance. In their view, the "incontinent" person does not act contrary to their own temperate judgement, but they merely reveal their real judgement at the last minute, having only deceived themselves into believing that they had good temperate judgement.

Good nutrition (sufficient blood glucose in particular), sufficient sleep, and [good health](#) make it easier to exercise self control. Some drugs (e.g. alcohol) make self control more difficult or at least less likely.

Peer pressure is notoriously erosive of self-control. For this reason, strengthening ones assertiveness and self-confidence may be an important factor in bolstering self-

control. On the other hand, peers who exercise self-control can perhaps [inspire](#) you to greater self control.

Sexual arousal is so notoriously effective at diminishing self control that it seems we have evolved to implement the heuristic that it is far better to have ill-considered sex with disastrous consequences than [not to have sex](#) at all.

There is [some evidence that self control is a resource that can be depleted with use](#) and that takes time to recharge. So for example if you have to use self control at task A, and then at task B, you will have a harder time than if you were just asked to use self control for task B. And yet, some evidence suggests that “self-regulation... can be strengthened like a muscle, which means that with regular ‘exercise,’ people can become less vulnerable to ego depletion effects.”^[6]

Mindfulness meditation may help to replenish self control reservoirs, at least in the short term.^[7]

The [Stanford marshmallow experiment](#) found a strong correlation between self-control exhibited by young children and the quality of their later life outcomes on a variety of measures. “Impulsivity” and “poor impulse control” are subjects of psychological investigation, and some interventions such as cognitive-behavioral therapy have shown some promise in improving impulse control.

Strengthening the [quasi-virtue of shame](#) may be a good way of making the negative consequences of incontinence more visceral so as to counter the temptation.

Rules of thumb like “don’t respond to an email while you’re angry” or “count to ten before you decide” can help give you the space you need to remember your long-term goals in the face of short-term temptations. Another example that a substance abuse counselor told me about was “play the tape forward”—envision the far-term consequences of the short-term temptation you are contemplating giving in to, so as to make those consequences more vivid, in the hopes that they will weigh more heavily against the temptation.

You might try to anticipate a future situation in which temptation will be strongest, so that you can practice in your imagination confronting and overcoming that temptation. That way at least the temptation will lose the advantage of taking you by surprise.

It is difficult to control what you are not aware of, so in some cases it can be helpful to increase your awareness of what you are trying to control by means of deliberate *monitoring* (carefully documenting e.g. how much money you spend and on what, how much you eat and of what, etc. depending on what you are trying to develop control over).

One school of thought holds that self-control is less an internal skill or trait like “willpower” and more a matter of [strategically adjusting one’s environment](#) to add friction to harmful temptations in order to make them less tempting. One way to do this is, when you commit to do the right thing, to establish an automatic penalty for later doing the wrong thing: the [Beeminder](#) app is one attempt to facilitate this process. B.F. Skinner created [a catalog of environmental interventions to influence self-control](#).^[8]

- See also: [How You Can Gain Self Control Without Self Control](#) (Spencer Greenberg)

Addiction and other edge cases

Addiction is either a particularly difficult example of lack of self-control or something that goes beyond mere lack of self control, depending on who you ask. Because the object of addiction comes packaged with strong reinforcing mechanisms (that only get stronger as dependence develops), it is hard to interrupt this with a will that lacks such enticements.

Addicts have painful insight into akrasia: solemn vows of sobriety at dawn that are broken by nightfall; a multitude of attempts to supplement self-control with techniques like “I will stop after two” or “beer and wine only this time” or what-have-you.

The first step in “Twelve-Step” programs is to admit that you are powerless over the object of addiction, that your self control has met its match. There is something of a paradox to this, in that you begin the path of controlling your addiction by surrendering and acknowledging that your will is weaker than it is. In [Alcoholics Anonymous](#), the addict gives up on trying to conquer alcohol with self control, and moves on to other strategies (e.g. peer support, a “higher power”, the steps).

Other examples of people whose actions don’t seem to align with their well-considered decisions are people with obsessive-compulsive disorders, Tourette syndrome, and things of that nature. While these seem to be pretty far afield from the usual lack of self-control that people complain about, it’s possible that things we learn about how to treat and control such disorders might be useful in more mundane cases.

1. ^

Aristotle, *Nicomachean Ethics*, book VII, chapter 10

2. ^

Aristotle, *Nicomachean Ethics*, book VII, chapter 7

3. ^

[Matthew 26:40-43](#): “Then he came to the disciples and found them sleeping, and he said to Peter, ‘So, could you not stay awake with me one hour? Stay awake and pray that you may not come into the time of trial; the spirit indeed is willing, but the flesh is weak.’ Again he went away for the second time and prayed, ‘My Father, if this cannot pass unless I drink it, your will be done.’ Again he came and found them sleeping, for their eyes were heavy.”

4. ^

[Romans 7:14-20](#): “For we know that the law is spiritual, but I am of the flesh, sold into slavery under sin. I do not understand my own actions. For I do not do what I want, but I do the very thing I hate. Now if I do what I do not want, I agree that the law is good. But in fact it is no longer I who do it but sin that dwells

within me. For I know that the good does not dwell within me, that is, in my flesh. For the desire to do the good lies close at hand, but not the ability. For I do not do the good I want, but the evil I do not want is what I do. Now if I do what I do not want, it is no longer I who do it but sin that dwells within me."

5. ^

Augustine of Hippo, *Confessions*, book VIII, chapter 7

6. ^

J.A. Bauermeister, et al. "[Assessing motivations to engage in intentional condomless anal intercourse in HIV risk contexts...](#)" *AIDS Education and Prevention* (April 2009)

7. ^

M. Friese, C. Messner, & Y. Schaffner, "[Mindfulness meditation counteracts self-control depletion](#)" *Consciousness and Cognition* (June 2012)

8. ^

B.F. Skinner, *Science and Human Behavior* (1953), chapter XV

Notes on Respect-for-Others

This post examines the virtue of **respect-for-others**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is this virtue?

The word “respect” is ambiguous; it covers several different things. For example: You can respect a person’s position or rank by granting them authority. You can respect a person’s reputation or skills or character or taste. You can respect the threat a potentially dangerous person or thing poses to you. You can show respect for someone as a form of showing submission to them.

The virtue of respect-for-others I mean to cover in this post is different. It has to do with understanding that other people have lives just as subjectively rich as your own, that they have their own perspectives, goals, desires, and priorities, and so forth, and that your own do not have objective priority over theirs.

This virtue is well summed up by the version of [Kant's Categorical Imperative](#) that goes like this: “So act that you treat humanity... always at the same time as an end, never merely as a means.”^[1]

There are a couple of ways people tend to describe how this variety of respect works. These are not mutually-exclusive, but people may emphasize one more than the other:

1. “I give every person some minimum baseline of respect that everyone deserves just by virtue of being a member of the human family, no matter who they are or what they’ve done.”
2. “I give everyone I meet a certain default amount of respect, and then adjust that amount up or down as I get to know them better.”

Related virtues

It seems odd to me that there isn’t a word in English that precisely encapsulates this virtue. Some related virtues that touch on respect-for-others include:

- concern, consideration, thoughtfulness, [compassion](#)
- sympathy, [empathy](#)
- civility, [politeness](#), tact
- acknowledgement, recognition, [appreciation, gratitude](#)
- liberality, tolerance
- [humility](#) (in the sense of not overvaluing oneself compared with others)

It also harmonizes with “humanism” and “philanthropy” in the sense of valuing human beings highly, relative to institutions, ambitions, or other-worldly values.

In human development

Children develop respect for others slowly, in stages, over many years. Early on, children have difficulty imagining that other people have their own perspectives and viewpoints or even their own versions of knowledge. Young children may see other people as extensions of themselves, and try to learn to manipulate them in the same spirit as they try to learn how to coordinate the movements of their bodies.

People with autism and Asperger's tend to have more difficulty forming a "theory of other minds," as do some people with schizophrenia.

But beyond just having the awareness of other independent minds, respect-for-others requires that those other minds be ungrudgingly allowed some independence from one's own projects and preferences. People with narcissistic personality disorder exaggerate their own importance or centrality relative to other people and expect other people to go along with that. People with antisocial personality disorder act as though they do not believe other people have any inherent value or that their preferences and pursuits are worthy of respect.

At the other extreme, people subject to abuse may become so fixated on understanding the motives of their abuser (in order to try to fend off the abuse) that they end up suppressing their own egos and becoming extensions of the ego of their abuser. [The Stockholm Syndrome](#) is one astonishing way this can play out.

Too much respect for others' points of view can lead to conformity pathologies such as those pointed out in the [Asch conformity experiment](#).

Some people seem never to confidently develop their own identities and viewpoints, and feel the need to assume off-the-shelf identities instead, or to merge their own identities with a charismatic figure or leader. They express borrowed opinions, assume fashionable tastes, speak in clichés, and so forth, seemingly under the delusion that they are not entitled to an identity of their own or that it would be too much trouble to maintain one. A particularly grotesque version of this is the sort of internalized *führerprinzip* displayed for example by [Adolf Eichmann](#), who adopted Hitler's values in place of his own and later tried to claim that he could for that reason assign the guilt for his actions in implementing the Holocaust to Hitler while remaining innocent himself.

What good is it? And the egoist objection

Respect for others is a sort of things-I-learned-in-kindergarten virtue. It's implied in the [Golden Rule](#) that has emerged in some form or other in folk ethics just about everywhere.

It plays an important role in other social virtues (e.g. friendship, [love](#), [trust](#), [justice](#)), and in some moral systems is the foundation on which the other virtues rest. For example, a person may be [honest](#) not so much because they love the truth as because they respect the person they are communicating with. Respect for others also is often found at the core of theories of political justice, in forms like "human rights," "inalienable rights," "equality before the law," and other such formulations.

But in spite of all of these credentials, is there a case for stopping short of respect-for-others? What if you were to acknowledge that other people have their own subjective experiences, projects, and priorities that are just as important to them as yours are to you, but not see this as any reason not to prioritize your own as being the only really important ones? At the very least, when the chips are down isn't it true that "every man for himself" rules the day? A straightforward [egoism](#) seems at first like it might be a reasonably strategic choice.

But even Ayn Rand, who disparaged altruism at every opportunity, included respect for others in her virtues. One's own self-interested pursuits ought to be undertaken, she wrote, with the understanding that other people are also entitled to their own such pursuits, and you should not expect them to be mere ingredients in your own plans: "[E]very living human being is an end in himself, not the means to the ends or the welfare of others — and, therefore... man must live for his own sake, neither sacrificing himself to others nor sacrificing others to himself."^[2]

Simone de Beauvoir pointed out how lonely and pointless the utterly egoistic viewpoint is. "If I were really everything, there would be nothing beside me; the world would be empty. There would be nothing to possess, and I myself would be nothing."^[3] She felt that "man can find a justification of his own existence only in the existence of other men,"^[3] and this only if we see other people as *complete* people like ourselves, not mere props or extras. What good is the admiration, love, and so forth, of people whose viewpoints we do not respect or cannot disentangle from our own? The egoist attitude that "winning isn't everything; it's the only thing" often seems to result in winning comparatively small and silly things in the grand scheme of things.

On the other hand, if you oppose egoism on the grounds that it actually won't work out well for you, that kind of sounds like you aren't really opposing egoism but some malformed and failed attempt at it. Maybe what you mean to say is that egoism and respect for others are compatible after all.

Who deserves respect?

If you buy that you ought to have respect for others, whom do you include in the set of such others? foreigners? heathen? children? babies? fetuses? animals? the disabled? bad people? sacred objects? the dead? nature? [φ-zombies](#)? nations? Do different classes of beings get different varieties of respect, or is it more all-or-nothing? What is it about others that makes them respect-worthy in this way, and do some people not have whatever that is, or do some non-people have it? Can you gain or lose it, or is it yours to keep once you have it?

Children are one tricky case. On the one hand, parents probably ought to respect their children as independent beings with their own preferences and characters, rather than trying inflexibly to fit them into preconceived molds. On the other hand, you wouldn't trust an immature child with a barely-formed view of the world to make major decisions about his or her destiny. In such a case, respect for the child seems to include an evolving and tentative sort of respect for a slowly-emerging autonomy. But we may want to pay more attention to the small ways in which we may show (and teach) disrespect for children, such as lying to them (e.g. about Santa or about where babies come from) or tickling them without their consent.

But as [important and interesting](#) as such questions are, in this post I want to side-step them. Assuming you believe that you ought to have respect for others, and assuming you have some adequate way of determining who those others are, what next?

What does “respect” entail?

Assuming you do respect someone in this manner, what does that amount to in a practical way? If you want to treat someone “as an end” rather than a means merely, how do you go about it?

Part of respecting someone is to respect them *as a person*: that is, being fully cognizant of their humanity rather than considering them as, for instance, a physical obstacle on the sidewalk between you and your destination. A friend of mine told me she is in the habit of giving people a little nod “in acknowledgement of their individual who-ness” as she passes them. “Often there is no response, but sometimes folks break out in a big grin and I feel like they appreciated having their selves respected, just for being.”

Another part of respecting someone is to respect them *as an individual* as opposed to a unit in an aggregate or a sample of a type. If you are thinking of someone primarily as a voter, a Native American, a human resource, or something of that nature, this can mean that you’ve already shoehorned them into some schema or project of your own as an interchangeable part, without allowing their own choices and interests to enter into it.

Another part of respecting someone is to respect them *on their terms*. This requires insightful attention. It might also present obstacles (for instance, if someone seems to demand respect in an unreasonable or unethical or overtaxing way).

Some ways people show respect: being courteous, giving people the benefit of the doubt (and being on guard against the [fundamental attribution error](#)), being tolerant of differences, being willing to share and take turns, exercising communication skills such as tact, being sensitive to people with vulnerabilities that you do not have direct experience with, respecting other people’s autonomy rather than [trying to make choices for them](#) or [manipulate](#) them or act on them without their consent, being aware of cultural differences (for example, in body language), not mocking or humiliating others or gossiping about them in their absence, and being helpful and cooperative unless there’s a good reason not to be.

Examples like those might be part of the respect package that a person with a strong sense of respect for others offers by default. For an example of more of a minimum baseline respect standard, the [non-aggression principle](#) is one concise formulation that is popular among political libertarians.

Obstacles to respect for others

One way I often see respect-for-others neglected is in commercial contexts. Customers will sometimes treat wait staff, cashiers, and such with no more regard than if they were vending machines. Or, employees will sometimes treat customers as merely potential sales.

Something about being paid-to-do-it seems to make some people willing to go way beyond the bounds of what they would otherwise find decent. For some forms of employment it almost seems *de rigueur* to treat people disrespectfully. In the wake of the [Milgram experiment](#), Milgram offered this interpretation of the results: “a person comes to view himself as the instrument for carrying out another person’s wishes, and he therefore no longer sees himself as responsible for his actions.”^[4] Something similar seems to happen to some employees, where they consider themselves to be not responsible for things that they do if they do them as part of their jobs.

It takes mental energy to model another person as a complex subjectivity with their own access to knowledge, their own models of the world, their own motivations, and so forth. We have to use guesswork and approximations and heuristics at the best of times. When our minds are also occupied with other tasks, these approximations can reduce to caricatures that may eventually be oversimplified so much that they might as well be mannequins. In order to respect people we have to permit them enough room in our mental models, and enough of our attention, so that they can appear to us as fully-dimensional people. A frequently encountered failure of respect-for-others is absent-minded inconsiderateness, in which a person whose mind is fully occupied with other things gives insufficient regard to the people around them.

One way to bolster respect for others, then, might be to periodically tune down whatever else is going on in your mind, look at the people around you, and attend to trying to understand them more vividly. The cashier who is checking out your groceries: does she appear relaxed or tense? do you think she is at the end of her shift or the beginning? is she new on the job or well-practiced? has she been standing for a long time or did she have a break recently? is she daydreaming or focused on her job? are there ways I put my groceries on the conveyor belt that made it easier or harder for her to process them? How does she answer when I ask “how has your day been going today?”

Another way lack-of-respect seems to bloom these days is in on-line interactions. Anonymity, pseudonymity, or even just being physically remote but virtually present, seems to bring out the worst in some people. If you’ve got a yen to be flamboyantly disrespectful to a stranger, by god you’re living in a golden age. You don’t even have to get up off the couch. You can be disrespectful to people by the thousands almost at the push of a button.

It is difficult even for otherwise well-behaved people to resist the temptation to, for example, share a video of some stranger embarrassing themselves in a particularly entertaining way. Is it respectful to help make someone notorious for some foible, weakness, indiscretion, or misjudgement... probably not. But if I were to judge myself by that standard, I’d fail the test.

If you are frequently *shown* disrespect — if people do not often reciprocate the respect you show for them — you will probably respond by giving people less respect by default and making them earn the rest. In this way, the typical standard of respect within a culture may decay. Sometimes subcultures develop that try to nurture and defend standards of mutual respect superior to those in society at large (fraternal orders, religious sects, the “[PLUR](#)” ethos, William S. Burroughs’s “Johnson family,”^[5] and things of that nature).

Thought experiments and games

There are some thought experiments that are designed to promote respect for others by evoking a “[had fate so decided, our positions might have been switched](#)” feeling. Most simply is just that: [imagine](#) what it would be like if you were in their place and they in yours.

A more complex version of this is a favorite of modern political philosophy: the “veil of ignorance” invented by John Rawls.^[6] Imagine that before you came into the world you had no idea who you would become, but you had a voice in what sort of world you would inhabit. What would be the ideal sort of political arrangement you would design from such an original position, if you knew you might end up assigned to any role within it?

My favorite is one that [Alan Watts](#) frequently returned to. I think he thought of it as more than a thought experiment: a revealed truth of some sort. Imagine that you are God, but being bored with being omniscient, omnipotent, and so forth, you decide to [invent Creation and then go there to hide from yourself](#), a bit like a king putting on grubbies to mingle with the commoners for a day. In this telling, God incarnates himself in each of us, hiding from himself so thoroughly that he forgets who he is and how he got here. The upshot of this is that while you are experiencing your life, including all of your triumphs and follies and pleasures and pains... the very same “you” is experiencing your neighbor’s life just as vividly. Imagine the respect you would feel for your neighbor if you knew that deep down you were them as well.

That thought experiment is one of those things I am tempted to believe not because I think there’s any good reason to believe that it is true, but just because I like the implications if it were true. I expect that means I will now have to do LessWrong penance of some sort.

Another way to build the skills of respecting other people might be through game play. You can’t be successful at chess, for example, if you can’t understand your opponent’s position and motivations. Role-playing games permit you to try on personalities and perspectives with goals and motivations unlike your own and may help you broaden your respect for different viewpoints. I wonder whether the many make-believe games of children — cops & robbers and the like — are in part exercises along those same lines.

1. ^

Immanuel Kant, *Groundwork of the Metaphysics of Morals* (1785)

2. ^

Ayn Rand, “The Objectivist Ethics” (1961) in *The Virtue of Selfishness* (chapter 1)

3. ^

Simone de Beauvoir, *The Ethics of Ambiguity* (1947)

4. ^

Stanley Milgram, *Obedience to Authority* (1974)

5. ^

Crops up in a few places in Burroughs's writing, a concept he borrowed from the author Jack Black (*You Can't Win*, 1926), e.g. *The Place of Dead Roads* (1983) "In this world of shabby rooming houses, furtive gray figures in dark suits, hop joints and chili parlors, the Johnson Family took shape as a code of conduct. To say someone is a Johnson means he keeps his word and honors his obligations. He's a good man to have on your team. He is not a malicious, snooping, interfering, self-righteous, trouble making person."

6. ^

John Rawls, [A Theory of Justice](#) (1971)

Notes on Piety

This post examines the virtue of **reverence** (a.k.a. “holiness”, “piety”, “sanctity”). I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

I am an atheist, and am addressing an audience in which, if I’m not mistaken, respect for the tenets of established [religion](#) is fairly low. But I want to explore reverence — in the spirit of [Chesterton’s Fence](#) — because it is common to many virtue systems across cultures and across time. Among the questions that concern me:

- Are there aspects of reverence that are valuable that rationalists can preserve and nurture in their own ways in their own traditions? [\[1\]](#)
- Is reverence perhaps so valuable that it is worth taking a “leap of faith” beyond the limits of rationalism in order to practice it?

What is reverence?

“Watch the stars in their courses and imagine yourself running alongside them. Think constantly on the changes of the elements into each other, for such thoughts wash away the dust of earthly life.” —Marcus Aurelius [\[2\]](#)

My first temptation is to ask whether reverence/piety is just a religiously-decorated variety of awe, elevation, or wonder, and to ask whether it can be replaced entirely by these secular alternatives.

Contemplation of the vastness of everything we know about, of the tremendous unplumbed chasm of the unknown, of the vertigo-inducing forever of infinity, of the mystery of why there is anything at all or any subjectivity with which to try to confront it... any of these things can induce a shudder of humble awe in the most dyed-in-the-wool atheist. Is that sufficient as the basis for reverence/piety?

One way I might think of reverence is as attention to *ultimate* values. You *instrumentally* value something if it helps you effect something else you value. But why do you value *that* thing? Maybe it too is instrumental, helping you to do or obtain something else. But if there are no ultimate values to terminate this chain, ones that you value for themselves and not instrumentally, this whole process becomes a comical tail-chasing.

To have an ultimate value means to take sides in a universe of flux and change (though I hear the Taoists reminding me that you can always take the side of flux and change).

Piety might in this view be considered those practices that bring our ultimate values to the forefront so that we can contemplate them and double-check that our instrumental values are in alignment with them.

Ultimate values are difficult to arrive at [rationally](#). If there is any irrefutable way to say of any particular ultimate value “and this is how I can prove it is ultimate” I don’t know what it is.

Maybe “God” in His various forms was invented in part to be a sort of ultimate value by-definition to serve for this purpose. If you are seeking an ultimate value, you might have criteria of wanting to choose the best thing that there is, the most important thing that there is, the most far-seeing perspective that can be had, the most influential thing that can be and so forth. If you do not have any idea what if anything satisfies those criteria, “God” seems a reasonable name for the placeholder. God, in this view, is the shape of the outline that the best ultimate values would have, even if we do not know what occupies that outline. By revering God we can devote ourselves to the best approximation of the ultimate values we can discover, while we continue to strive to discern what those values are more clearly.

I sense that I've gone out on a limb here. I've also been bending over backwards to accommodate piety without having to acknowledge faith in an actual, real, honest-to-God God.

I don't think a subjunctive God like the one I contemplated would satisfy most of the traditions that have given a place of importance to piety/reverence in their virtue systems. When I see people stress reverence and piety, typically they seem to have a pretty good idea of what occupies the space where God belongs, and it's a particular deity or pantheon from human mythology, whose characteristics and concerns are specified in uncanny detail. Maybe the human mind abhors a vacuum; the idolatrous temptation to fill-in-the-blanks is just too great to resist.

Related virtues

Aside from [awe](#), [elevation](#), or [wonder](#), which I mentioned earlier, other virtues closely related to piety/reverence include [devotion](#) and [faith](#).

What good is it?

In one sense, piety is just the name for whatever attitude it is appropriate for a person to have toward the ultimate (abstracted away from questions of what that attitude consists of and what or Who that ultimate is). So asking what good is it is kind of odd. It would be like asking a computer programmer what good is debugging, really. The good of it is built in to the definition of it.

But there are also ways in which religiosity may be instrumentally valuable. In certain religious frameworks, of course, piety is rewarded and impiety punished according to the terms of the religion. Those sorts of rewards and punishments are notoriously difficult to verify or test from within the mortal, mundane world, however.

There are also a variety of secular, potentially-testable benefits that have been claimed for religiosity, such that it:

- correlates suggestively with virtues like altruism,^[3] compassion,^[4] forgiveness,^[5] kindness,^[6] gratitude,^[7] and with prosocial values in general,^[8] and provides a framework in which people cooperate for community benefit^[9]
- is linked to optimism, hope, and happiness,^[10] and to physical and psychological well-being,^[11] and helps people cope with life's curveballs^[12]
- shields children from engaging in various harmful behaviors^[13]
- improves the quality of marriages^[14]

(I pulled these references from *Character Strengths and Virtues: A Handbook and Classification* by Peterson & Seligman, and from *Applied Positive Psychology* by Lomas, Hefferon, & Ivtzan.)

It can be difficult to tease out which aspect of religiosity may cause which benefit. Is it the belief, the attitude of reverence, the strictures, the congregation of peers, the institutional structures?

One group of researchers tried to investigate this.^[15] They divided people into four quadrants based on their having high or low religiosity (institutional religious involvement), and high or low spirituality (earnest search for the sacred). The groups of people who had the highest spirituality had the highest well-being scores (on measures of “self-actualisation, meaning in life, and personal growth initiative”); the worst-off were those with high religiosity but low spirituality. This suggests that sincere piety, rather than some incidental social/institutional artifact of religion, is what leads to at least some of religion’s benefits.

How to become more pious

Advice on how to become more reverent/pious differs in its details from religion to religion, but frequently includes things like the following:

- Prayer or contemplation that brings God / the object of reverence / your ultimate values to mind and prompts you to orient yourself appropriately in relation to it.
- Recitation / reading / listening to other pious people or to traditional scripture / prayers / chants that have a track record of inducing piety / reverence.
- Rituals, symbols, fasting, particular clothing or dietary choices, etc. that serve to periodically remind you of the ultimate value, whatever it is.
- Acts done in conformance with your highest values and with a consciousness that this is why you are doing them.
- Renunciation of instrumental values that do not serve the ultimate value, or of other ultimate values that on reflection aren’t worth the candle.

I’m aware of one study that attempted to *induce* piety via a series of exercises designed to psychologically implant “sanctification” on some object and thereafter use that object as a focus for sacred contemplation.^[16] It claims that this intervention increased measures of psychological and subjective well-being.

1. ^

A. de Botton *Religion for Atheists: A Non-Believer’s Guide to the Uses of Religion* (2012)

2. ^

Marcus Aurelius, *Meditations* 7.47

3. ^

R. Wuthnow, & V.A. Hodgkinson, *Faith and philanthropy in America: Exploring the role of religion in America’s voluntary sector* (1990)

M.D. Regnerus, C. Smith, & D. Sikkink, "Who Gives to the Poor? The Influence of Religious Tradition and Political Location on the Personal Generosity of Americans toward the Poor" *Journal for the Scientific Study of Religion* (1998), 37(3)

H. Lovell Smith, Anthony Fabricatore, & Mark Peyrot, "[Religiosity and Altruism Among African American Males: The Catholic Experience](#)" *Journal of Black Studies* (1999)

4. ^

Robert Wuthnow, *Acts of Compassion: Caring for Others and Helping Ourselves* (2012)

P. Steffen & K. Masters "[Does compassion mediate the intrinsic religion-health relationship?](#)" *Annals of Behavioral Medicine* (2005)

5. ^

M.S. Rye, K.I. Pargament, M.A. Ali, G.L. Beck, E.N. Dorff, C. Hallisey, V. Narayanan, & J.G. Williams, "Religious perspectives on forgiveness," in *forgiveness: Theory, research, and practice* (2000)

M.E. McCullough & J.E.L. Worthington "[Religion and the forgiving personality](#)" *Journal of Personality* (1999)

6. ^

Christopher G. Ellison, "Are Religious People Nice People? Evidence from the National Survey of Black Americans" *Social Forces* (December 1992)

7. ^

N. Krause "[Religious involvement, gratitude and change in depressive symptoms over time](#)" *International Journal for the Psychology of Religion* (2009)

8. ^

Jacqueline S. Mattis, Robert J. Jagers, Carrie A. Hatcher, G. Dawn Lawhon, Eleanor J. Murphy, & Yohance F. Murray, "[Religiosity, volunteerism, and community involvement among African American men: An exploratory analysis](#)" *Journal of Community Psychology* (19 June 2000)

9. ^

Kenneth I. Maton & Kenneth I. Pargament, "The Roles of Religion in Prevention and Promotion" *Prevention in Human Services* (1987)

K.I. Maton, & E.A. Wells, "Religion as a Community Resource for Well-Being: Prevention, Healing, and Empowerment Pathways" *Journal of Social Issues* (1995)

Andrew Billingsley, *Mighty Like a River: The Black Church and Social Reform* (1999)

Andrew Billingsley, & Cleopatra Howard Caldwell, "The Church, the Family, and the School in the African American Community" *The Journal of Negro Education* (1991)

C. Eric Lincoln & Lawrence H. Mamiya, *The Black Church in the African American Experience* (1990)

10. ^

Sheena Sethi & Martin E. P. Seligman, "[Optimism and Fundamentalism](#)" *Psychological Science* (1993)

Sarah French & Stephen Joseph, "Religiosity and its association with happiness, purpose in life, and self-actualisation" *Mental Health, Religion & Culture* (1999)

Jeffrey S. Levin & Robert Joseph Taylor, "Panel Analyses of Religious Involvement and Well-Being in African Americans: Contemporaneous vs. Longitudinal Effects" *Journal for the Scientific Study of Religion* (1998)

A. Moadel, et al. "Seeking meaning and hope: Self-reported spiritual and existential needs among an ethnically-diverse cancer patient population" *Psycho-Oncology* (1999)

11. ^

Neal Krause, "[Religion, Aging, and Health: Current Status and Future Prospects](#)" *The Journals of Gerontology Series B Psychological Sciences and Social Sciences* (1997)

Jeffrey S. Levin "Religious Research in Gerontology, 1980–1994: A Systematic Review" *Journal of Religious Gerontology* (1998)

12. ^

Paul J. Handal, Wandamarie Black-Lopez, & Stephanie Moergen, "Preliminary Investigation of the Relationship between Religion and Psychological Distress in Black Women" *Psychological Reports* (1989)

Kenneth I. Pargament, *The Psychology of Religion and Coping: Theory, Research, Practice* (1997)

David R. Williams, David B. Larson, Robert E. Buckler, Richard C. Heckmann, & Caroline M. Pyle, "[Religion and psychological distress in a community sample](#)" *Social Science & Medicine* (1991)

13. ^

Byron Johnson, David Larson, Spencer Li, & Sung Joon Jang, "[Escaping from the Crime of Inner Cities: Church Attendance and Religious Salience among Disadvantaged Youth](#)" *Justice Quarterly* (2000)

Michael Donahue & Peter Benson, "Religion and the Well-Being of Adolescents" *Journal of Social Issues* (2010)

Kenneth I. Maton & Elizabeth A. Wells, "Religion as a Community Resource for Well-Being: Prevention, Healing, and Empowerment Pathways" *Journal of Social Issues* (1995)

H. C. Stevenson, "Managing anger: Protective, proactive, or adaptive racial socialization identity profiles and African-American manhood development" *Journal of Prevention & Intervention in the Community* (1997)

14. ^

G. H. Brody, Z. Stoneman, D. Flor, & C. McCrary, "Religion's role in organizing family relationships: Family process in rural, two-parent African American families" *Journal of Marriage and the Family* (1994)

A. Mahoney, K.I. Pargament, T. Jewell, A.B. Swank, E. Scott, E. Emery, & M. Rye, ["Marriage and the spiritual realm: The role of proximal and distal religious constructs in marital functioning"](#) *Journal of Family Psychology* (1999)

15. ^

I. Ivtsan, C.P. Chan, H.E. Gardner, & K. Prashar ["Linking religion and spirituality with psychological well-being: Examining self-actualization, meaning in life, and personal growth initiative"](#) *Journal of Religion and Health* (2011)

16. ^

E.D. Goldstein ["Sacred moments: Implications on well-being and stress"](#) *Journal of Clinical Psychology* (2007)

Notes on Loyalty

This post examines the virtue of **loyalty**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is loyalty?

These are two senses of loyalty:

- **Loyalty as reciprocity** — A is loyal to B in this way because either A has reason to feel gratitude to B for some previously-granted favor, or expects to be able to get such a favor from B at some time in the future. “You pulled the thorn from my paw, and now I shall certainly come to your aid.”
- **Loyalty as partiality** — A is loyal to B in this way by reliably making the promotion and defense of B’s interests a priority to A when that would not be the case absent the loyalty. “I would never rat out a fellow member of the Water Buffalo lodge.”

Loyalty may suggest both “I’ll be in your corner” (unlike other non-loyal people), and “I’ll be in *your* corner” (not in your adversary’s), but sometimes one more than the other.

Phrases like that one that describe loyalty commonly include descriptions of body language and relative positions of bodies: “I won’t turn my back on you,” “Now I know where you stand,” “I’ve got your back,” “He stayed by her side through it all.”

Issues of loyalty provide the story arc of many a popular classic — for example movies like *Casablanca*, *Yojimbo / A Fistful of Dollars*, or the Star Wars films (will Annakin go over to the dark side? will Han abandon his comrades in their hour of need?). Betrayal and false loyalty define many a classic villain (e.g. Macbeth). Loyalties that make incompatible demands have been a staple of tragedy at least [since Sophocles](#) and the Mahabharata. We seem to take particular interest in stories that involve shifts in loyalty, hidden loyalties being uncovered, loyalties being put to the test, the disloyal getting their comeuppance, and that sort of thing. This suggests that careful attention to the loyalties of those around us may have been an important skill to have in the history of our species.

Synonyms and related virtues

The words “faithfulness” and “fidelity” are sometimes used more-or-less synonymously with loyalty (especially in the context of marriage vows). “Fealty” and “allegiance” cover something similar in the context of loyalty expressed upwards in hierarchies.

“Patriotism” sometimes gets used as a synonym for the loyalty a person feels toward their nation. “Filial piety” includes a specific variety of loyalty practiced towards ones

parents. “Solidarity” is a sort of implied loyalty that similarly-situated people are supposed to feel toward one another. “Teamwork” includes a sense of loyalty to the team itself and its goals.

Loyalty often gets discussed in combination with nearby-virtues like [commitment](#), dependability/reliability, and [duty](#). When people are unswervingly loyal to principles, ideas, and ideals, this can be a [reasoning](#) failure of unwise intellectual rigidity; but sometimes “loyalty” is used metaphorically in this context to describe devotion, consistency, integrity, and other such virtues.

Loyalty is important to the virtue of friendship (“The ground for the steadfastness and constancy for which we are searching in friendship is faithfulness.” —Cicero^[1]). A loyal friend is sometimes described as a “true” friend: one who has been tried and has passed the test.

In a professional context, when you agree to promote your client’s or customer’s interests as part of the contract or as part of the ethical obligation of the job, this sort of loyalty is sometimes called “fiduciary responsibility.”

Sometimes loyalty is used informally to describe a mere preference or habit (“a loyal Starbucks customer”). Other times we express a sort of loyalty to tradition or to our ancestors (“just like my grandparents did, and their grandparents before them” or “as the founding fathers would have wanted”).

Loyalty can conflict with other virtues — most obviously virtues like [impartiality](#), [objectivity](#), and justice, but really any virtue against which loyalty might plead the cause of vice. For this reason, some philosophers have given loyalty the stink-eye, seeing it as more of a temptation than a virtue.

What does loyalty commit you to?

What exactly loyalty demands of us is usually pretty vaguely defined. What loyalty consists of is often conveyed through anecdotes and exemplars (sometimes of the disloyal) rather than through rules.

This can make it seem like loyalty is less a compelling commitment and more of a post-hoc excuse for what a person wanted to do for other reasons. How do you choose between “I am loyal to you, so I must” and “I am loyal to you, but I won’t”? Loyalty may induce you to incur opportunity costs: you come to the assistance of whatever you are loyal to at the cost of working on your own pursuits. This gets trickier when loyalty encourages you to do things you would otherwise find actually un-virtuous or fully vicious.

People and institutions that rely for their strength on the loyalty that people have toward them do what they can to strengthen that loyalty. They may try to eliminate rival loyalties by demanding that you make loyalty to them paramount: “you cannot serve both X and Y,” “you’re either with us or against us.” They may, as the United States does with its schoolchildren, ask you to pledge your allegiance over and over again.

A loyalty-dependent institution like this can broadcast its strength by demonstrating the extremes its fanatics are willing to go to to show their loyalty. For this reason, they may ask people to signal their loyalty in various ways. Oaths, pledges, vows, insignia,

binding rituals, and things of that nature are legible ways to signal loyalty. But because they are easily-accomplished they may not be very effective gauges. More expensive signals are more reliable for this purpose, and so sometimes people are called upon to prove their loyalty by things that may seem absurd to outside observers: believe the unbelievable, defend the indefensible, assert the incredible, humiliate yourself, take the blame for something you didn't do.^[2] You can best prove your loyalty by doing something that is costly, that goes against your own interests, and that otherwise violates your moral code: something that you would obviously never do except for your loyalty. (And once you have done so, even though such abusiveness suggests that maybe your loyalty is misplaced, [the sunk-cost fallacy](#) may help cement your loyalty further.)

People may exploit the vagueness of what loyalty commits you to, by asking you to be loyal in a way that explicitly commits you to X, but then asserting at some later time that you implicitly committed yourself to Y & Z as well. Open-ended or unspecified commitments are especially tricky. "You said you'd be there if I needed you, and now I need you to help me hide this body."

Because of this potential for abuse and for conflicts with other virtues, loyalty is a virtue that requires strong bodyguards in the form of wisdom, discernment, foresight, vigilance, and [caution](#). If you are going to be fiercely loyal, you should take special care in deciding what to be loyal to. If you ask favors of the Godfather, expect to hear "Someday I will call upon you to do a service for me" in return.

Loyalty in tension with justice

Loyalty usually implies partiality, which is a problem if you value loyalty but also value impartiality and objectivity. For this reason, we suspect the judgement of people who have expressed (or suspected) loyalties that might induce them to put their thumbs on the scale.

People and institutions with more power, authority, and resources can use those things to extort, command, or purchase more loyalty, which they can then trade in for more power, authority, and resources. This can create a dynamic in which these things flow to where they are already concentrated, in a way that can seem unjust. However, places where lots of power, authority, and resources come together are notoriously dens of intrigue and back-stabbing, so maybe this dynamic is ultimately unstable.

Part of what is exceptional about Christianity is its emphasis on solidarity with the downtrodden as a way of demonstrating loyalty to Jesus — flipping that dynamic of demonstrating your loyalty through acts that benefit those who already have more than their share: "For I was hungry and you gave me food, I was thirsty and you gave me drink, I was a stranger and you welcomed me, I was naked and you clothed me, I was sick and you visited me, I was in prison and you came to me."^[3]

When you put energy and resources into displaying loyalty to (for example) a football team, those are resources that could be spent instead on those who have more genuine need. So maybe there's an effective-altruism argument for reducing the influence of such loyalty as well.

Although loyalty can be in tension with justice, it can also be a way of honoring justice. If loyalty is owed, then disloyalty is the unjust failure to honor a debt. The

disloyal are sometimes described as being *unjust* in their betrayals. This is especially true when loyalty is a sort of reciprocity (you came through for me in a pinch, so now you can count on me).

Loyalty as a coordination mechanism

If coordinated group effort is important to the success of some endeavor, loyalty (to the cause or to the institution or to the leader) is one mechanism for helping to ensure that individual efforts are appropriately focused on the common task. “Teamwork” is a variety of loyalty in which the members of the team value the goals of the team as a whole over their own personal goals, and behave accordingly.

“We must all hang together, or, most assuredly, we shall all hang separately.”
—Benjamin Franklin, to his fellow-revolutionaries, on the signing of the Declaration of Independence^[4]

The loyalty of patriotism is sometimes defended from charges that it is irrationally partial by people who say that strong nations are good, and loyalty is necessary to making a nation strong, so you should not try to judge whether your nation is *worthy* of your loyalty but you should be loyal to it in order to benefit your nation and increase its worth.

Loyalty is a “force multiplier”. One legend holds that during negotiations with the enemy on the brink of battle, the leader of the Hashashin abruptly ordered one of his soldiers to leap from the window of the room to his death. The soldier complied without hesitation and without a word. The representative of the enemy of the Hashashin realized from this how devoted and formidable an enemy he was facing, and so war was averted.

To share in the benefits that come with coordinated group action, it can be useful for individuals to signal that they are “team players” with robust senses of loyalty. I wonder if the subconscious reason people often ostentatiously display loyalties to things like sports teams, brands, and so forth, is that these things signal that they are capable of forming strong loyalties, and in this way they encourage other people to join with them in mutually-beneficial alliances.

In [prisoners-dilemma](#)-type games, a reputation for loyalty can help game players optimize their play.

Loyalty as a way of forming identity

Loyalty is a component of *belonging*, which people tend to value. “I am an American” may describe a mere accident of birth; “I am a *loyal* American” seems to bind me together with other Americans in a joint project. People often define themselves in part by the loyalties they have adopted. If you think of yourself as a Freemason, for example, or a Marine (*semper fi!*), or a husband or wife, you have an identity that comes necessarily packaged with certain expectations of loyalty.

Demonstrations of loyalty, declarations of loyalty, symbols and tests of loyalty, and the like, are ways of policing in-group/out-group boundaries.

In our eagerness to belong, people sometimes adopt loyalties (or pantomime as though they have) to ephemeral and arbitrary things, and for the most tissue-thin reasons. Once you start, for example, harmlessly rooting for the home team or being [true to your school](#), it can be hard to remember that your home team or school isn't really objectively better or more noble or more worthy. The teacher who as an experiment divided her class up by eye color and encouraged eye-color-solidarity among them was astonished to see "what had been marvelous, cooperative, wonderful, thoughtful children turn into nasty, vicious, discriminating little third-graders in a space of fifteen minutes."^[5]

Whether or not you can keep a critical, objective head about you while remaining loyal to a person, team, or cause, is a tough nut to crack. "Blind" or "unthinking loyalty" is usually looked down on, there is an honorable place for the "loyal opposition," and sometimes your most loyal friends are the ones who aren't afraid to tell you what you didn't want to hear.

Loyalty as extended reciprocity

When loyalty has been *earned* (e.g. through services rendered), it is sometimes seen as a form of gratitude. Expressions of loyalty can be acknowledgments of indebtedness, or that the original favor has not been forgotten.

Loyalty is sometimes seen as a potential resource that can be "cashed in" in a more concrete way at some future time. You might offer such loyalty in exchange for someone's help if you don't have a better way to incentivize them. If you cultivate a reputation for steadfast loyalty, the loyalty you offer at such times will have a higher value and you presumably can obtain more for it.

Sometimes, institutions will use this sort of mechanism as a mutual-insurance policy. For example, I understand that Masons typically take an oath to come to the aid of any other Mason in distress.

Conclusion

In what has become an alarming pattern with these virtue explorations, I picked up "loyalty" thinking that it seemed simple enough and that I had a pretty good handle on what it meant, but then the more I investigated the more complex it revealed itself to be.

1. ^

Marcus Tullius Cicero, *Lælius de Amicitia* (44 BC)

2. ^

see e.g. Tyler Cowen, "[Why Trump's Staff Is Lying: It's both a loyalty test and a proclamation of power](#)" *Bloomberg Opinion* 23 January 2017

3. ^

[Matthew 23:35-36](#)

4. ^

very possibly apocryphal

5. ^

["A Class Divided"](#) PBS *Frontline* (2003)

Notes on Compassion

This post examines the virtue of **compassion**. It is meant mostly to explore what others have learned about this virtue, rather than to give my own opinions about it, though I've selected what I found interesting or credible according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about compassion and how to improve in the practice of it.

What is compassion?

If you browse the literature about compassion you'll see a lot of hair-splitting about what it is: whether it is an emotion or a feeling or an affect or an attitude or a response, for instance. Trying not to get too lost in the weeds, and acknowledging that there are differences of perspective that may be very important in some contexts, here is my stab at it:

Compassion concerns our feelings and actions towards people who are suffering or unfortunate. It has three linked components:

1. Take notice of another person's suffering.
2. Become motivated thereby to relieve that suffering.
3. Take action with the intent of relieving that suffering.

The first step requires that you become aware of the suffering, which may be a matter of luck (you happen upon someone who is obviously suffering) or one of skill (you discern through subtle signs in someone's manner that they're suffering, or you take pains to learn about obscure suffering taking place out of sight). This implicates the further virtues of curiosity, imagination, sensitivity, and sympathy as well as some complex cognitive skills involved in understanding another person's needs, motives, and emotional states.

The second step distinguishes compassion from mere care-giving (of the sort a person might do professionally or out of duty, without necessarily doing it compassionately). In this step, learning about suffering triggers concern and an urge to relieve the suffering. In some people this happens quickly and subconsciously and seems almost automatic, the way smelling baked bread might make you hungry, or hearing James Brown might make you want to [get up offa that thing](#). People who feel this way may say that they themselves suffer sympathetically with the original sufferer, and this may be part of the motivation for wanting to relieve them. If you feel upset at the suffering you see but are not motivated to relieve it, you may be experiencing "alarm" or "distress" but not the sort of *concern* that is likely to motivate you to a compassionate response.

The third step involves translating that intention to relieve suffering into action. Compassionate people do this more or less skillfully, but this skill, unless artlessness rises to actual negligence, is usually seen as its own virtue, distinct from compassion which is satisfied by compassionate intent. If in step three you turn away *from* rather than toward the sufferer, as in an attempt to remove yourself from the suffering and stop being menaced by it, you are probably experiencing something more like "disgust" or "horror" or "contempt" than compassion.

There are a couple of directions that step three can go in. You might comfort the sufferer, for one. This typically involves a characteristic facial expression, vocal tone, and variety of physical touch that suggests and effects comfort, along with reassuring vocal content and maybe efforts to increase the sufferer's immediate safety and security and their distance from whatever is distressing. The other direction bypasses comfort for problem-solving: you devote your own resources to removing the threat, serving the needs, or otherwise addressing the external cause of the suffering. You might for this reason consider compassion itself to consist of steps #1 and #2, but for compassion to be *successfully consummated* that it then transition into something like comfort or problem-solving as a distinct but connected activity.

If you want to get an overview of the science around compassion, I found helpful "Compassion: An Evolutionary Analysis and Empirical Review" *Psychological Bulletin* 2010 May; 136(3): 351–374 [doi:10.1037/a0018807](https://doi.org/10.1037/a0018807).

Related virtues

There are several other virtues that are closely related to compassion. These are some:

- **Care** has two meanings that are each components of compassion. First, the sense of “care” in phrases like “I don’t care” — as indicating whether one takes notice of something or is willing to attend to it. Second, in the sense of “to provide care for” which may have connotations of nurturing, protectiveness, provision, and such (but which may be done for motives other than those that activate compassion).
- **Concern** (and sometimes “worry”) means to take interest in something’s well-being, but may stop short of the action part of compassion.
- If you exercise **muditā**, you delight in another person’s good fortune, and are pained by their misfortune. You might think of it as the opposite of both envy and *schadenfreude*.
- **Consolation** (“comfort”, “solace”, “succor”) is offering comfort to someone who is suffering so as to lessen that suffering, and is one sort of action compassion might prompt. Suffering caused by grief and loss is the specialty of consolation.
- **Kindness** is less specific than compassion. It is not necessarily motivated by suffering or misfortune, but might be motivated by needs or wants of any sort, or by nothing in particular (“random acts of kindness”).
- **Mettā** (“*maitri*”, “loving-kindness”) includes compassion but is broader and also includes things like benevolence, amity, and goodwill.
- **Consideration** (“thoughtfulness”) is a sort of preemptive compassion in which you imagine possible future suffering of others, with an eye toward avoiding or alleviating that suffering. For example, if you notice someone with bad knees is suffering from standing a long while at some event, compassion might prompt you to go get them a chair. If you know someone who will be attending the event has bad knees, thoughtfulness might prompt you to make sure a chair is on hand for them beforehand.
- **Sympathy** is understanding the feelings and perspectives of another person, and **empathy** takes that a step further by attempting to actually feel those feelings and perspectives as the other person feels them.
- **Love** is a word with many definitions, but in some of its uses it also encompasses compassion.
- **Pity** is a feeling of regret for another person’s misfortune, sometimes with a connotation of condescension towards the sufferer. (Nietzsche distrusted pity,

thinking of it as a sort of contagion by which other people's suffering infects those around them. He also suspected the motives of people who act from pity, suggesting that they demean and objectify those they pity.)

- [Nying.je](#) is compassion decidedly without pity or condescension, and with as much of a cognitive as an emotional prompting, though it itself is an emotion. It is an emotional conclusion of a combination of empathy and [reason](#).
- [Mercy](#) ("clemency", "indulgence", "leniency") is compassion exercised towards those who are in our power and towards whom we are justified in being (or otherwise inclined to be) punitive or harshly just.

What is compassion good for?

"What do we live for, if it is not to make life less difficult for each other?" —George Eliot^[1]

Compassion, or one of the virtues that includes or is included in compassion, is part of many virtue traditions from many times and cultures. It is also important in many religious traditions. For example: An interpretation of the incarnation and crucifixion of Jesus is that it was undertaken from a compassionate desire to feel the suffering of and effect the salvation of mankind; this was also a compassionate sacrifice by God of his own son ("For God so loved the world, that he gave his only begotten Son..."^[2]). Jesus instructed his followers to be compassionate, even to outsiders (e.g. [the Good Samaritan parable](#)). Christians were instructed: "Carry each other's burdens, and in this way you will fulfill the law of Christ."^[3] The four noble truths of Buddhism concern awareness of suffering and how to relieve it. The Bodhisattva Vow begins (in one version): "Suffering beings are numberless, I vow to liberate them all." "The compassionate, the merciful" are among the titles of Allah.

Compassion is almost always depicted and described as an admirable trait. Our words for describing compassionate people typically have a positive connotation, and even when describing someone who is compassionate unwisely or to excess (e.g. a "softie" or "pushover") are not very harsh. In contrast, people without compassion are described by terms like "unfeeling," "indifferent," "cold-hearted," "unkind" that are not neutral but terms of disparagement. So if nothing else, practicing compassion is a ticket to esteem.

Compassion, whatever its instrumental value, is also valued in itself as being one component of a flourishing human life. In this view, to human well is to human compassionately.

Compassion is also presumably good for the target of the compassion: the person or persons whose suffering the compassionate person aims to relieve. (Though this is conditional on the skill of the compassionate person, and can be complicated by things like [codependency](#) in which compassionate-like behavior can be harmful.) Sometimes compassion is described as being (or as best being) purely altruistic.

On the other hand, some people are skeptical of altruistic virtues, or suspect that traits like compassion emerged from a "slave morality" that stunts people who identify with it. One variety of this criticism of altruistic compassion is this: if we place too much value in helping the least fortunate people tread water, we will fail to devote ourselves to grander pursuits and so will waste our potentials.

People who practice compassion do not always describe it as being altruistic in the sense of being self-denying or self-sacrificing. One interpretation of compassion is that the compassionate person starts at baseline, feels sympathetic suffering, then tries to relieve that suffering so they can return to baseline. If you interpret compassion as what happens *before* you act (stopping at step #2: the sympathetic feeling toward another person's suffering), compassion seems like it should be unpleasant.^[4] But compassionate people often report that when they behave compassionately this brings them *above* baseline. Perhaps a better way of conceptualizing this would be to say that as a compassionate person helps someone rise from a sorrowful position to a better one, the compassionate person can ride on their coattails and improve their own state as well. If so, compassion has benefits both for the relieved sufferer and for the person providing the relief: the practice of compassion is inherently rewarding.

Compassion also seems to be a crucial ingredient in other things people value, such as friendship, romantic relationships, and parenting. Also: If we live in a culture in which we expect that people behave compassionately, this is a sort of social safety net, which can make us feel more secure in taking risks, or feel less insecure about old age and debility.

The mechanism of [reciprocal altruism in humans](#) suggests that by offering compassion you are more likely to receive the benefits of compassion from others if you are in need, so it is also potentially valuable to the compassionate person in this way.

How did compassion happen?

In 1759, philosopher Adam Smith published *The Theory of Moral Sentiments* in which he argued that the human moral feelings that prompt us to actions are not the results of applying abstract ethical rules, but that our ethical rules are post-hoc heuristics approximating preexisting moral emotional reactions. Smith thought these moral emotions were implanted in us, and were calibrated to be the way they are, in order to help us survive, thrive, reproduce, and raise children effectively. Smith didn't know it, but he had written a book of evolutionary psychology.

Darwin's *Origin of Species* came a century later (followed by *The Expression of the Emotions in Man and Animals*), and helped to fill in some of the blanks about how these moral emotions came to be implanted and calibrated.

Compassionate-seeming behavior is found in many other species as well. In animals that are born unable to fend for themselves, for example, the pattern of the child crying out in need or distress, which causes the parent to attend and try to relieve the distress, is common and unremarkable. Less prominent but also observable (and often adorable) in other animals is compassionate-seeming behavior directed towards non-offspring (even occasionally non-relatives and members of other species).

Speculation and "[just-so stories](#)" abound concerning how compassion in humans and other animals may have evolved. Mechanisms like "[kin selection](#)" and "[reciprocal altruism](#)" have been suggested as ways in which such altruistic-seeming behavior could emerge through natural selection. Darwin himself considered a "[group selection](#)" explanation.^[5]

Human children are helpless, and then immaturely inept, for twice as long as any other primates. From what we know of other extinct hominids, this is the result of a process of increasingly delayed maturity (*Australopithecus afarensis* matured earlier

than *Homo habilis* who matured earlier than *Homo erectus*, who matured earlier than *Homo sapiens*). This could have necessitated a dramatic expansion in the capacity for and an increased prioritization of compassion in human parents and other kin. This in turn may help to explain why human compassion is so important to us and why it seems to have expanded to extremes that seem hard to justify by evolutionary pathways (humans may feel compassion for people on the other side of the world, for fictional characters, for the souls of the dead, for members of other species, etc.) Perhaps in humans it was more evolutionarily important that we become more compassionate quickly than that we do so in a precisely-targeted way.

Because for human children to survive and thrive they have to have compassion shown toward them often and for a long period of time, the amount of compassionate behavior we see modeled for us is greater than it would be otherwise. To the extent that our behavior is based on mimicry of what we observe, therefore, this could also have caused a more general expansion of compassion: amplifying any less-dramatic effects of genetic evolution. Human children apparently begin to show behavior characteristic of compassion (e.g. soothing others in distress) as early as age two.

Displaying compassionate behavior might be a good way of advertising to potential mates that you have what it takes to care for children competently (it also advertises that you have enough surplus resources that you can afford to behave altruistically), so [sexual selection](#) may have also been a factor both in expanding compassion and in increasing the esteem people feel for the compassionate.

Human susceptibility to compassion has not gone unnoticed by other species: Housecats seem to have learned to make sounds, specific to their interactions with humans, that mimic the sounds of needy babies and prompt us to exhibit compassionate behavior towards them.^[6]

Here are a few resources on the evolution of compassion:

- Frans de Waal: [“The Evolution of Empathy” Greater Good Magazine](#) (his book [Primates and Philosophers: How Morality Evolved](#) goes into this in more depth).
- Dacher Keltner: [brief TEDx video](#) and [longer talk](#) on the evolution of compassion.
- Robert Wright: [The evolution of compassion \(TEDSalon video\)](#).

Partial compassion

Compassion in its natural form in humans is not impartial. We tend to be more or less likely to exhibit compassion towards people and to be willing to exert more or less effort on their behalf based on a variety of factors. These factors have started to come under scientific investigation, and results so far suggest that we have some peculiar heuristics baked in. For instance:

- As the number of people we learn are suffering increases, our willingness to help them seems to *decrease*. For example, people seem to be more willing to donate more money to save the life of one child than to save the lives of two children.^[7]
- People are more apt to feel distress about and to help a particular identified individual victim than an anonymous one, and more also than one particular victim if that victim is identified as being part of a larger group of victims.^[8]
- We are more apt to offer help to a certain number of needy people if that number is closer to the whole number of people in need. In other words, if you

can help five of eight starving refugees you are more likely to do so than if you can give just as useful help to five of eighty.^[9]

- How much we feel compassionate towards members of another species correlates to how recently our species and theirs had a common ancestor.^[10]

In general, our willingness to behave compassionately to others seems to fall off as they become more distant from us in terms of family-relation, geographic proximity, cultural background, appearance, language, and so forth.

People who see this as a flaw in our moral reasoning say that we should try to flatten this curve such that we would exhibit as much compassion for people who are more different and distant from us — that we should give compassion “to each according to their need” rather than according to accidents of proximity or similarity. This is sometimes termed “expanding the circle of moral concern” and the influential ethical philosopher Peter Singer has done excellent philosophical and rhetorical work in that area.^[11]

People may also withhold compassion from an expectation that someone else will step in to do so in their place. The [bystander effect](#) takes hold when multiple people are able to offer assistance but each of them waits for someone else to take the initiative, and as a result the assistance never takes place. People sometimes also fail to respond compassionately out of a vague sense that “the government” must have some department that is responsible for doing it instead of them.

Being worthy of compassion

Another way we sometimes ration our compassion is by withholding it from people whose suffering seems self-inflicted or deserved. The suffering of people who seem to have earned it or brought it upon themselves can even lead to responses that seem to be the opposite of compassion: rather than feeling sympathetically sad, we feel joyful *schadenfreude*; rather than offering soft words of reassurance, we laugh. We may even feel the urge to *inflict* suffering on people we think have earned it.

(Similarly, Aristotle paired compassion with indignation and thought they were both worthy feelings. Compassion is taking pain in another person’s unearned misfortune; indignation is taking pain in another’s unearned good fortune. Both situations are an unhappy awareness of an unjust imbalance.)^[12]

We also may resent showing compassion to people who do not reciprocate by showing us gratitude, or who demonstrate a lack of compassion themselves.

One reason why we might temper our compassion in these ways is that it discourages people from manipulating compassion in order to be parasitic.

Whether our sense of compassion *ought* to be tempered by our sense of justice in these ways is controversial. Some who have considered the subject think that compassion ought to be blind to questions of desert or justice, or at any rate should operate independently from such questions:

“We must learn to regard people less in the light of what they do or omit to do, and more in the light of what they suffer.” —Dietrich Bonhoeffer^[13]

Effective altruism

The “[effective altruism](#)” movement takes for granted that our compassion ought to be impartial and altruistic in the way it relieves suffering, and tries to make it more efficiently so.

The efforts to “expand the circle of moral concern” that I mentioned earlier are part of the effective altruism project. It also tries to promote *more* compassionate behavior: suggesting that there is so much tractable suffering in the world that anyone with excess resources can and ought to do the most good by devoting the bulk of those resources to addressing the suffering. Finally, the effective altruism movement diligently pursues the question of how to reduce the most suffering per amount of effort, so that if you follow their advice you will get the most bang for your buck.

There are similarities and differences between effective altruism and compassion. Effective altruists try to become aware of suffering; once aware their concern prompts them to help; and then they try to offer genuine help to alleviate the suffering. That seems like it checks all the boxes. And yet in many ways it does not resemble human compassion as it is usually encountered and practiced. It is relatively unemotional and affectless, almost never involves physical touch or soothing, and only rarely evokes the sort of cathartic relief of having immediately helped someone that face-to-face compassion offers. If you throw out that much bathwater, it’s best to double-check for the baby. Depending on how you look at it, effective altruism might be a peculiarly rational and generous form of compassion, or it might be something different: perhaps a particularly rational and compassionate form of generosity.

One thing effective altruism highlights is that unless we apply rational scrutiny to our generously compassionate impulses, we may not actually be very helpful. This can certainly happen when we donate money to wasteful or grossly suboptimal charities, or donate our time in some something-a-thon that produces few practical results for all of the feel-good hoopla. Well-meaning compassionate interventions can backfire if they are not critically scrutinized (for example, a food relief program that undercuts and thereby retards the growth of markets that can supply food more reliably). But we can also be ineffective or counterproductive in our more personal compassionate endeavors. A parent who rushes in to rescue a child from every mishap and scrape may be “overprotective” in an ultimately unhelpful way. It is good to help a person who suffers, but also good for a person to develop the skills to handle suffering on their own. Compassion is much more helpful if it is served with a healthy dose of discernment and wisdom. Otherwise, well, at least your heart is in the right place, for what that’s worth.

Self-compassion

“Self-compassion” is a term that describes a way of processing one’s own suffering in a way that is analogous to how one might show compassion for someone else. This would be in contrast to wallowing in the suffering, suppressing/denying it, or amplifying it, and it is supposed to be a healthier way to go about it.[\[14\]](#)

How to become more compassionate

"This is no easy task... there is no blessing or initiation — which, if only we could receive it — or any mysterious or magical formula or mantra or ritual — if only we could discover it — that can enable us to achieve transformation instantly. It comes little by little, just as a building is constructed brick by brick or, as the Tibetan expression has it, an ocean is formed drop by drop.... Nor should the reader suppose that what we are talking about here is the mere acquisition of knowledge. It is not even a question of developing the conviction that may come from such knowledge. What we are talking about is gaining an experience of virtue through constant practice and familiarization so that it becomes spontaneous. What we find is that the more we develop concern for others' well-being, the easier it becomes to act in others' interests. As we become habituated to the effort required, so the struggle to sustain it lessens. Eventually, it will become second nature. But there are no shortcuts." —Tenzin Gyatso (the Dalai Lama)^[15]

Because compassion is a complex, compound process, there are many ways it can fail. For example: You might not become aware of suffering (e.g. via difficulty in reading others, narcissism, inattentiveness), you might not be moved by the suffering of others (indifference, pitilessness, egocentrism), or you might not be moved to act effectively (neglect, lack of confidence, histrionics). If you struggle with developing compassion, therefore, it may be helpful for you to examine yourself closely to try to determine if one of these steps in particular is your weak link or if you typically fall short in a particular way.

If you have difficulty becoming aware of suffering, for example, you might try techniques like [active listening](#) or you might try change how you greet people: with a "how are you doing?" that sounds (and appears, and is) more genuinely inquisitive. You could also volunteer for a hospice or homeless shelter or in some other way put yourself into a situation in which human suffering is likely to become apparent to you if you are attentive. There are some tools out there designed to help people become better at reading the emotions of others in their facial expressions and body language, and these may help you to become more aware of suffering if this is something that does not come to you naturally.

Even if you do not feel as though you struggle to become aware of suffering, it can be useful to work on this area, as you sometimes have to be very astute to see suffering in some people (some people hide their suffering well; some symptoms of suffering masquerade as other things).

Some people deliberately avoid becoming aware of suffering because they find it upsetting or frightening. This habit makes it more difficult to practice compassion.

"Shun not the scenes of human misery, nor flee from the abode of distress and poverty; for if we desire to be capable of having compassion for the sufferings of an unfortunate brother, we must be acquainted with the various scenes of misery which this world exhibits." —Baron Knigge^[16]

If you do not find the suffering of others to be upsetting — if it does not prompt you to deviate from your routine — the typical advice for fixing this failure mode involves using your imagination: "Put yourself in their shoes." Take some time to imagine that you are suffering what they suffer, as evocatively as you are able. If you deliberately kick-start your empathy in this way when you notice suffering, eventually it may come to life on its own without such deliberate attention.

There are varieties of meditation, such as karuṇā meditation, mettā meditation, or the “exchanging self and others” variety of meditation, that use imagination to evoke and strengthen compassion in the meditator.^[17]

You may be *afraid* of being compassionate: afraid of being contaminated by another person’s suffering; afraid that if you step forward to help someone you will be taking on more of an obligation than you are prepared to meet; afraid of initiating what amounts to an intimate relationship with an unknown quantity; afraid of getting trapped in a pity-party; afraid of being taken advantage of; afraid that if you try to help you will bungle things.

If fears like these are obstacles to your practice of compassion, you may want to work on courage more generally. Becoming skillfully assertive in setting personal boundaries can also remove some of these threats.

To learn how to strengthen compassion in yourself and others, it might be worth paying attention to the ways in which institutions intentionally try to *degrade* compassion. For example, compassion can interfere with the pursuit of certain military goals. If you feel compassion for people, you may be reluctant to kill them, maim them, torture them, orphan their children, destroy their property, or participate in other such important aspects of national policy. Much deliberate erosion of compassion of this sort exploits the biases described earlier: it highlights and exaggerates how *different* people of the enemy nation are from you, and how *deserving* they are of whatever suffering you are meant to inflict on them.

A (stereotypically masculine) failure mode in *acting* compassionately is to reflexively try to fix the problem that is causing the suffering rather than trying to comfort the person suffering. This is not always the wrong thing to do, but there are a couple of reasons why over-reliance on this sort of response is a problem: First off, usually the suffering person is much more aware of their situation and of the pros and cons of their possible responses to it than the responder is. When the responder compassionately makes obvious suggestions for (or attempts at) fixing the problem, this can be seen as condescending and insulting (“if it were that easy, I wouldn’t be in this fix”). Secondly, there are many situations of suffering that cannot be fixed, but must simply be endured. These too call for a compassionate response, but if the only sort of response you know how to offer is a “let’s fix it” response, you’ll be caught flat-footed.

A more subtle mistake is to try to help the suffering person the way you would want to be helped if you were in their situation. While such a response is probably a good first approximation for how you can best help, everybody is a little different, and you should recognize that what would be helpful for you might not be for everyone. It is more effective to be flexible and be willing to adjust your methods of assistance depending on how they are received.

“Do not do unto others as you would have them do unto you—they might have different tastes.” —George Bernard Shaw

While I was researching this post I also came across compassion it which says it has “Stanford-trained compassion facilitators” who lead seminars on boosting compassion. I don’t know any more about it than that.

George Eliot, *Middlemarch* (1872), in the words of the character Dorothea Brooke

2. ^

[John 3:16](#) “For God so loved the world, that he gave his only begotten Son, that whosoever believeth in him should not perish, but have everlasting life.”

3. ^

[Galatians 6:2](#)

4. ^

See, for example, P. Condon & L. Feldman Barrett “Conceptualizing and experiencing compassion” *Emotion* (2013)

5. ^

Charles Darwin, *Descent of Man, and Selection in Relation to Sex* (1871)

6. ^

See, for example, David Grimm, “[Cat Purrs Evoke Baby Cries](#)”, *Science* (13 July 2009)

7. ^

D. Västfjäll, P. Slovic, M. Mayorga, & E. Peters, “[Compassion Fade: Affect and Charity Are Greatest for a Single Child in Need](#)” *PLoS ONE* (2014)

8. ^

Tehila Kogut, & Ilana Ritov, “[The ‘Identified Victim’ Effect: An Identified Group, or Just a Single Individual?](#)” *Journal of Behavioral Decision Making* (July 2005)

9. ^

D. Fetherstonhaugh, P. Slovic, S. Johnson, et al. “[Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing](#)” *Journal of Risk and Uncertainty* (1997)

10. ^

A. Miralles, M. Raymond, & G. Lecointre, “[Empathy and compassion toward other species decrease with evolutionary divergence time](#)” *Scientific Reports* (2019)

11. ^

For example Peter Singer, [The Life You Can Save](#) (2009).

12. ^

Aristotle, *Rhetoric*, book II, chapter 9

13. ^

Dietrich Bonhoeffer, "After Ten Years" *Letters and Papers from Prison* (1942)

14. ^

See Susan David, "[How to be kinder to yourself](#)" *ideas.TED.com* (21 February 2020) for a brief introduction to the idea.

K.D. Neff "[Self-compassion: An alternative conceptualization of a healthy attitude toward oneself](#)" *Self and Identity* (2003)

K.D. Neff & C.K. Gerner "[A pilot study and randomized controlled trial of the mindful self-compassion program](#)" *Journal of Clinical Psychology* (2013)

15. ^

Tenzin Gyatzo, *Ethics for the New Millennium* (1999)

16. ^

Adolf Freiherr von Knigge, *Practical Philosophy of Social Life* (1788), chapter XIV

17. ^

S. Salzberg *Loving-Kindness: The Revolutionary Art of Happiness* (1995)

B.L. Fredrickson, et al. "[Open hearts build lives: Positive emotions, induced through loving-kindness meditation, build consequential personal resources](#)" *Journal of Personality and Social Psychology* (2008)

Notes on Wisdom

This post examines the virtue of **wisdom**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

Singing the praises of wisdom at LessWrong has a [bringing coals to Newcastle](#) feel to it. After all, isn't this community all about working hard and passionately to hack through the jungle of bias, illusion, and ignorance in search of the hidden temple of Athena?

So I was tempted to skip over wisdom and work on writing up some other virtue instead. But I'm hoping that by exploring wisdom as-a-virtue I can illuminate some facets of it that otherwise receive less attention here.

Two varieties of wisdom

There are two senses of wisdom that are found in some virtue traditions:

1. *phrónēsis*, or “practical wisdom” (sometimes translated [“prudence”](#)), which concerns knowing how the world works, and reasoning well about how to pursue goals effectively (and about which goals are worth pursuing — which sometimes gets separated out into “conative wisdom”)
2. philosophy, which concerns a more big-picture understanding of “what it’s all about,” whether or not there seems to be any way to make practical use of that understanding

They are both important: *Phrónēsis* without philosophy can make you merely clever; while without *phrónēsis*, philosophy can leave you with your head in the clouds, unable to bring your wisdom down to earth where you can make it matter.

“The title wise is, for the most part, falsely applied. How can one be a wise man, if he does not know any better how to live than other men? — if he is only more cunning and intellectually subtle?” —Thoreau^[1]

Philosophy is also sometimes considered an important end in itself. Aristotle thought it was the richest and most satisfying activity for people to engage in, and reasoned that it was the pastime of the gods.

The person with the *virtue* of wisdom habitually and regularly prioritizes thinking and behaving wisely. Which raises the question: why *wouldn’t* you? You might at first think that the only reason why you would deliberately think or behave *unwisely* is because you believe mistakenly that you are being wise.

That is *one* way you can go astray: you might understand the wise course of action based on the sort of situation you are in, but mistakenly believe you are in some other sort of situation; or vice-versa, you might understand the situation you are in well enough, but be mistaken about how to confront situations of that sort wisely. But

people are also deflected from wisdom by being overwhelmed by emotions like fear or anger, or by sensations like pleasure or pain. For this reason, virtues like courage, endurance, [self-control](#), and [temperance](#) can come to the assistance of wisdom.

Wisdom and mistakes

It is a popular belief that we gain wisdom (or gain it most effectively) by learning from our own mistakes.

“Wisdom is a virtue of old age, and it seems to come only to those who, when young, were neither wise nor prudent.” —Hannah Arendt^[2]

On the other hand, learning from other people’s mistakes may be the more prudent way to go about it (#LFMF!). LessWrong is in part a collection of dead-ends marked by warning signs, pointing out the mistakes in reasoning that others have been waylaid by.

But you typically learn other people’s mistakes from other people’s failures, which may leave your own artisanal mistakes unchallenged. If you are willing to strap on your theories and go into battle with reality [until you lose](#), you will be more likely to discover and shed your worst theories. This takes [courage](#), confidence, [industriousness](#), and [a willingness to fail and to admit failure](#).

Wisdom and age

Wisdom is popularly associated with age. This is one way it is distinguished from *intelligence*, which (by some measurements) typically peaks in early adulthood. That said, children and young people who are “wise beyond their years” are also a common trope, and metrics of wisdom designed by psychologists fail to find the expected correlations between wisdom and age.^[3]

Wisdom is often described as a variety of *perspective* that benefits from a wider or longer familiarity with the variety of things life tosses up. (See also: [Moderation](#), [Balance](#), and [Harmony](#).) A wise person looks at the big picture. Where an intelligent person may be the first to say “I know how we can solve problem x ” a wise person will be the first to notice “ x is not really the problem we should be focusing on.”

Another way intelligence and wisdom are sometimes contrasted is when intelligence is considered as an individual skill of mental sharpness and agility, but wisdom as more of a collective and long-term project of cultural assimilation. In this way of putting it, individuals may develop intelligence on their own as intelligent animals, but they *tap into* wisdom by intelligently observing and reflecting on the institutions, aphorisms, myths, customs, exemplars, and so forth that previous generations have assembled.

Surfing less unwisely

“The fool doth think he is wise, but the wise man knows himself to be a fool.”
—Shakespeare^[4]

At least since Socrates, wisdom has been associated with epistemic humility. The “LessWrong” name itself nods at that tradition: To be more wise, assume that you are

wrong, try to figure out where and how, patch that up as best you can, lather, rinse, repeat. Don't be too proud of the nuggets of wisdom you have dug up, but occasionally peer into the vast voids of ignorance, the blank spaces on the map. Imagine those things that *could* be true that would mean utterly overthrowing most of what you currently suspect to be true. Don't become attached to your best guesses or too inclined to round off a high probability into a certitude, but always prefer reality to your favorite hypothesis.

Wisdom seems to have less to do with arriving on the firm ground of confident understanding, and more to do with learning to surf the unstable edge of profound uncertainty: neither clinging to the barely-buoyant flotsam of belief nor being pulled out into a sea of nihilism by a undertow of skepticism.

Mystical vs. rational wisdom techniques

To understand and make our way in the world around us, we try to systematize, to find regularities, to discover cause-and-effect relationships, and so forth. We create a map, using our knowledge of the territory that we have passed through, to help us predict the territory we are about to enter. By extrapolating from suggestive patterns in the world, our maps can illuminate things we do not experience directly, and can suggest places to look to discover more than we might have stumbled upon on our own. [Habits of rational thinking](#) help us to keep our maps from misrepresenting the territory, and warn us about where our maps might be misleading even when they are as accurate as we can make them.

Mystical wisdom techniques suggest a different way to go about it: rather than just improving your map and your map-reading, take some time also to look directly at the territory and improve the quality of your vision. The advantage of this approach is that you lose the compression artifacts and other errors that come from trying to reconstruct the territory from the map. A disadvantage is that while maps can sometimes be shared, visions have to be turned into maps before they can be — and by the time you have turned your vision into a map, there may be little to recommend it when compared with maps arrived at through more rational methods.

1. ^

H.D. Thoreau, ["Life Without Principle"](#) *The Atlantic Monthly* (1863)

2. ^

Hannah Arendt, "Isak Dinesen" *Men in Dark Times* (1968)

3. ^

Christopher Peterson & Martin E. P. Seligman *Character Strengths and Virtues* (2004) pp. 189-190

4. ^

William Shakespeare, *As You Like It*, act V scene 1 (Touchstone)

Notes on Temperance

This post examines the virtue of **temperance**. It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is temperance?

Temperance is not a word that you hear a lot these days, and when you do, it's often in the context of historical discussions of the "temperance movement" that culminated in the experiment with alcohol prohibition. But temperance used to be central to discussions of ethics and personal excellence. It was one of the four "cardinal virtues" of ancient Greek philosophy, for example.

To be temperate is to have a well-regulated set of desires, meaning that the desires themselves are good ones, that they are well-proportioned (none are exaggerated to unhealthy levels), and that the way we respond to them is appropriate. This of course raises the questions of which desires are the healthy ones, how to know how much of a desire is too much, and what is the proper way to respond to a desire. Another difficult question is to what extent our desires are tractable, or whether perhaps they are just givens that we have to work with or around.

Aristotle suggested a comparison between temperance and courage: Courage enables us to keep our wits about us and behave honorably in the presence of frightening things (rather than behaving shamefully from cowardice); temperance enables us to do this in the presence of desirable things (rather than behaving shamefully from self-indulgence or covetousness).

He also drew a distinction between temperance and self-control. While a person with self-control is able to resist the temptation to do something they are sorely tempted to do, a temperate person is not so very tempted in the first place: they have better-regulated desires and do not typically need exceptional self control to overcome them.

In addition to the ordinary failure of intemperance, you may also respond to desire with an unhealthy asceticism or anorexia as a sort of counterfeit temperance. Such a thing suggests either that you are so intemperate that you cannot trust yourself to respond to desire appropriately, or that you are ironically giving into desire but in a pathologically inverted way.

Can we shape our desires?

You sometimes hear expressed the opinion that our desires are innate characteristics rather than choices. This became a major point of contention in the LGBT&c rights movement, and its pushback against the term "sexual preference" in favor of "sexual orientation." (It is not unusual to hear someone say "and that's when I *realized that I was sexually attracted to women*" but you almost never hear someone say "and that's when I *decided to be sexually attracted to women*," for instance.)

“De gustibus non est disputandum” (there's no accounting for taste), goes the saying, suggesting that when we get down to the level of desire and raw preference, we leave the arena of argument and [reason](#) and choice for a place where the irrational holds sway and we must simply take at face value what we find.

But it is not so simple as that. For one thing, our desires do change over time, and while some of this is not under our control (hello puberty!), some seems to be.

Your desires may become more refined: For example, if you have a yen for the manly-men-in-action genre, a college literature class or an influential friend may help you see delights in *For Whom the Bell Tolls* such that you can never go back to [The Executioner](#) with the same appreciation.

You may learn something about an object of desire that makes it less desirable. For example, you might learn something about how animals are raised and slaughtered or see a video about the manufacture of some mass-made meat slurry and lose the desire you used to have for fast food burgers. The guy you've been flirting with might tell a racist joke and lose all appeal to you. It may dawn on you that horoscopes are hogwash, and your desire to know the astrological signs of those you meet evaporates.

It's not unheard of to become critical of your own desires and overthrow them. (“Wait a minute... why do I keep doomscrolling when it only makes me angry. I quit.”)

You might discover that you have been *mistaken* in your desires. For example: maybe I thought I desired the sweet taste of tobacco, and then one day I realize that all along what I really desired was the relief from nicotine withdrawal. Or, more subtly and psychoanalytically: I thought I really loved ice cream, but what I really desired was my mother's love which I'd learned to associate with ice cream. [We don't always know our own desires well](#) and shouldn't just take them at face value. We can become more self-aware about our desires: what do we *really* desire as opposed to what we have learned to assume is the medium through which we can fulfill that desire.

People sometimes use intemperate self-indulgence to self-medicate, to relieve stress, to self-comfort, or even to self-punish. In such cases the explicit desire is just an excuse that masks something under the surface.

Some desires seem to *infect* us to our detriment, and with work we can expel them. The nicotine addict who beats the addiction and comes to hate the smell of cigarettes, for example.

The sculpting of desire is the science of the advertiser and the propagandist. (“You must love Big Brother. It is not enough to obey him: you must love him.”^[1]) This also suggests both that it is possible to change our desires and that maybe we ought to be paying closer attention so that our desires are not changed on someone else's behalf.

How should we shape our desires

If we *can* shape our own desires, how can we know *how* to do so? Is there some objective metadesire we can use as our lodestone? Are there other criteria we can use? There have been many attempts to answer this question; here are a few:

A Buddhist perspective is that our desires are typically out-of-joint in this way: they aim for an unachievable stasis in which pleasant states are fixed in place and unpleasant states are kept forever at bay. This cannot be accomplished, and so if we stick with such desires we will always be frustrated, and thereby desireify ourselves into unhappiness. Pleasant states are all impermanent and will slip through our grasp no matter how tightly we hold onto them, and unpleasant states (e.g. sickness, old age, and death) are inevitable ingredients of our destinies. The key, then, is to really grok this flux and impermanence such that you replace your unrealistic desires with something more in tune with reality. Temperance in Buddhism culminates in nirvana: the quenching of desire.

Aristotle felt that the key was to align your desires with what is good for you. If you desire those things that make you a flourishing, thriving human being, then your desires will guide you to your best potential (at least to the extent that accidents of fate allow). These proper “objects” of desire, thought Aristotle, were not really objects at all, but personal characteristics: the virtues. The virtuous person acts virtuously because they *desire* the virtue (they don’t, in other words, act virtuously in a spirit of self-denial, or in opposition to their desires). He thought that it was difficult to mold desires correctly in people once they have hardened into adulthood, and that people do not necessarily hit upon the virtues instinctively on their own, and so it is important to educate children in the virtues from an early age. Temperance to Aristotle is the alignment of your desires with the virtues, and the pursuit of those desires in a measured way.

Stoic ethics was all about changing desires. We are frustrated in our desires, and thereby become disturbed, taught Epictetus, whenever we desire things that are outside of our control. When we desire the weather to be a certain way, or for the plague to spare us, or for some distant politician to make the right decision, or for someone else to love us, we set ourselves up for disappointment. Instead, he counseled us to accept everything that is outside of our control with absolute equanimity: this, exactly this, whatever this is, is the playing field I have been sent to play on. What is in my control is how I play, and there alone will I exercise my desire. Temperance to Epictetus might be summed up by the [Serenity Prayer](#): “God, grant me the serenity to accept the things I cannot change; the courage to change the things I can; and the wisdom to know the difference.”

Responding to the call of our desires appropriately

Having the right desires is part of temperance. Another part is responding proportionately and appropriately to the desires we have. Desire cocoa puffs if you must, but do not [go cuckoo](#) for them.

When we are very young, the circuit between desire and pursuit is very short. Part of maturity is to add more deliberation between desire and action. I want this... but do I *really* want it? do I want it more than what I would have to give up to obtain it? do I want to be the sort of person who indulges such a want? does this interfere with other things I want even more? ...okay then.

The word Aristotle used for “intemperate” was the same as the Greek word for “unchastened”—the intemperate person is in that way *immature*: they haven’t grown past the candy-grabbing stage of development.^[2]

But how do you go about maturing in this way if you didn't get the right tutelage as a child?

Some people have reported success using forms of mindfulness meditation to bring desire into focus and defang it: When you feel a desire, really stare it down. Examine it from all angles. How do you know it is a desire? How is it prodding you to act and what does *that* feel like? And so forth. Eventually under such a withering gaze, the desire may melt away or show itself to be not so very substantial as it was when you gave into it reflexively.

Sometimes people feel a disconnect between their short-term and long-term desires, where the immediacy of what short-term desires promise overwhelms the more delayed gratification of long-term desires when they conflict. I have heard of people using imagination to frequently make the results of long-term goals more salient, or visualization to model the process of spurning short-term temptations in the service of long-term goals (["implementation intentions"](#)), each of which may help you when temptation strikes.

Some people are more *impulsive* than others. They're more likely to indulge their whims right away rather than thinking it through first. I don't know much about impulse control issues, but it is something that psychologists study and treat, so if you think this is something you have trouble with, they might have the answers.

For particularly strong but unwanted desires, such as addictions and other obsessive behaviors, there are a variety of approaches. [Alcoholics Anonymous](#), for instance, has developed by trial-and-error over generations from the efforts of insightful and desperate people, and contains nuggets of hard-won wisdom about desire and our power over it that are worth careful study.

Our screen time (such as gaming and social media) is being deliberately engineered in increasingly addictive ways. If we want to lead healthy, fulfilling lives, we will need to attend more closely to the state of our desires, cravings, and habits.

(Some of this touches on the related virtue of [self-control](#), which I cover in a separate post.)

1. ^

George Orwell, *1984* (1949) part 3, section IV

2. ^

Aristotle, *Nicomachean Ethics*, book III, chapters 10-12

Notes on Fitness

This post examines the virtue of **fitness**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

Why is fitness a virtue?

This is a good time to restate what is meant by a virtue, in the classic, Aristotelian sense: A virtue is a person's characteristic habit that tends to promote or exhibit that person's human flourishing.

Flourishing is affected by lots of other things, including the actions of other people and the luck of the draw. But as far as characteristic habits go, those that have a typical tendency of promoting one's human flourishing are the virtues; those that typically degrade one's human flourishing are vices.

To the extent that you are sick, debilitated, or disabled, your human flourishing is that much less. This does not in itself say anything about your moral worth or virtuousness: being sick, debilitated, or disabled is not in itself blameworthy.

But, if you have characteristic habits that have a tendency to make you sick, debilitated, or disabled, those characteristic habits may be vices (assuming they don't have offsetting benefits in terms of human flourishing) and you would be wise to replace them with corresponding virtues: characteristic habits that promote wellness, strength, and capability.

Fitness as a virtue is not so much measured by how healthy you are so much as how healthy your *habits* are. Your health may be in large part a matter of what cards you were dealt, but you can play those cards well or poorly, and that's where the virtue of fitness comes in.

Related virtues

Some closely-related or synonymous names for this virtue are strength, vigor, hardiness, vitality, health, fettle, shape, well-being, and robustness.

In the context of [endurance](#), fitness goes by names like resilience, grit, toughness, and fortitude.

When fitness is reflected in action, it may be called energy, liveliness, or tirelessness.

Fitness comes to the aid of other virtues by giving you more strength, energy, and capacity to carry them out, while distracting you less with aches & pains & worries. It helps the intellectual virtues by keeping the mind sharp. As has become especially apparent during the CoViD epidemic, fitness also comes into play in the social virtues: the healthier you are, the less likely you are to directly be a burden to others or to the

health care system, and wise health practices make it less likely that you will contribute to the infection of others. Fitness is, in ways like this, [considerate](#).

How to improve the virtue of fitness

“[N]o pains, expense, self-denial, or restraint, to which we subject ourselves, for the sake of health, is too much. Whether it require us to relinquish lucrative situations, to abstain from favorite indulgences, to control [intemperate](#) passions, or undergo tedious regimens; whatever difficulties it lays us under, a man who pursues his happiness [rationally](#) and [resolutely](#), will be content to submit to.”

—William Paley^[1]

When we abruptly suffer through illness or accident, we remember the value of health and we would do a lot to get it back. But we’re less likely to apply this same calculus to efforts to build, preserve, and maintain health when we’re already doing okay, or when we’ve been declining more gradually.

This is a shame, for a number of reasons. For one thing, it’s often easier to avoid falling into a hole than to dig yourself out once you’ve fallen in. If you’re healthy-ish already, it’s not all that difficult to incorporate healthy exercise, good sleep, attention to diet, and so forth into your lifestyle. If you’re unhealthy, all of those things can be more of an uphill climb: you’re not firing on all cylinders, so everything takes more effort.

It can be surprisingly difficult to get good advice about fitness. There are lots of exercise, diet, and supplement fads out there, for example, and if you judge them based on how confidently their promoters tout them, they’ll all sound essential. You would probably be wise to focus first on things for which there is broad expert consensus about how valuable they are for the typical person. When it comes to the more controversial specifics, tread carefully and skeptically.

Without going too deep into specifics, the following are some of the categories of ways in which you can develop healthy habits of the sort that compose the virtue of fitness:

- **Diet.** Eat food with the range of nutrients your body needs, in a way that your digestive tract can manage. Avoid harmful foods, toxins, or excess. Be aware of any individual peculiarities you have (allergies, food intolerances) and work around them.
- **Food handling safety.** Develop an understanding of and adhere to good practices of food preparation and storage to avoid microbial contamination. This includes proper sanitation in food preparation, how to cook various types of food sufficiently, and how to store food safely in the short and long term.
- **Avoid poison.** Don’t smoke, drink little if any alcohol, beware of toxic pollutants, maintain clean air indoors, wear protection if handling toxins or if they’re in your environment. Understand the medications you take and their interactions with each other and with foods.
- **Water.** Drink enough water, from uncontaminated sources.
- **Weight.** Watch your weight. Make an effort to lose weight if you’re trending toward obese. Investigate unexplained abrupt weight changes.
- **Aerobic exercise.** Get some, [on the regular](#).
- **Movement.** Protect your range of motion and joint health by incorporating range-of-motion and flexibility exercises. Take care not to engage in repetitive

motion for excessive amounts of time, or to maintain a single (e.g. seated) posture for excessive durations.

- **Sunlight.** Get outdoors from time to time. Use good-quality sunscreen and in other ways take care to avoid deleterious effects from solar radiation.
- **Sleep.** Get sufficient, continuous sleep of good quality.
- **Stress.** Get yourself out of situations that cause chronic stress, and/or learn good stress reduction techniques.
- **Risks.** Take preventative steps to mitigate not-uncommon risks: e.g. wear a seat belt while in a car, mitigate tripping hazards around the home, don't do things with ladders that will wind up on YouTube, etc.
- **Travel.** Investigate possibly unfamiliar health risks of places you travel to (parasites? venomous animals? predators?).
- **Preventative health care.** Get your check-ups, have your teeth cleaned, do the blood tests your doctor orders, keep your vaccinations up to date, take time out to get those mammograms, colonoscopies, etc. on schedule.
- **Education.** Develop some basic understanding of biology, anatomy, physiology, metabolism, and so forth, and the terminology associated with those disciplines, so you have more insight into your health, and the vocabulary you need to describe it to specialists.
- **First Aid.** Have some understanding of how to deal with uncommon acute issues you might encounter, such as heat stroke or a bleeding wound.
- **Infection.** Behave sensibly during epidemics or around people with infectious diseases (mask-wearing, hand-washing, condoms, etc.). Use a bed net if you're in a malaria-prone region.
- **Basics.** Have a warm and dry place to sleep, clothing sufficient to protect you from the elements, shoes on your feet, a place to wash up.
- **Grooming.** Wash your hands regularly, bathe from time to time, brush your teeth, floss, trim or clean your nails, treat any superficial infections, abrasions, or cuts you have.
- **Friendship.** Make sure there is someone in your life you see frequently who you have the sort of relationship with such that they might say something like "you're looking pale today," "you've lost a lot of weight," "what's that thing on the back of your neck?" etc.

I encourage you to contribute more-specific advice, if there is any you feel you can vouch for, in the comments.

1. ^

William Paley, "Human Happiness" *The Principles of Moral and Political Philosophy* (1785)

Notes on Sincerity and such

This post examines a cluster of virtues that includes **straightforwardness, frankness, sincerity, earnestness, candor**, and **parrhēsia**. I hope it will help people who want to know more about these virtues and how to nurture them. I am a technical writer by trade and have developed some strong opinions about the value of, and the techniques of, clear and accurate communication, and so I will also draw on that experience to inform this post.

What are these virtues?

“[L]et sincerity and ingenuousness be thy refuge, rather than craft and falsehood: for cunning borders very near upon knavery. Wisdom never uses nor wants it. Cunning to the wise, is as an ape to a man.” —William Penn^[1]

These virtues have to do with communicating in a way that is clear, precise, efficient, and useful. They show respect for those you are communicating with by “giving it to them straight” and not forcing a lot of second-guessing and interpretation.

I briefly mentioned some of them in my post on the related virtue of [honesty](#), but now I want to look at them more closely.

In short, what these virtues have in common is “saying what you mean, and meaning what you say” (but also not talking a lot of rot that’s not to the point). If honesty covers “the truth,” the rest of these virtues help to cover “the whole truth, and nothing but the truth.”

Other closely related virtues include trustworthiness, reliability, and authenticity.

In opposition to these virtues are things like beating around the bush, candy-coating, ambiguity, euphemism, flattery, winks-and-nods, insinuations, exaggerations, incantations, ostentation, deflection, pretension, evasion, [false modesty](#), irony, sarcasm, manipulativeness, insincerity, flippancy (making light of serious matters), changing the subject, playing rhetorical [motte and bailey](#), or being “all hat and no cattle.”

There is some tension between these virtues and the virtues of tact and discretion (see below).

These virtues span a spectrum of outspokenness. On the more reserved end, you may rarely speak, but when you do, you speak sincerely and straightforwardly to the point. Towards the middle, you may try to anticipate what people would want to know and, with frankness and candor express this, warts and all, whether they ask or not. At the unreserved extreme, you may feel compelled to reveal those things that people *don’t* want to know but need to be confronted with: this is the *parrhēsia* that made the Cynic philosophers notorious (and sometimes unwelcome).

I think maybe if we all exhibited *parrhēsia* we’d get sick of it pretty quick, but in small doses it’s valuable. It’s the person with a bit of *parrhēsia* who is the first to call out someone on their racist joke or sexist assumption, or to mention the elephant in the

room, or to laugh at the [emperor's new clothes](#), or to confront someone about their drinking problem while everyone else keeps to the conspiracy of silence.

Being sincere isn't always about what you communicate, but sometimes about what you won't. If you feel the need to be mysterious, if you like to keep people guessing, if you present yourself as something of a code and judge your friends by their ability to crack it... well, you might want to consider [how to be more straightforward](#) instead.

But what about tact and discretion?

"[He] looked from me to the forms and back again, giving me the exact kind of smile of someone who, on Christmas morning, has just unwrapped an expensive present he already owns." —David Foster Wallace^[2]

Tact has to do with communicating in a way that will not hurt feelings, step on taboos, or in other ways be [impolite](#) or off-putting. Discretion can mean steering clear of topics that might raise hackles or open old wounds, or it can also mean keeping secrets and not being a blabbermouth about things that weren't your business to begin with.

These things seem at first glance to be in conflict with candor. One possibility is that they are, and that maybe this shows that tact, discretion, frankness, and candor are not all virtues after all. Another possibility is that they are all virtues, but that we should not always expect virtues to fit together flawlessly in a mutually-compatible way: they are after all not commandments handed down from on high, but merely generalizations about human character traits shaped by generations of folk psychology. Another possibility is that they are compatible after all, but that it takes a little extra discernment and skill to make them work together nicely.

Let's consider this last possibility:

There's a fine line between giving someone an unpleasant answer accompanied with "a spoonful of sugar to help the medicine go down" and giving them the answer so candy-coated that the bitter truth can no longer be tasted at all. Sometimes not talking about something is a peaceful way to agree to disagree, or to mind one's own business; other times it can be complicity with a foolish demand to ignore the elephant in the room.

This is to say that tact and discretion *can* certainly be deployed in the service of insincerity and cover-up; but that doesn't prove that they are necessarily always deployed that way. If you can learn to be tactful in a sincere way, to be candid about your discretion, you will have found a way to improve both sets of virtues.

Consider the phone-call-ending phrase "I guess I should let you get back to work." It usually means, more frankly, "I think we're done now; let's end this call." The first phrase is often (and usually pretty obviously) insincere, if only a little grating; the second phrase somehow seems too blunt, maybe a little rude, implying that you're eager to be rid of an unpleasant duty. I can't help but think that a skillful person, well-practiced in both sincerity and tact, could come up with more graceful ways to bring such a call to an end.

Consider also [this remarkable essay by philosopher Agnes Callard](#). She writes of some sort of trauma she endured long ago, and of some sort of neuroatypicality she

experiences, but she [steadfastly](#) refuses to give names to either of those things. She wants to talk about them, but if she names them, she suspects we will use those names to apply a familiar template to her and her experiences, and then we will interpret what she says according to that template.

“And that means I can’t talk to you. No one can sincerely assert words whose meaning she knows will be garbled by the lexicon of her interlocutor....

“[I]t chafes at me that you have decided that if I want to talk... with you, I have to follow your rules, and let you trample all over me.”^[3]

Callard here uses her discretion, even her blunt *lack of candor*, to be straightforward and sincere in a way that she feels would be otherwise unavailable to her.

But what about irony, sarcasm, and stuff like that?

“When people speak in a very elaborate and sophisticated way, they either want to tell a lie, or to admire themselves. You should not believe such people. Good speech is always clear, clever, and understood by all.” —Leo Tolstoy^[4]

People use irony and sarcasm, understatement and hyperbole, parody and caricature, modest proposals and other such rhetorical devices to express themselves creatively in different registers. This can be entertaining, witty, and clever of course, but also sometimes insightful and poignant and biting in a way that would be difficult to match with more straightforward ways of speaking.

This raises the question of whether the virtues in this collection are ones that threaten to make us dull and to limit our expressive range.

If you are hesitant to give up these shades of your conversational palette, consider instead that there may be better and worse ways to use them. For example, if you speak ironically but in a way that is understood as such by those you are speaking with, that’s very different from speaking ironically in a way that some of your audience gets, and gets to feel superior about, at the expense of those whose heads you’re speaking over.

If you use hyperbole in a fun way, as another form of shared-irony, that may be innocent enough. But if you use it excessively or unthinkingly — if you seemingly can never be concerned by what you see in the news without being “deeply troubled” by it, for instance — consider recalibrating your verbiage.

I suspect that this is one of those cases where we will have to rely on the spirit of the law rather than any firm prohibitions. Use these rhetorical registers, but use them carefully, and question your motives for using them. If you resort to caricature to make a joke or simplify an example, maybe that’s all fun and games, but if you use it to reinforce a stereotype or to paint a grey area in black-and-white to hide its complexity, consider that you may have taken a step too far.

There is a lot of rotten use of this stuff going around these days. Someone says something insincere and offensive, and the next thing out of their mouths is something along the lines of “I was only putting it out there,” “That’s just what some

people are saying,” “I only wanted to see how you’d react,” “Hah; you thought I was being literal,” “I sure triggered the outgroup with that!” Don’t be that guy.

But what about flirtation?

It were as
possible for me to say I loved nothing so well as
you: but believe me not; and yet I lie not; I
confess nothing, nor I deny nothing.

—Beatrice, in *Much Ado About Nothing* [5]

Flirtation is a form of communication that is indirect and ambiguous and that stubbornly talks around the main topic without addressing it directly. You remain coy and veiled, hinting and feinting at your intentions and feelings rather than stating them outright. Rather than being sincere and straightforward, you create and amplify uncertainties. Irony abounds and the earnest are out of their depth: Playing hard-to-get is sometimes the only way to get got.

Flirtation has been likened to a game. As in a game, the players are *sincere* in trying to play well; but the game itself is a sort of make-believe. The moves in the game refer to the game, and not to the world outside. (The hockey player does not play the game in order to put the puck in the net, but puts the puck in the net because that is how you play the game.)

Telling someone who is flirting that they ought to be more straightforward and candid is like telling a courting peacock he’ll never get off the ground flapping his tail feathers like that, or like telling someone playing Monopoly that they should probably switch to real money if they want to have any hope of buying seaside property in New Jersey. It’s missing the point.

My guess is that flirting of this sort is ubiquitous, cross-cultural, and ancient. But I may be overgeneralizing from the culture I grew up with and what it has found meaningful from other cultures and times. If flirting is one of the essential things human language is for, and it operates by its own set of rules, perhaps it is best to sandbox it appropriately, like our other games, and play it as best we can. But if flirting is merely a sort of inessential insincerity we sometimes allow ourselves to indulge in during courtship, maybe we should see if we can disarm it with surprising candor and begin our romances on a more sincere note. That [might work out better](#). To me, it’s an open question. I’ve experimented with both modes, and with many experiments under my belt by now, all I can say for certain is that you’d be a fool to take any romance and relationship advice from me.

But these days, with an increasing percentage of couples meeting through on-line dating sites, [6] it seems that at least some flirting has become moot. There’s little point in being coy with the person you’ve met through looking-for-love-dot-com. The secret’s out.

I had finished writing this section before I remembered the whole pickup artist scene and its ruthless cultivation of insincerity in pursuit of the ol’ in-and-out. I don’t want to dwell on it, but one way of looking at it is that it just takes some of this logic of flirtation to extreme conclusions: if flirtation includes pretension, misdirection, flattery, and the like, why not just declare no-holds-barred and play to win on your terms?

But what about culture jamming?

For a long while I was a fan, scholar, and practitioner of “culture jamming,” and I admiringly cataloged historical pranks, hoaxes, frauds, impostures, counterfeits, tricksters, trolls, fakery, hacktivism, performance art, fauxvertising, forgeries, scams, modest proposals, and things of that sort at a 20th century web 1.0 site I called [Sniggle.net: The Culture Jammer’s Encyclopedia](#).

I explained:

Most of this site highlights deception, but it’s not because I have a thing for liars and cheats. I think there’s a brand of immunizing deception that helps us to expose and correct the lies we tell ourselves and the webs of falsehood that make up our societies. Harmless fibs can remind us that we’ve dropped our guard and let the Big Lies in. [\[7\]](#)

In an interview I doubled down on the therapeutic explanation of culture jamming:

A whole lot of the evil of the last century was conducted by people who followed rather sheep-like the twisted consensus reality of their societies. What the trickster does is to find flaws in that consensus reality and to construct creative performances to exploit and uncover those flaws. If this happens enough, perhaps people will come to develop an instinctive distrust of consensus reality and will be more likely to see reality as it is. [\[8\]](#)

I don’t think time has been kind to my theories. The increasing prominence of outrageous lies and fakery has not immunized people against their effects or caused people to become more vigilant, as I had naïvely hoped might be the case. Instead of the collapse of illegitimate authority causing people to think for themselves, people seem to have responded by saying “well, I guess there’s no way to know *what* nonsense to believe, so I’ll just believe whichever nonsense I like the best.”

I still think there is something useful in the art of culture jamming, and I still admire a well-delivered hoax in a good spirit. I wasn’t *all* wrong, I don’t think. But today, older and wiser (I hope), I choose to err on the side of sincerity instead.

But what about framing?

“Political language — and with variations this is true of all political parties, from Conservatives to Anarchists — is designed to make lies sound truthful and murder respectable, and to give an appearance of solidity to pure wind.” —George Orwell [\[9\]](#)

Framing (or “spin”) is the attempt to fit revealed or asserted facts into a rhetorical framework in such a way that they will lead people to desired conclusions or away from undesirable ones. When this is called “spin” it usually implies purposeful dishonesty; when it is called “framing” its proponents sometimes claim it can be done in the service of clarity and honesty, or to defend against spin.

However, I more often see framing deployed as a way of trying to manipulate the audience into adopting a certain belief without putting that belief forward and

defending it explicitly: in other words, as just spin with a new name. Here, for example, is George Lakoff, a scholar of framing, explaining a textbook example:

The phrase “Tax relief” began coming out of the White House starting on the very day of Bush’s inauguration. It got picked up by the newspapers as if it were a neutral term, which it is not. First, you have the frame for “relief.” For there to be relief, there has to be an affliction, an afflicted party, somebody who administers the relief, and an act in which you are relieved of the affliction. The reliever is the hero, and anybody who tries to stop them is the bad guy intent on keeping the affliction going. So, add “tax” to “relief” and you get a metaphor that taxation is an affliction, and anybody against relieving this affliction is a villain.[\[10\]](#)

So he recommends this frame instead:

It is an issue of patriotism! Are you paying your dues, or are you trying to get something for free at the expense of your country? It’s about being a member. People pay a membership fee to join a country club, for which they get to use the swimming pool and the golf course. But they didn’t pay for them in their membership. They were built and paid for by other people and by this collectivity. It’s the same thing with our country — the country as country club, being a member of a remarkable nation.[\[10\]](#)

This is just trading one unstated and poorly-defended set of background assumptions for another. To me, describing tax cuts as “relief” seems more accurate and honest than describing them as “reduced country club membership fees,” but they’re both examples of manipulative spin. A more straightforward and respectful way of discussing this issue would be simply to describe the proposed tax law changes, what effects they could reasonably be expected to have, and whether or not you think that would be a good thing.

This then brings up the larger debate — which I don’t want to wade into now — about whether one has to leave one’s virtues behind and fight dirty when one steps into the ring of political action.

But to sum up: If you are thinking of your argument in terms of what framing and spin to apply to it, you’ve probably left the field of candor, frankness, sincerity, and so forth, and it’s time to take a U-turn if you want those virtues back.

Manipulative framing and other forms of insincerity are so ubiquitous in political and culture-war discussions that I think it’s usually best to avoid them entirely so as not to be infected. Mute every pundit. Kill your television. Ignore politicians. “Read not the *Times*. Read the Eternities.”[\[11\]](#)

Appendix: The trouble with passive voice sentences

Professionally I am a technical writing consultant in the software industry. A company will sometimes call me in to fix things when nobody can understand the documentation that was written by the engineers who created the software. If I have time, one of the things I will do for such a company is write up (or occasionally present a class on) “technical writing for software engineers in one easy lesson.”

That lesson is essentially this: *hunt down and kill every passive voice sentence*. If they can grok that, they're 80% of the way to being able to write reasonable technical documentation themselves.

Quick grammar review: An *active voice* sentence has a subject, verb, and object: “Elon unwisely tweeted insider information.” (Imperative sentences are an exception: they do not have an explicit subject, but the subject is implicitly whomever the sentence is uttered to.) A *passive voice* sentence may leave out the subject entirely: “Insider information was unwisely tweeted.” A missing subject means missing information, which means ambiguity, which makes trouble for precise technical communication.

“The transmit box should be checked.” Who is responsible for checking it? Do I need to check it somehow? or ought I to make sure someone else has checked it? or is it not being checked a suspicious error condition I should be on guard for? There is no way for me to know based on that passive voice description.

Passive voice sentences are often accurate without being complete or precise. Their accuracy makes them seem non-problematic to those who write them, while their imprecision makes them non-helpful to those who read them. People often write in the passive voice out of laziness, or because it can sound a little more formal and academic and so has a false air of sophistication to it. But people sometimes also use it to hide their ignorance or to sweep things under the rug.

The so-called [“past exonerative tense”](#) usually takes advantage of the blame-dodging obscurity of the passive voice. It’s most notoriously put to use by police department press releases that test the limits of grammar to avoid straightforwardly describing officer misconduct.^[12] “Handcuffed Man Fatally Shot Inside Police Cruiser” reads the headline. “The man was shot multiple times by the officer’s service weapon, the police spokeswoman said.”^[13]

In summary: I recommend that you look very skeptically on any passive-voice sentence you intend to deploy that obscures the subject — not just in technical writing, but in any communication that you want to be clear and precise.

1. ^

William Penn, *Some Fruits of Solitude In Reflections And Maxims* (1682)

2. ^

David Foster Wallace, *The Pale King* (2011)

3. ^

Agnes Callard, [“I Don’t Want You to ‘Believe’ Me. I Want You to Listen.”](#) *New York Times* (30 November 2020)

4. ^

Leo Tolstoy, *A Calendar of Wisdom* (1910), February 15

5. ^

William Shakespeare, *Much Ado About Nothing*, act IV

6. ^

Alex Shashkevich, [“Meeting online has become the most popular way U.S. couples connect, Stanford sociologist finds”](#) *Stanford News* (21 August 2019)

7. ^

David Gross [“About sniggle.net”](#)

8. ^

David Gross [“Interview With a Sniggler”](#) (ca. 2006)

9. ^

George Orwell, [“Politics and the English Language”](#) *Horizon* (1946)

10. ^

Bonnie Azab Powell, [“Framing the issues: UC Berkeley professor George Lakoff tells how conservatives use language to dominate politics”](#) *UC Berkeley NewsCenter* (27 October 2003)

11. ^

H.D. Thoreau, [“Life Without Principle”](#) *The Atlantic* (October 1863)

12. ^

Radley Balko, [“The curious grammar of police shootings”](#) *Washington Post* 14 July 2014

13. ^

[“Handcuffed Man William Green Fatally Shot Inside Prince George’s County Police Ofc. Michael Owen’s Cruiser”](#) *CBS Baltimore*, 27 January 2020

Notes on Justice (as a virtue)

This post examines the virtue of **justice**. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

Defining what justice *is* is a major topic of ethical philosophy and of political philosophy. There is no way I can do justice (see what I did there?) in a brief post like this one. Instead I want to concentrate on the *virtue* of justice. So, setting aside for now what justice *is* or what it *consists of*, what does it mean to be characteristically, habitually just: to exhibit the virtue of justice?

Justice as a virtue

Justice is usually thought of as something that exists *between* people: as a balance, with opposing claims of different people weighed against each other.

Justice, as a virtue, is different. It is something within a particular individual (though it often exhibits itself through how that individual behaves towards others).

It is not *contractarian* in the social-contract sense of being willing to submit to the restraint of justice in exchange for the security of knowing that your neighbors are similarly restrained. It is also not merely personal *obedience* or *acquiescence* in the face of a particular standard of justice.

In a person without the virtue of justice, justice is a double-edged sword. If that person is cheated unjustly, they may hope they can appeal to justice to right the wrong. But if that person ends up on the favorable side of a shady deal, they may hope that justice remains thwarted so that they can get away with it.

A person with the virtue of justice evaluates such situations differently. To such a person, justice itself is something they value: the life they consider to be most worth living is a just life. The possession of a just character is worth more to them than the possession of things successfully swindled away from someone else unjustly. Because of this desire to be just, the person with the virtue of justice has developed the habit of characteristically seeking out just outcomes.

An example of someone exhibiting the virtue of justice would be someone who notices that they were inadvertently handed a \$20 instead of a \$10 in change at the cafe that morning, so they return at lunchtime to give back the extra money. Even though they did nothing wrong, even though nobody but them knows of the unjust state of affairs, even though that injustice worked to their material advantage, they feel the desire to set things right and are willing to go out of their way to do so.

It is not that the just person does not value money, but that they value “money acquired by just means” as a more richly-defined package deal. One way of understanding this might be to compare it to the difference between having a trophy on your shelf that you had won compared with one that you had purchased, borrowed, or stolen. It isn’t the trophy merely, but the trophy-earned-in-victory, that is important.

This works the other way around too. Someone who does not value justice may pause to think, when they have been swindled in some small way, that it would be too much

bother to fight back: the cost/benefit is too high. Someone who loves justice may fight back “just for the principle of the thing” because the cause of justice is itself a benefit worth fighting for, which changes the cost/benefit ratio.

The person with the virtue of justice also may approve of just outcomes and just behaviors (and disapprove of their opposites) even when they do not concern that person directly. They are disgusted or angered at the sight of injustice, and feel the urge to step in and do something about it. Cicero, for example, thought “[t]here are... two kinds of injustice: the one, on the part of those who inflict wrong, the other on the part of those who, when they can, do not shield from wrong those upon whom it is being inflicted.” So “he who does not prevent or oppose wrong, if he can, is just as guilty of wrong as if he deserted his parents or his friends or his country.”^[1]

Related virtues

There are a number of virtues that are closely related to justice. These include:

- [censure / chastisement / judgment](#) — the willingness to call out injustice in others; [righteous anger](#) is another form this can take.
- [epieikeia](#) (sometimes translated “equity”) — a corrective to inflexible letter-of-the-law justice that makes it conform better to the spirit of the law. Aristotle put it this way: “It is equity to pardon human failings, and to look to the [intentions of the] lawgiver and not to the law; to the spirit and not to the letter; to the intention and not to the action; to the whole and not to the part; to the character of the actor in the long run and not in the present moment; to remember the good rather than evil, and good that one has received, rather than good that one has done; to bear being injured; to wish to settle a matter by words rather than by deeds; lastly, to prefer arbitration to judgment, for the arbitrator sees what is equitable, but the judge only the law, and for this an arbitrator was first appointed, in order that equity might flourish.”^[2]
- [fairness / impartiality](#) (a.k.a. objectivity?)
- [sportsmanship](#) — defending “fair play” even at the expense of advantage or victory.
- Justice sometimes gets packaged up into such things as “probity,” “[honor](#),” “chivalry,” “rectitude,” “propriety,” or “righteousness.”
- To do justice *well*, you need virtues like wisdom, discernment, perspective, [phrónēsis](#), the ability to see things from different people’s points of view, and so forth.
- [rectificatory justice](#) has to do with the art of putting things right again when injustice has unbalanced the scales (amends, atonement, recompense, rectification, redress, reparation, restitution, step #9 of the Twelve Steps^[3]).

Obstacles to being just

Greed and ambition are common obstacles to justice ([see Cicero again for a good discussion of this](#)). If you crave material goods, power, position, or fame so much that you are willing to get them unjustly, good luck to you in trying to be just.

This suggests that the virtue of [temperance](#) (having desires of the right sort, and calibrating them well) is important to the exercise of justice. In addition, a sudden and vivid temptation may overwhelm you in the heat of the moment and cause you to

make an unjust [decision](#), even if, if you had time to carefully weigh the temptation against justice, you would have decided more justly. If this is the case, rather than temperance, the virtue of [self control \(restraint, continence\)](#) is what you need. This can be a challenge in particular for people who easily fly into a rage and make rash choices while angry (in which case the discipline of [anger management](#) may be helpful).

In general, if you value anything more than you value justice, if you are given the opportunity to obtain that valued thing but only by unjust means, it will be difficult to resist that temptation. For example: What if you value [honesty](#), but honesty would require you to give truthful testimony that would hurt the case of the more just party in a lawsuit? What if you value filial piety, but justice would require you to side against your parents and with someone they have wronged? Virtues that make incompatible demands are the stuff of [tragedy](#) because, apparently, they are the stuff of life as well.

Someone who values justice above all else, particularly someone whose sense of justice is unrefined by *epeikeia* and untempered by mercy, is sometimes depicted as inhumane and harmful. [Javert](#) from *Les Misérables* is one good example of this.

Rather than thinking of justice as a virtue you either have or don't have, maybe it makes more sense to think of it as a value you rank more or less highly among your values.

If you think you are up against an unjust adversary or are making bargains with an unjust person, you may be tempted to be unjust yourself in defense, or even in offense ("screw them before they have a chance to screw me").

You may be tempted away from justice by leniency, by squeamishness about being harsh or judgmental, by being a nebbish and letting people walk all over you, or by preferring to complain about how unjustly you have been treated than to see to it that you are treated justly. You may prioritize forgiveness and charity to the extent that justice decays. On the other hand, the intoxication of righteous anger can also lead to unjust excesses of vengeance. [Aristotle's "golden mean"](#) theory seems to come into play here.

"Someone asked: 'What do you think about the principle of rewarding enmity with [kindness](#)?'

" 'With what, then, would you reward kindness?' asked the Master. 'Reward enmity with just treatment, and kindness with kindness.' " —Confucius^[4]

Finally, one may inadvertently be unjust by being deceived by irrelevancies or cognitive biases. My favorite example of this comes from economist Thomas Schelling, who used to present his students with two hypothetical tax code examples: one in which couples with children qualify for a tax credit, and another in which couples without children are charged an additional surtax. He would ask, in the first case, would it be just to give poor families a bigger credit than rich families? and in the second case, would it be just to make the childless poor pay a bigger surtax than the childless rich? Students typically thought that the first hypothetical would be just, but the second would be unjust. But if you do the math, you can see that these are just different *descriptions* of the exact same outcomes. [See this page for the details.](#)

That said, there is a big difference between having the right understanding of what is just and having the virtue of justice. Aristotle put it this way:^[5] A science concerns a

subject matter in which your knowledge and skill can help you aim for opposite extremes: for instance, a doctor knows the science of health, and this knowledge would be equally useful to her in healing someone and in harming them. A virtue, on the other hand, goes in only one direction — having the virtue of [courage](#) doesn't make it easier for you to be cowardly, for instance. Aristotle asserts that justice is a virtue in this sense, rather than a science. So justice can't be summed up by learning the rules of what makes one thing just and another thing unjust (which, presumably, could help you do either one were you so inclined, and which, incidentally, seems to be most of what ethical philosophy concerns itself with these days). Instead it is about having just desires and doing just acts because you have a just character. However, Aristotle also devoted a book of his *Nicomachean Ethics* to the intricacies of the science-part of justice — the study of what is and isn't just — and he said that he thought of it as a more complex subject, more difficult to master, than medicine.

Justice as a social benefit

If you are just and I am just, it is relatively frictionless for us to cooperate. A handshake-deal will do. We don't need to consult lawyers, draw up contracts, keep a close eye on each other, and so forth.

On the other hand, you can get a large temporary advantage by pulling a fast one on someone who thought you were a just person and dealt with you based on that assumption. (This may be another reason why just people do not merely behave justly towards others, but are willing to go to the mat to enforce justice against others who have behaved unjustly.) But that burns the bridge behind you, at least with that person.

If there is no mutual assumption of just character at all, then some sorts of cooperation are off the table entirely and others become considerably more costly.

In short, there seems to be a sort of [prisoner's dilemma](#) at work. [David Hume](#) thought that this is how justice emerged: people thought through this prisoner's dilemma, realized the mutual advantages of justice as a social norm, and so began to enforce it in various ways. [\[6\]](#)

Why be just?

The classic argument against the virtue of justice is to say that there is no good reason to value justice itself. If having a reputation for justice makes you more esteemed, makes people [trust](#) you more, makes it easier to make business deals, and so forth, that's nice, but it's *those* things that are valuable, not the justice itself. If you could get those things more cheaply by only pretending to be just, and then taking unjust advantage when you can get away with it, in this view, you ought to do so.

If you say you value certain things and want to obtain or effect them, *but only if you can do so justly*, you just make those values harder to obtain and so unnecessarily frustrate yourself. The unjust person, without those hobbles, is more apt to achieve their values.

This, or something like it, was the argument of Glaucon in the *Republic*, who put forward the ["Ring of Gyges"](#) thought experiment to drive the point home (imagine having a ring of invisibility, while wearing which you could get away with all sorts of

mischief without getting found out).^[7] He also asked us to consider an entirely just man who has nonetheless through some misfortune acquired the reputation of a scoundrel, in contrast with his opposite: a completely unjust man who has managed to convince everyone that he is pure as snow. Who is more fortunate? Has the just man's embodiment of his ideal of justice done him any good? does it give him any comfort? Is the unjust man's lack of justice any problem for him?

The difficulty of showing that justice can be an ultimate value is, I think, just a special case of the general difficulty of providing reasons for choosing ultimate values. If you value wealth and admiration, say, and tell me that because justice can get in the way of this justice isn't itself very worthwhile, can't I just turn this around and ask you why you value wealth and admiration? Don't you really want what wealth and admiration permit you to obtain, such as material resources? And aren't those really valuable in order to obtain pleasure and ward off pain? Maybe hedonism is your only valid value? But what do pleasure and pain matter in the big scheme of things? Doesn't the lowliest worm wriggle around in pursuit of pleasure and aversion to pain? Don't you want your life to mean something more than that? And so you go round and round, justifying one value on the basis of another, trying in vain to find some value that logically ends the sequence and gives you a firm foundation.

It seems you have to *choose* rather than *discover* the values by which you will live your life. If you don't choose deliberately, you'll just be tugged back and forth by circumstances, urges, and whims. If you do choose, you have to decide. The virtue traditions serve as good menus of values to choose from, and justice is one that comes with some strong endorsements.

1. ^

Marcus Tullius Cicero, *De Officiis* I.7

2. ^

Aristotle, *Rhetoric*, book I, section 13

3. ^

"[We] Made direct amends to such people wherever possible, except when to do so would injure them or others."

4. ^

Analects of Confucius, XIV.xxxvi

5. ^

Aristotle, *Nicomachean Ethics*, book V, section 1

6. ^

David Hume, *A Treatise of Human Nature* (1739-1740)

7. ^

Plato, *Republic*, book II

Notes on Industriousness

This post examines the virtue of **industry** (a.k.a. “assiduity,” “enterprise,” “industriousness,” or “[productivity](#)”). It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I’ve been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope this post (and your comments on it) will be helpful to people who want to know more about this virtue and how to nurture it.

“Lose no time: be always employed in something useful; cut off all unnecessary actions.” —Benjamin Franklin[\[1\]](#)

What is industry?

Industry is the virtue of habitually devoting one’s sustained effort earnestly toward worthwhile goals. It is contrasted with such things as idleness/laziness/sloth, but also with things like aimlessness, dabbling, dithering, flitting, being ironic (un-earnest), triviality, and wasting time. The self-motivating impulse behind industry is sometimes distinguished from industry itself and called by names like “drive,” “eagerness,” “enthusiasm,” or “[ambition](#).”

“There is a perennial nobleness, and even sacredness, in work. Were he never so benighted, forgetful of his high calling, there is always hope in a man that actually and earnestly works: in idleness alone there is perpetual despair.” —Thomas Carlyle[\[2\]](#)

Related virtues

Some virtues that are in the same ballpark as industry include:

- [Duty](#) (responsibility, diligence): understanding what you *ought* to be doing, and then devoting yourself to that in particular.
- [Engagement](#): participating earnestly in life, rather than being a mere spectator.
- [Focus](#): resisting distraction and maintaining a goal over time (though for some types of goal, being able to multi-task and rapidly shift focus is important).
- [Purposefulness](#): a belief that your endeavors matter, a sense of mission or calling.
- [Service](#): exerting oneself for the sake of something beyond oneself.

The virtue of [determination](#) / follow-through is also associated with industry, but it can sometimes be counterproductive if taken to extremes. If you are determined to push something through to the finish line no matter what, you may be exhibiting a stubbornness or bullheadedness that interferes with industry: you continue to devote yourself to an unproductive enterprise because you’re unwilling to give up or to change your mind or because the sunk cost fallacy has captured you.

“Never shirk the proper dispatch of your duty, no matter if you are freezing or hot, groggy or well-rested, vilified or praised, not even if dying or pressed by other

demands. Even dying is one of the important assignments of life and, in this as in all else, make the most of your resources to do well the duty at hand.” —Marcus Aurelius^[3]

What is industry good for?

Most obviously, industry helps you achieve goals and accomplish things that you value.

It is also sometimes supposed to have some inherent value, aside from the value of whatever it is that it helps you to do or produce.

Purposeful, earnest action can contribute to a sense that your life has meaning. If you are idle and aimless, the questions of “what am I doing here? what is the point of my being alive?” are more likely to gnaw at you.

“The sleep of a laboring man is sweet, whether he eat little or much...”
(Ecclesiastes 5:12)

People seem commonly driven to do *something* goal-oriented. If you do not choose something meaningful to do, you may spin your wheels doing something just to pass the time.

“[W]ork... is a necessary condition of life... If a man does not work at necessary and good things, then he will work at unnecessary and stupid things.” —Tolstoy^[4]

If you are idle and lazy, work is *hard* and *daunting*. You may find that just doing the basic tasks you need to do to get through the day is irritating or difficult. By contrast, if you habitually put your shoulder to the wheel and apply yourself to an enterprise you find valuable, the more mundane chores of life are less likely to bother you.

“Whether you like to or not, acquire the habit of working hard, then you won’t have to work hard. Idleness does not make work easy, it ensures that work will be hard.” —Crates^[5]

Purposeful effort can itself be pleasurable. It is enjoyable and invigorating to be “in the zone” — applying yourself thoroughly to the task at hand.

“[T]hose pleasures are most valuable, not which is most excellent in the fruition, but which are most productive of engagement and activity in the pursuit.”
—William Paley^[6]

How can you develop industry?

“[L]et him who gropes painfully in darkness or uncertain light, and prays vehemently that the dawn may ripen into day, lay this other precept well to heart, which to me was of invaluable service: ‘Do the duty which lies nearest thee,’ which thou knowest to be a Duty! Thy second duty will already have become clearer.” —Thomas Carlyle^[7]

It is hard to be industrious when you are mentally or physically fatigued. If fatigue makes it difficult for you to apply yourself, you may need to attend to physical [fitness](#),

general health, nutrition, sleep, and exercise. [Stress](#), depression, and other psychological ills can also make it difficult to concentrate or to exert productive effort. Some drugs can mimic or exacerbate the effects of fatigue (others, such as caffeine, modafinil, and amphetamines can at least temporarily alleviate them; and some people swear by more exotic nootropics for this purpose).

You may find it easier to accomplish things if your life is more orderly. [Some attention to tidiness might increase your effectiveness.](#) Some [time management tips](#) might also help.

A common complaint these days is of short attention span, and a feeling that thousands of things are [tugging at us all the time](#) to try to get our attention. It is hard to roll up your sleeves and get down to business when a vibrating phone, blinking icon, and clanging inbox are all lining up to take swings at you. So many global crises really are very important and really do demand attention, but it is hardly possible to even stay up-to-date on all of them, much less address them with any competence, unless you narrow your focus and allow some of them to remain ignored.

You may want to [reconsider the utility of some of the distractions you have invited into your life](#). How much is all of that doom-scrolling doing for you? Could you maybe leave that next flame war unfought? Do you really need to make it to the boss room of Ultimate Myth Hammer XVII? Is all of that trivia-collecting about your favorite band / celeb / sports team / whatever really the best use of your time? Perhaps a few moments installing ad-blockers and tightening up app notification settings and abandoning dead-end screen-habits and played-out hobbies today would save you lots of future scattered thinking and allow you to set your sights higher.

Sometimes we unwisely use distractions as more-or-less deliberate techniques of [procrastination](#) (and then blame the distractions for it).

Warren Buffet had this bit of advice:^[8] Make a list of your top 25 goals. Then pick out the most important five of those. Then avoid putting in any work at all on the other twenty things until you've taken care of those five.

Ivy Lee had a different technique but one that seems to have a similar aim:^[9] Each day, start by listing the top six things you have to do, ranking them by their importance. Begin working only on the most important one (not for example the easiest, most time-sensitive, most well-understood, or some other such tempting heuristic), and keep at it until you're done, and only then move on to the next one.

There are an asston of such simple-secrets-of-titans-of-industry tricks (and similar "Getting It Done" advice) out there, though, and it's hard to tell if they're genuinely useful or just a genre of stories. On the one hand, advice is so bountiful that if you wade into it you're eventually bound to stumble upon something you'll find useful. On the other hand, there's just so much to wade through.

One piece of advice that seemed sensible to me is that you periodically remind yourself of your big, important goals when you've been spending a lot of time buried in the nitty-gritty of some difficult subtask. "Keep your eyes on the prize," as the song goes.

1. ^

Benjamin Franklin, *Autobiography* (1791)

2. [^](#)
Thomas Carlyle, *Past and Present* (1843), chapter XI: “Labour”
3. [^](#)
Marcus Aurelius, *Meditations* VI.2
4. [^](#)
Leo Tolstoy, *A Calendar of Wisdom* (1910), March 7
5. [^](#)
Crates of Thebes, *letter to Hermaiscus*
6. [^](#)
William Paley, *The Principles of Moral and Political Philosophy* (1785), chapter VI: “Human Happiness”
7. [^](#)
Thomas Carlyle, *Sartor Resartus* (1836), chapter IX: “The Everlasting Yea”
8. [^](#)
James Clear [“Warren Buffett’s ‘2 List’ Strategy: How to Maximize Your Focus and Master Your Priorities”](#)
9. [^](#)
James Clear [“The Ivy Lee Method: The Daily Routine Experts Recommend for Peak Productivity”](#)

Notes on Duty

This post examines the virtue of **duty**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What good is duty?

"If we have our own why of life, we shall get along with almost any how." —Nietzsche^[1]

A duty is often described as something of a burden (e.g. in phrases like "it is my unfortunate duty...", "duty-bound"). A duty can be a sort of lingering, unfinished business: When one "discharges" a duty, it implies a sort of relief.

So it's worth asking: in what sense is duty a virtue? Why is it a good thing you would want to seek out in your life?

To say that you have a duty is another way of saying that you have a *purpose* and that that purpose involves your agency (it's not a passive purpose). To lack a purpose of this sort can lead to despair, while having a purpose can contribute to [hope](#).^[2] People who do not know their duty are more likely to be plagued by questions like "what am I here for?" "why bother to get up in the morning?" They may wallow in frivolity or a "bullshit job," wondering vaguely what is the point of it all.

As a simple example of how duty can improve well-being, consider this study:

Some [elderly nursing home] residents (the experimental group) were given a plant to take care of, while others (the comparison group) were also given a plant, but were told *the staff* would take care of it. After three weeks the experimental group had significantly higher wellbeing levels, including sociability, vigour, and self-initiative.^[3]

What is duty?

"Many persons have a wrong idea of what constitutes true happiness. It is not attained through self-gratification but through fidelity to a worthy purpose." —Helen Keller^[4]

The word "duty" derives etymologically from words meaning "due" and "debt" — it suggests something that is owed, an obligation. It usually means that some sort of *service* is owed (as opposed to some sort of thing, or some cash equivalent, except in some specific usages like "customs duties"). If someone is asked "what is your duty?" they typically respond with a verb-phrase: a description of an action.

If you have a duty, it is your *responsibility* to see that the duty is carried out. Duty and responsibility overlap, and may be nearly synonymous. "Duty" has more of a sacred ring to it, with "responsibility" being more down-to-earth and secular (responsibility also has a backwards-looking sense to it that duty doesn't have, as in "who was responsible for putting that there?"). A dutiful or responsible person is one who is characteristically diligent in carrying out her or his duties, and this is where duty-as-a-virtue comes in.

There are two parts to being dutiful: 1) knowing your duty, and 2) carrying out your duty conscientiously.

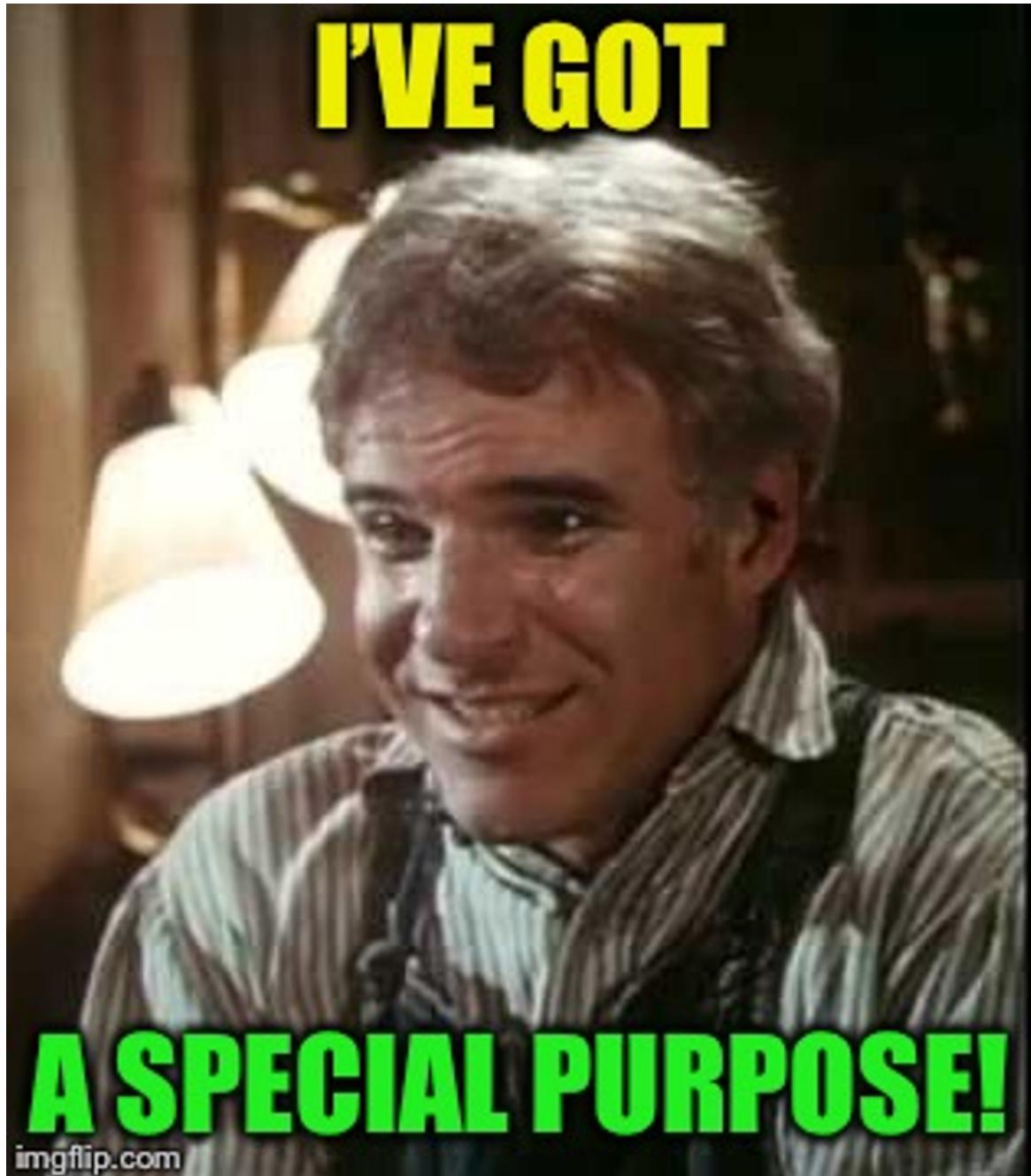
Knowing your duty

"Noble examples stir us up to noble actions; and the very history of large and public souls inspires a man with generous thoughts. It makes a man long to be in action, and doing something that the world may be the better for; as protecting the weak, delivering the oppressed, punishing the insolent." —Seneca^[5]

How do you know your duty? Sometimes you take on duties explicitly, for instance if you offer to take care of a neighbor's dog while they're on vacation, or if you take a vow or make a promise. Professions and offices typically come with duties inseparably attached to them: if you take on such a role (e.g. lawyer), you take the associated duties too (e.g. faithfully defend your client's legal interests). Religions often come packaged with certain duties (e.g. Mormon missionary service, the Muslim hajj, Catholic confession); "penance" is a variety of religious duty designed to expiate sins.

You may voluntarily take on a duty from a sense of engagement or purposefulness. This is a way of unreservedly jumping into life as a *participant* rather than a by-stander, and means that you have to put philosophical skepticism aside and commit yourself. Krishna's counsel to Arjuna in the *Bhagavad Gita* might be interpreted in this vein. Sometimes in myth and story, a duty will take the form of a *quest* or a *mission*. The paradox of asserting your freedom by choosing to bind yourself to a duty is also something that [the existentialist philosophers](#) liked to chew on.

Conscience is sometimes described as a sort of sixth-sense that allows you to *perceive* your duty; other times as a kind of agent that *tells* you your duty. Similarly, some people claim to *hear* the "call of duty" —they discover their "calling." In these senses, your duty was there all along, and you just needed to become aware of it.



Other duties are implicit: if you're raising a child, it's your duty to take care of their basic needs, and nobody expects you to first take an explicit vow or search your conscience in order to do this. Which duties are implicit and bind people without their explicit consent, and what those duties require of people, are part of what defines a culture. What do parents owe their children; what does filial piety demand from children; what are the duties of serfs and lords toward one another; what does a just government owe its citizens, and what obligations can it justly place on its citizens; [what do nobles owe society](#); what are our religious duties; are there universal duties that everyone owes their fellow-man? Questions like these are part of the intracultural negotiation of what it means to be a good person.

[Loyalty](#) usually includes duty, and this can be a mixture of explicit and implicit duties.

The [applause](#) light glow around the word “duty” can be used in manipulative ways. An institution that wants to conscript you into its service will sometimes try to reframe this as your duty rather than its imposition: “jury duty” for example. This rhetorically attempts to put the burden of objecting to the conscription on you, rather than the burden of justifying it on the institution.

Doing your duty

“Never shirk the proper dispatch of your duty, no matter if you are freezing or hot, groggy or well-rested, vilified or praised, not even if dying or pressed by other demands. Even dying is one of the important assignments of life and, in this as in all else, make the most of your resources to do well the duty at hand.” —Marcus Aurelius^[6]

Once you know what your duty is, in general terms, to be dutiful you must then carry it out. This requires at least the following:

1. **[Resolve / commitment / dedication](#)**. This translates knowing your duty into actions that fulfill that duty. Resolve is a sort of hand on the crank that keeps turning until your duty is done.
2. **Diligence**. If you know your duty in general terms, or in terms of the goal it is meant to meet, you may still need to do some work to break that down into what that duty requires of you specifically and right now. To do your duty, you need to put in that work.

“[L]et him who gropes painfully in darkness or uncertain light, and prays vehemently that the dawn may ripen into day, lay this other precept well to heart, which to me was of invaluable service: ‘Do the duty which lies nearest thee,’ which thou knowest to be a Duty! Thy second duty will already have become clearer.” —Thomas Carlyle^[7]

3. **Conscientiousness**. You keep your duty at the forefront of your thoughts, and judge all of your actions at least in part by how they further or interfere with fulfilling your duty.
4. **[Industriousness](#)**. You put in the work, doing what it takes to accomplish your duty.

In some circumstances, you might fail in spite of having done your duty (my duty was to defend the bridge, which I did honorably, but our position was nonetheless overrun by stronger enemy forces); in others, failure might mean precisely that you have failed in your duty (my duty was to defend the bridge, but I unthinkingly neglected to guard against saboteurs and the bridge was demolished).

Related virtues

Aside from commitment, conscientiousness, dedication, diligence, engagement, filial piety, loyalty, purposefulness, resolve, responsibility, and service, which I’ve already mentioned, other virtues that are related to duty include patriotism & [citizenship](#), accountability, and reliability.

1. ^

Friedrich Nietzsche, *Twilight of the Idols* (1888)

2. ^

Viktor Frankl, *Man’s Search for Meaning: An Introduction to Logotherapy* (1963)

3. ^

J. Rodin & E.J. Langer [“Long-term effects of a control-relevant intervention with the institutionalized aged”](#) *Journal of Personality and Social Psychology* (1977), as described in T. Lomas, K. Hefferon, & I. Ivitan *Applied Positive Psychology* (2014) p. 105. That paper also made claims about a dramatic effect of the intervention on mortality; see [this page](#) for some skepticism about that.

4. ^

Helen Keller, journal, December 10, 1936

5. ^

Lucius Annaeus Seneca (the Younger), *De Beneficiis*

6. ^

Marcus Aurelius, *Meditations* VI.2

7. ^

Thomas Carlyle, *Sartor Resartus* (1836), chapter IX: “The Everlasting Yea”

Notes on Social Responsibility

This post examines the virtue of **social responsibility**. It is meant mostly as an exploration of what other people have learned about this virtue, though I've been selective about what I found interesting or credible, according to my own inclinations.

I wrote this not as an expert on the topic, but as someone who wants to learn more about it and to become better at it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

I am going to mostly shy away from writing about specific contentious things that some people consider to be socially responsible. I don't intend to stake out a position in the "social justice" debates, for example. I plan instead to consider this question: whatever it is that socially responsible people do, how do you become the sort of person who characteristically does that?

(I may not be able to totally avoid the temptation to get on the soapbox and harangue a bit, however.)

What is this virtue?

"That which isn't good for the hive isn't good for the bee." —Marcus Aurelius^[1]

"Man is by nature a political animal."^[2]

"The polis is a community of equals for the sake of a life which is potentially the best."^[3] —Aristotle

"In no other realm does human excellence approach so closely the paths of the gods as it does in the founding of new and in the preservation of already founded communities." —Cicero^[4]

People participate in groups of all sorts: internet discussion fora, churches, fandoms, nations, audiences, mobs, corporations, neighborhoods, book clubs, posses, gangs, unions, classes. When things go well, we get a lot of value from such participation. Some groups are better than others at contributing to the thriving of those in them; some can even be harmful. Any particular group may have aspects that work better or worse in this way. Groups also may be more or less sustainable: there are potential costs to be borne in starting a group, in the disintegration of a group, and in keeping a group going.

All of which is to say that the flourishing of each of us as individuals is influenced by the health of the groups in which we participate.^[5] The health of these groups in turn is in large part a function of the decisions made by the people who make up these groups.

The person with the virtue of social responsibility recognizes this, and works earnestly to sustain, defend, and improve the groups that person participates in (or perhaps in extreme cases, to sabotage, disband, or replace those that are harmful and irredeemable) and the environments in which those groups operate.

Related virtues and vices

Other virtues in the same general bailiwick include teamwork/cooperation, leadership (and its flip side: obedience/submission?), [loyalty/fidelity](#), and [duty/responsibility](#).

Rhetoric/persuasion and charisma/influence can make you more effective at shaping cooperative efforts. Those virtues are also important because persuasion via discussion, debate, and reason are the classic alternatives to coercion by violent force in times of political conflict. Hannah Arendt identified a third alternative to those two: authority.^[6] Commanding authority, or seizing the mantle of authority, can also be a way to wield political influence.

The virtue of social responsibility is sometimes called “civic” virtue or “citizenship” (because the latter also refers to a particular legal status, this can be confusing). The Boy Scouts, for example, have merit badges for “citizenship in society,” “citizenship in the community,” “citizenship in the nation,” and “citizenship in the world,” that address social responsibility. Sometimes civic virtue at the national level is described as [true patriotism](#).

Social responsibility can strengthen and be strengthened by a sense of belonging or community or solidarity. You are more likely to feel responsible for something when you feel that it is (at least partially) yours. And taking more responsibility for something can give you more of a stake in it.

Alexander Solzhenitsyn suggested that [courage](#) is an important component of social responsibility. “[W]e have gotten used to regarding as valor only valor in war (or the kind that’s needed for flying in outer space), the kind which jingle-jangles with medals. We have forgotten another concept of valor—*civil valor*. And that’s all our society needs, just that, just that, just that!”^[7]

One way of setting and defending high standards in a group is through the judicious deployment of [judgment and righteous anger](#).

Having initiative, being [ambitious](#) or bold, and demonstrating confidence can improve social responsibility. How many problems go unaddressed simply because it’s “not my job” for anyone to address them? The phenomenon of [bystander apathy](#) can cause even acute and obvious problems to be neglected because nobody is bold (or presumptuous?) enough to be the first person in the group to break out of the status quo trance. This is the case even for small things: That flickering light bulb in the hallway, that squeaky ceiling fan that bugs everybody a little bit... All it would take would be someone willing to say “that’s *my job today*” and do something about it.

With all of the deliberate disinformation, urban legends, propaganda, and [bullshit](#) floating around these days and causing harm and unnecessary social friction, there may be a case for adding “epistemic responsibility” to social responsibility (see: [honesty](#), [sincerity](#), and [rationality](#)). Francisco Mejia Uribe made this case not long ago in an essay for *Psyche*.^[8] He believes that the liberal norm of wanting to make room for everyone to have their own moderate and restrained opinions of whatever sort they care to—to calmly agree to disagree—isn’t good enough anymore: “the stakes of credulity are simply too high.” Instead, we have “the moral obligation to believe only what we have diligently investigated.” A responsible citizen “in her capacity as communicator of belief... has the moral responsibility not to pollute the well of collective knowledge and instead to strive to sustain its integrity.”

That is reminiscent of Sissela Bok's argument that I mentioned in my discussion of the virtue of [honesty](#): that the negative consequences of a lie include not only its harm to the person deceived, but also its corrosion of the general culture of truth-telling. [9]

Maybe that's just a specific example of a broader phenomenon: The social virtues in general—including honesty but also things like tolerance, helpfulness, [justice](#), [gratitude](#), hospitality, and [courtesy/civility](#)—all tend to contribute to healthy social environments. The practice of such virtues also exemplifies them for others and demonstrates standards of behavior, and so has ripple effects that can benefit the group and those in it.

Some schools of thought suggest that the most socially responsible thing you can do is to become more personally virtuous, rather than to concentrate on specifically community-oriented activity. The key to healthier communities, according to this school, is the level of virtue in the members of (or at least the leaders of) those communities. [10] If you are full of vices, your idea of how to improve the groups that you are part of will be affected by that (how can I change this group so it better-coddles my vices or shields me from their negative effects?). If you are full of virtues or at least desire them, you will want to change the group to be more encouraging of or amenable to the virtuous sort of person or the person who strives for virtues. In this school of thought, social responsibility and political science are secondary—for minor tweaks around the edges—while what you really need for flourishing communities are flourishing people.

Sometimes social responsibility *is* just a matter of behaving personally in a way that has good social side effects. But other times it requires more specialized skills of coordination and persuasion: To make things better, it is not enough that you alone do the right thing. If everyone in the village shits in the convenient shitting field upstream, then you digging a latrine for your family in some more sanitary location may be a virtuous and even socially responsible thing to do, but it won't do much about the cholera epidemic. In such a case, the solution is necessarily done in cooperation (or potentially in conflict) with others.

Aristotle saw ethics and politics as mutually-reinforcing: you developed the virtues in large part so that you could contribute as a citizen to making the polis an excellent and stable one, and you did that in order that the polis would be fertile soil in which a thriving, virtuous citizenry could grow and in order that there could be an appropriate political context in which you could successfully practice virtues like justice. [11]

Social responsibility is sometimes described as a particularly important virtue for the financially fortunate and privileged, in a [noblesse oblige](#) sort of way. For example, Aristotle's ethics included the virtue of munificence—a sort of lavish generosity exhibited by the rich on behalf of their communities. [12] A wealthy Athenian might finance and tastefully produce a festival or theatrical performance, erect a public building or religious shrine, or host visiting foreign dignitaries in style.

On the vice side, you might lack social responsibility because you feel powerless over group dynamics: believing that you can only be responsible for yourself, and whatever behavior of the group emerges does so from an unpredictable chaotic process, or due to the machinations of more sophisticated and powerful operators than yourself. A sort of helplessness or failure of imagination might be the vice at the base of this, or it could be just the considerable cognitive challenges of understanding complex social machinery. The ways to effectively help shape a community are not always obvious. It takes some skill just to identify which levers to pull or dials to turn. (In some

communities this may even be more-or-less deliberately obscured, for example in a city where the ordinary rubes petition the mayor while the sharp operators who know how to get things done take the city manager out for cocktails.)

Another vice that interferes with social responsibility is [pessimism](#) or indifference: “ah, so my group is going to hell in a hand-basket, isn’t that just the way of the world? wuchagunnadu?” That highlights another possible benefit of social responsibility: an invigorating feeling of empowerment and agency. If you make an effort to be socially responsible, presumably you do so because you feel that your decisions matter, that you have some amount of influence.

Some people are recluses who do not seem to get much value from group participation, and so have less motivation to contribute to the health of groups. This perhaps implicates virtues like [amiability](#), community, and connection.

Some people are social vandals—they like to see what happens when things break or people get hurt, so they deliberately sabotage groups in various ways in order to see sparks fly. This suggests a lack of virtues like [benevolence](#) and [respect for others](#).

Some people seem to identify their self interest with a short-sighted and narrowly-defined selfishness. They do not see how their influences on the groups they belong to contribute to their own thriving in indirect ways. They may go through life burning bridges behind them and constantly scramble to scrape together a pittance of social capital. It takes long-term thinking and imagination to follow the diffuse cause-and-effect by which you contribute to health of the group, and the health of the group contributes to your well-being, without any obvious and immediate quid-pro-quo.

There is some tension between social responsibility, which is often accompanied by [loyalty](#) and some partiality, and virtues like independence, [impartiality](#), and cosmopolitanism.

There is also some tension between social responsibility and [humility](#). To exert social responsibility may imply that you have confidence (or, less flatteringly, hubris) about how to best influence a group and those in it. It can be difficult to know how bold or how humble you should be in such circumstances. You might feel inhibited from trying to help influence a social group because you are unsure if you are a fully-fledged member of the group, or have sufficient authority within it, or have enough of a grasp of the group’s history and dynamics to understand the effects of your interventions.

Some aspects of social responsibility

Social responsibility vs. social engineering

A thin line separates social responsibility from the darker arts of social engineering or social manipulation. Many of the same skills that can make you better at social responsibility—imagination, insight, patience, charisma, influence, leadership, ambition, and so forth—can also help you parasitize groups such that others bear the costs while you reap the benefits. Even a bank robber is deliberately changing the dynamics of his social environment to better contribute to his idea of human thriving.

Social responsibility paradigmatically concerns win-win scenarios; social engineering, zero-sum or parasitic ones. But this is not a hard-and-fast rule. The *Nudge* school of

social engineering ostensibly aims at socially responsible ends, for instance.^[13] Some socially responsible efforts may be beneficial to the group, and in general to those in it, but nonetheless result in winners and losers. And even what looks like a broadly win-win scenario to you may not appear that way to others with different priorities. Expedient group compromises frequently compromise standards of fairness, or cannot possibly satisfy diverse standards of fairness. And sometimes things operate differently at different levels: for example, you work to strengthen your group in an internally win-win way in order that your group may successfully compete with other groups in zero-sum competitions.

Some people behave as though they have a model of the world in which most all social arrangements are zero-sum, and that therefore one's social efforts ought to be directed toward figuring out how to get on the winning side and get other people on the losing side as much as possible. Such people's attempts at social responsibility are bound to be stunted by such a worldview.

One way of distinguishing social responsibility from social engineering might be to recall Aristotle's distinction between a virtue and a science.^[14] A science concerns a subject matter in which your knowledge and skill can help you aim for opposite extremes: for instance, a doctor knows the science of health, and this knowledge is just as useful to her in curing someone or in poisoning them. A virtue, on the other hand, goes in only one direction—having the virtue of courage doesn't also make you a more proficient coward, for instance. So social engineering might be regarded as the science of social manipulation; social responsibility as the virtue of striving to promote the health of one's social environment.

But when I look at it that way, I wonder if there really is such a virtue. Maybe what I've been calling the "virtue" of social responsibility is merely the science of social engineering combined with virtues like benevolence and respect for others, or with a robust understanding of self-interest.

Collective action challenges

Obstacles to social responsibility include collective action problems like those of the [prisoners' dilemma](#) / free rider / tragedy of the commons sort. What if you can best thrive personally by making decisions that degrade the social environment? Or, to look at it the other way, what if the best contribution you can make to the health of the social environment is one that would leave you personally worse-off?

In unique or unusual cases, you may just have to prioritize and make a judgement call. However, if there is a class of problems like this, or if they repeat themselves, it may turn out that one of the best things you can do for your community is to "go meta" and try to design some sort of workaround for the collective action problem itself rather than merely addressing the individual case in which the problem arises.

It is unfortunately common to see collective action problems depicted as the social sciences equivalent of a dead end: Once you foresee "a tragedy of the commons situation" you have been defeated and you must immediately turn around and head the other way. But collective action challenges are ubiquitous and solutions to them are varied and ingenious.^[15] Economist [Elinor Ostrom](#) was an avid researcher of how different communities have invented ways to resolve collective action problems. She concluded:

At any time that individuals may gain from the costly action of others, without themselves contributing time and effort, they face collective action dilemmas for which there are coping methods. When de Tocqueville discussed the “art and science of association,” he was referring to the crafts learned by those who had solved ways of engaging in collective action to achieve a joint benefit. Some aspects of the science of association are both counterintuitive and counterintentional, and thus must be taught to each generation as part of the culture of a democratic citizenry.^[16]

Ostrom believed that typical civic education shortchanges citizens by overprioritizing “elections, political parties, and what politicians and public officials do” and that we should also try to learn more about these varied, painstakingly developed, sometimes unintuitive grassroots methods of conflict resolution and of building more capable social organizations.

Care for the commons is one way to exhibit social responsibility. You may correct a misspelling on Wikipedia, pick up litter in the parking lot, or buy carbon offset credits, for example, as attempts to preserve or improve the commons. Eliezer Yudkowsky’s essay [“Well-Kept Gardens Die By Pacifism”](#) is a plea for people to adopt a stronger ethos of social responsibility in the context of on-line commons in particular, which can devolve into awfulness if they are not vigilantly protected against trolls and vandals and other such fools. If we want delightful commons, someone has to put in the work to maintain their delightful nature: someone who appreciates the commons and who also has the social responsibility to recognize that this maintenance is their job too.

Some on-line commons have formally codified some of their best practices for socially responsible behavior. Wikipedia, for example, has a [code of conduct](#) and a [huge variety](#) of other [principles](#) (e.g. [“assume good faith”](#), [“be bold”](#)). We may be able to improve our contributions to the general on-line commons by thoughtfully reviewing advisories like these.

Levels of social responsibility

Social responsibility applies to different sorts of groups, and to groups of different sizes. Different skills and techniques of social responsibility work better in different contexts. The skills you need to negotiate the allocation of household chores harmoniously among roommates are different from those you would use when trying to desegregate a school district or defend against a Russian invasion. Some communities are brief, others are durable, and this too means different skills come into play. A community of commuters sharing a bus between stop A and stop B is an ephemeral one, but one for which social responsibility can still be relevant.

You can expect individual efforts of social responsibility to be diluted according to the size of the group. A child’s tantrum can interrupt a church service, but hundreds of thousands of protesters can fill the streets and still fail to change a national policy.

Sometimes what you do has effects at multiple levels. For example, how you behave in the *LessWrong* comments section affects a particular comments thread, the *LessWrong* community, and how that community serves as an example for the larger internet community. Sometimes your efforts might have conflicting effects on different levels—maybe even deliberately so, as with advocates for centralization or

decentralization of political power who may want to empower one level and disempower another.

We are members of intertemporal communities as well. We maintain a “conversation” of sorts with past authors and other cultural creators. Our language & culture & institutions & such came into our hands as shaped by past generations. We in turn pass our own influences on to generations that will succeed us. To the extent we recognize ourselves as participants in this generational torch-passing, we also can engage in social responsibility to the community we have in common with our ancestors and with our posterity. Helping the children we raise to understand their culture and to participate confidently in it is one way we can do this.

Effective altruists and people who try to mitigate existential risks can be thought of as trying to optimize social responsibility at the largest scale: the community of conscious beings from here to eternity.

The trouble with government

“Suppose a problem in psychology was set: What can be done to persuade the men of our time—Christians, humanitarians or, simply, kindhearted people—into committing the most abominable crimes with no feeling of guilt? There could be only one way: to do precisely what is being done now, namely, to make them governors, inspectors, officers, policemen, and so forth; which means, first, that they must be convinced of the existence of a kind of organization called ‘government service,’ allowing men to be treated like inanimate objects and banning thereby all human brotherly relations with them; and secondly, that the people entering this ‘government service’ must be so unified that the responsibility for their dealings with men would never fall on any one of them individually.” —Tolstoy[\[17\]](#)

“The fate of the country does not depend on how you vote at the polls—the worst man is as strong as the best at that game. It does not depend on what kind of paper you drop into the ballot-box once a year, but on what kind of man you drop from your chamber into the street every morning.” —Thoreau[\[18\]](#)

Standing, institutionalized governments hold out the promise of being tools that people can leverage for socially responsible action. Indeed, they tend to present themselves as *the* legitimate tools for such a purpose. In government-oriented propaganda about civic responsibility, it can be neutered to things like being law-abiding, paying your taxes, voting, petitioning, and jury duty. This can atrophy the sort of civic responsibility that is most needed: instead of deliberating or acting together, people merely vote or petition (or complain about the government). They lose the skills they need to work together on large-scale problems; they assume such problems are “out of our hands.” If some shared social responsibility becomes a bureaucrat’s job title, that can have the effect of making it everyone else’s “not my job.”

On the other hand, if there are existing institutional channels for political action, you don’t have to reinvent the wheel when you need to make collective decisions. Because of their familiarity and regularity, such channels can make it easier for people to learn some modest forms of political engagement. For example, de Tocqueville observed that “Town meetings are to liberty what primary schools are to science; they bring it within the people’s reach, they teach men how to use and how to enjoy it.”[\[19\]](#)

Sometimes standing government is justified under the assumption that it *causes* social responsibility, or is a prerequisite for it. Without the overwhelming coercive capability of government forcing us to be civilized, according to this theory, we would become bestial and savage in our civic nakedness: “nature, red in tooth and claw”^[20] would reemerge, in some *Lord of the Flies* or *Mad Max* way, and, until we can bring ourselves back under the protection of our authorities and institutions...

In such condition there is no place for industry, because the fruit thereof is uncertain... no account of time, no arts, no letters, no society, and which is worst of all, continual fear and danger of violent death, and the life of man, solitary, poor, nasty, brutish, and short.^[21]

Some inadvertent experiments in which institutional government is temporarily disrupted because of natural disaster have undermined this theory. The sociological subdiscipline of “disaster studies” has revealed that bloody tooth and claw is hardly the only backstop following the collapse of institutional political authority. Rather:

Just as many machines reset themselves to their original settings after a power outage, so human beings reset themselves to something altruistic, communitarian, resourceful, and imaginative after a disaster... [W]e revert to something we already know how to do.^[22]

Institutional government and other standing institutions can also degrade individual social responsibility (as Tolstoy pointed out in the quote above). People sometimes don’t consider the ramifications of their behavior if that behavior is legal (or legally mandated), is part of their job description, or is directed by those above them in an institutional hierarchy. “Good Germans” for example, fooled themselves into pursuing respectability in the eyes of very *irresponsible* social institutions, as an apparently palatable counterfeit of social responsibility.

I work with an all-volunteer, private, donation-supported non-profit that provides services to homeless people in my community. In this role I have also been struck by how comparatively *inefficient* government-provided solutions are. Often paltry results emerge from expensive and well-staffed government projects ostensibly designed to help homeless people. This experience has made me more skeptical of proposed solutions to societal problems that boil down to “let’s get the government to do it.” I’ve become knee-jerk about slapping an implied “maybe if all else fails” onto the front of that.

In short: government strikes me less as a useful mechanism for enacting social responsibility, and more as something unfortunate that seeps in and fills the cracks where social responsibility has failed. And I should probably get off my soapbox now.

Consumer responsibility

When you purchase goods or services, you help make profitable—and thereby incentivize—the way those goods or services were provided. A provider may cause harmful things to communities, or on the other hand may contribute to communities above and beyond the value of the goods and services provided. Your purchase may also incentivize these additional effects.

For example, if you buy a bargain bicycle at the swap meet from a vendor who “doesn’t ask questions” about where the used bikes he purchases and refurbishes

come from, you help make the black market in stolen bicycles profitable and thereby incentivize bike theft.

Abolitionists who realized how much of the cotton and sugar trade was dependent on slave labor felt the need to boycott such goods or to instead consume painstakingly-differentiated "[free produce](#)" alternatives.

If our purchasing decisions send the message "I don't care how you do it, but I want it cheap," should we be surprised when the companies who win the bidding war for our custom and who then dominate the market turn out to be the ones that have cut corners? Doesn't this suggest that if we want better behavior from the suppliers of our lifestyles that we'll have to pay a little more attention?

I recently participated in a discussion group that mulled over the topic of consumer responsibility, and we mostly threw up our hands in frustration at the topic. The argument in favor of greater consumer responsibility is intuitively attractive. But when we started trying to consider what criteria we would use, and how we would go about investigating whether, for example, a pair of sneakers or a smart phone was more-or-less responsible of a purchase than its competitor... we began to despair. If we tried to do this for every purchase we made, it seemed we'd have a full-time job of it. There are groups that try to do this work for you—for example, those that provide certifications like "[fair trade](#)", halal/kosher, American Humane, or union labels—and sometimes particular brands will stake their reputations on a certain ethos of responsibility—but you still have to find those whose criteria you agree with, and you have to trust them to do their job well.

On the other hand, most of us reported that we did confidently or semi-confidently make some of our shopping decisions from motives of consumer responsibility. And some of us used general heuristics that we felt were reasonably reliable guides even if they were not powerful enough to capture all that could conceivably be at issue.

How to become more socially responsible

The most discussed and studied methods to improve social responsibility seem to be those designed for schoolchildren. All of the interventions described in the positive psychology handbook *Character Strengths and Virtues* were targeted either at juveniles or at criminal offenders. [\[23\]](#)

I see this pattern a lot as I learn about virtue-related interventions and studies, and I find it discouraging. Adults often seem unduly self-congratulatory. In particular, I do not think there is good reason to be sanguine about the level of social responsibility in the typical adult.

In addition, the typical school is a very authoritarian place, and so it mostly teaches an authoritarian-friendly variety of social action (e.g. ask the teacher, complain to the principal). There are some nods toward teaching children techniques of self-organization or giving them authority to make collective decisions (e.g. boy scouts, school clubs, model UN). And some varieties of school put higher priority on student initiative. Children also inevitably learn on their own how to make group decisions in their occasional islands of freedom from adult control (e.g. which games to play at recess, who gets which roles in those games, how to decide on the rules). But it seems

to me that interventions designed to help people become more socially responsible ought more often to be remedial education targeted at adults who have been subjected to such intense and lengthy authoritarian control, rather than as a half-hearted corrective directed at the children who are still otherwise subject to it.

(Interventions designed for schoolchildren, because they are often mediated through taxpayer-funded government-run schools, inevitably also get tangled up in culture war stuff, which can make them very stupid indeed.)[\[24\]](#)

One of the more well-studied and frequently-deployed varieties of social responsibility education is [“service learning.”](#) Students engage in hands-on community service projects, and simultaneously are asked to reflect more consciously on the agency of the individual in the health of the community. This is almost exclusively studied as an intervention on schoolchildren, but I expect the basic technique could be used profitably by adults who want to improve their social responsibility. In any case, I’m often tempted, when I hear someone complain about this or that aspect of society, government, the environment, public policy, corporate misbehavior, and so forth, to recommend that the complainer take a crash course in getting their hands dirty and helping to do something about it.

For what it’s worth, I found my own social responsibility went up several notches after I read the autobiography of American anarchist and all-around ornery character Ammon Hennacy. His idiosyncratic variety of anarchism was all about setting and following high standards of social responsibility. Here’s a taste:

A lady wrote a letter to the local paper about a dead cat on the street, and bemoaned the fact that no one came to remove it. A week later she wrote again and the cat was still there. In an anarchist society, each one would be responsible, and would not have to write letters to papers, or to call the cops, to have something done: They would do it themselves.

Coming home from helping my friend Joe Craigmyle pick oranges and grapefruit the other night, I mentioned this lady and the cat, and said that, the Sunday before, I had seen a dead cat on the lateral on my way to the bus, but being late, I did not stop to remove it. On my way back in the afternoon, after hundreds of cars had passed, and numerous Mexicans going to the bus, I noticed that the cat was still there, and stopped to throw it off of the road. As we were talking, we noticed a two-by-four with four spikes sticking up on the highway. We swerved around it, and were a quarter of a mile past, when Joe said, as an afterthought to my remark, that this would cause somebody some trouble: “I’ll back up and you can throw it in the ditch.” In my mind, then, Joe, who has not been much of a man of action, rose from a one-cylinder to a two-cylinder anarchist.[\[26\]](#)

Something that I have found helpful in improving my social responsibility is this: I’ve attached a flag to the sentiment “somebody oughta.” Whenever I find myself thinking “somebody oughta fix that thing” or “somebody oughta do something about that” or something of that nature, this flag goes up and that reminds me to investigate further. Who exactly is this somebody who oughta? How do they become aware of what they oughta do? How do they prioritize among the things they oughta do? What are the constraints that they have to work with? How do they marshal the resources to do what they oughta do? What are the obstacles in their way, and why are they there? Or is there no such somebody? Is it perhaps *me* who oughta do something? If so, how would I go about doing it? Who are the stakeholders and how would I get buy-in? How could I obtain the necessary expertise and material? Whom might I enlist to help, and

what arguments might they find persuasive? What additional future responsibilities does my solution entail, and who can take those on?

You can amplify the effect of your socially responsible impulses by coordinating your efforts with others. For example, Benjamin Franklin's [Junto](#) was both a small cooperative continuing education and self-improvement club, and a mechanism for initiating and coordinating schemes of civic improvement. America's first lending library, a volunteer fire department and volunteer militia, the University of Pennsylvania, and the Pennsylvania Hospital all got their start as discussions in the Junto.

A thought experiment that some people find motivating when developing their social responsibility is this: Imagine your life as a whole and ask "is the world a better place for my having lived my life the way I did?" This is implicitly an answer to the question "what's it all about?", which is a notoriously difficult one to answer confidently. (In any case, I falter and blush when I try to state my best guesses out loud.) But that answer strikes me as worth trying on for size anyway.

1. ^

Marcus Aurelius, *Meditations* VI.liv

2. ^

Aristotle, *Politics* I.ii. And we're not the only one; see, e.g., Frans de Waal's *Chimpanzee Politics* (1982)

3. ^

Aristotle, *Politics* VII.viii

4. ^

Cicero, *De Legibus*

5. ^

Or, in social-scientist-speak, how "salutogenic" the groups are, e.g. S. Graeser "[Salutogenic factors for mental health promotion in work settings and organizations](#)" *International Review of Psychiatry* (2011) pp. 508-515

6. ^

Hannah Arendt, "What is Authority?" *Between Past and Future* (1968) pp. 91-141: "Since authority always demands obedience, it is commonly mistaken for some form of power or violence. Yet authority precludes the use of external means of coercion; where force is used, authority itself has failed. Authority, on the other hand, is incompatible with persuasion, which presupposes equality and works through a process of argumentation. Where arguments are used, authority is left in abeyance."

7. ^

Alexander Solzhenitsyn, *The Gulag Archipelago, 1918-1956: Volume 1: An Experiment in Literary Investigation* (1973)

8. ^

Francisco Mejia Uribe, "[To be a responsible citizen today, it is not enough to be reasonable](#)" Psyche 12 January 2021

9. ^

Sissela Bok, *Lying: Moral Choice in Public and Private Life* (1978)

10. ^

e.g. that school of thought described in James Hankins *Virtue Politics: Soulcraft and Statecraft in Renaissance Italy* (2019)

11. ^

Aristotle *Politics* I.i–ii, III.iv, VII.i–iii; *Nicomachean Ethics* II.i, V.vi, X.ix

12. ^

Aristotle, *Nicomachean Ethics* IV.ii

13. ^

Cass R. Sunstein, *Why Nudge?: The Politics of Libertarian Paternalism* (2014)

14. ^

Aristotle, *Nicomachean Ethics* V.i

15. ^

One of the takeaways from David Graeber's & David Wengrow's newish (2021) book *The Dawn of Everything: A New History of Humanity* is that historic and prehistoric human political and social organization in the face of collective action problems has been spectacularly inventive and flexible, in a way that can be obscured by homogenizing social science categories like "hunter-gatherer bands" or "matrilineal families" or "male dominance hierarchies" or what-have-you.

16. ^

Elinor Ostrom "The Need for Civic Education: A Collective Action Perspective" 1998

17. ^

Leo Tolstoy, *Resurrection* (1899)

18. ^

Henry David Thoreau "[Slavery in Massachusetts](#)" (1854)

19. ^

Alexis de Tocqueville "The American System of Townships and Municipal Bodies"
Democracy in America vol. 1, ch. 4 (1835)

20. ^

Alfred, Lord Tennyson "In Memoriam A.H.H." (1850), Canto 56

21. ^

Thomas Hobbes, *Leviathan: or The Matter, Forme and Power of a Commonwealth Ecclesiasticall and Civil XIII.ix* (1651)

22. ^

Rebecca Solnit, *A Paradise Built in Hell: The Extraordinary Communities that Arise in Disaster* (2010)

23. ^

Christopher Peterson & Martin E.P. Seligman *Character Strengths and Virtues: A Handbook and Classification* (2004), chapter 16. Peterson & Seligman cluster "citizenship, social responsibility, loyalty, and teamwork" and say they have in common "...a feeling of identification with and sense of obligation to a common good..." (p. 370)

24. ^

See for example [Senator Rick Scott's recent "An 11 Point Plan to Rescue America"](#) which promises both "Kids in public schools will say the Pledge of Allegiance, stand for the National Anthem, and honor the American Flag... Public schools will teach our children to love America because, while not perfect, it is exceptional, it is good, and it is a beacon of freedom in an often-dark world" and, in the same set of bullet points: "We will not allow political or social indoctrination in our schools."

25. ^

Two illustrative quotes from Hennacy:

- "An anarchist is someone who doesn't need a cop to make him behave."
- "Oh judge! Your damn laws! The good people don't need them, and the bad people don't obey them."

26. ^

Ammon Hennacy *The Book of Ammon* (1970), free ebook available [here](#).

Notes on Prudence

This post examines the virtue of **prudence**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

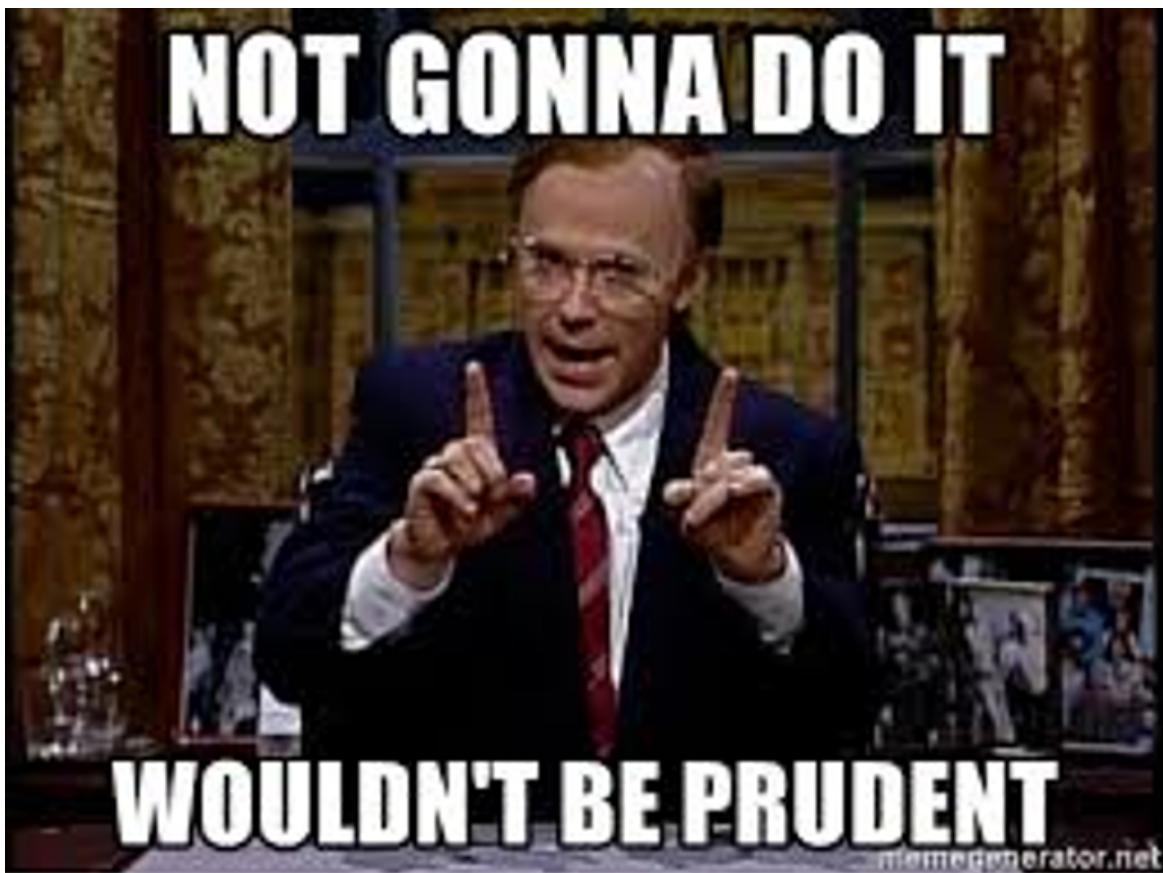
What is prudence?

Prudence is one of [the four cardinal virtues](#). From there it became part of [the seven traditional Christian virtues](#). It turns up again and again in virtue traditions. I can't very well ignore it.

And yet... the word "prudence" has gone through such a dramatic shift in meaning that it's difficult to know how to tackle this one.

"Prudence" was a common English translation of the Greek word *phrónēsis*, which has implications that range from having how-to skills to things like choosing your goals wisely and exercising good judgment when picking paths to those goals. In short, it is [wisdom](#) applied to practical, real-world [decision-making](#), where the rubber meets the road.

When prudence was incorporated into [the traditional Christian virtues](#), it was via the Latin word *prudentia*, which can mean things like [rationality](#), insight, discernment, foresight, wisdom, or skill. Again, though, the focus is on the quality of your process of making practical decisions, so this isn't too far off.



Dana Carvey as President G.H.W. Bush on *Saturday Night Live*

But nowadays when you call someone “prudent” you usually mean that they are [cautious](#): they plan ahead, look before they leap, avoid taking unnecessary risks, save for a rainy day, and that sort of thing. The word now has an old-fashioned sound to it, and is rare enough as a compliment that it’s sometimes even deployed as an insult, to imply that the “prudent” person is over-cautious, timid, afraid to take chances, or reluctant to innovate. (The resemblance of the word “prudence” to the etymologically distinct word “prudish” has also contributed to giving the word a stuffy connotation.)

Because of this meaning shift, when you see someone singing the praises of “prudence” it’s important to investigate further to find out which sort of prudence they’re praising. Sometimes authors will even drift from one definition to the other without seeming to realize that they’re doing so. ^[1]

The authors of *Character Strengths and Virtues: A Handbook and Classification*^[2] found such close kinship between the virtue of prudence and the [five-factor personality model](#) factor of [conscientiousness](#) that they tend to use the latter as a proxy for the former (this also because much more psychological research has been done about the well-defined conscientiousness factor than about prudence, which lacks such a consensus definition for the psychological community to coordinate their research efforts around).

Prudence as practical wisdom / decision theory

The science of what is a rational decision to make, given certain goals and constraints and uncertainties, is called [Decision Theory](#). It is complex and interesting and I am thankful that

there is a marvelous [Decision Theory FAQ](#) on LW so I don't have to try to summarize it myself.

Prudence (in the sense of “practical wisdom”) might be considered decision theory put into practice. Being practically skilled at making rational decisions is something that goes beyond theoretical understanding of good decision-making processes.

Aristotle explained the difference this way:^[3] While it's possible for a young person to be a savant with a genius understanding of something like mathematics, prudence seems to be something that must be acquired through long experience. This is because expertise in mathematics largely requires an intellectual understanding of abstract universals, while prudence requires actual encounters with real-life particulars. When you teach a young savant a mathematical truth, she grasps it as a truth immediately; but when you teach a truth of prudence, the same student may have reason to be skeptical and to need to see that truth exemplified in real-life examples first before she can internalize it into her worldview.

You exercise prudence when you:

1. Recognize that you are faced with a decision and are not indifferent to the outcome.
2. Use a skillful process of evaluating your alternatives to come up with the best choice.
3. Follow through on that decision by actually acting as you have decided to act. (This may also involve the virtue of [self-control](#).)

Psychologist Barry Schwartz has made prudence (in the sense of practical wisdom) a focus of his work. Here are links to videos of some of his talks on the subject:

- [“Our loss of wisdom”](#) (TED talk, 2009)
- [“Using our practical wisdom”](#) (TED salon, 2011)
- [“Practical Wisdom”](#) (Talks at Google, 2011)
- [“Practical Wisdom”](#) (Knowledge at Wharton, 2012)
- [“Doing the Right Thing for the Right Reason”](#) (University of British Columbia, 2012)

In part what Schwartz is doing is pushing back against theories that what we need to do to improve society is to create better rules and institutions on the one hand, or cleverly manipulate incentives on the other. He believes, and says that his research supports, that those things are insufficient. To make things better, you need to improve not the incentives or structures that people act within, but the characters of the people themselves.

If I squint and turn my head at an angle, this looks to me like the practical version of the theoretical ethics debate between deontologists, consequentialists, and virtue ethicists. Deontologists might advocate better rules and institutions; consequentialists might argue for the importance of incentives; and virtue ethicists emphasize the need for character.

Practical techniques of practical wisdom

Decision theory can sometimes be difficult to put into day-to-day practice. The simplifications that make it easier to analyze as-theory can make it impractical to apply in real life. What stands in the way of good decision-making is often not the lack of a good theory, but human biases and blind spots that cause us to neglect or ignore relevant data or possible scenarios.

A variety of techniques have been developed that are meant to correct for this.^[4] These include:

- Recharacterize yes/no binary decisions as decisions that may have multiple options.
^[5] (Prematurely framing the question as “should I do X or shouldn't I?” may blind you to alternatives Y and Z.)

- If some decision begins to look inevitable, imagine a world in which that decision were somehow impossible and imagine what decision you could then come to in that world. [6]
- Conduct a “[premortem](#)”: [7] Imagine a future state in which your preferred decision has turned out to be the wrong one after all. Why did it fail? What went wrong that you failed to anticipate? [Red Teams](#) and [devil’s advocates](#) are other methods of trying to uncover unexpected weak points of what seem to be strong decisions.
- Consider sketching out a [decision tree](#) or a [decisional balance sheet](#) to help make sure more of the factors of your decision (and the interactions between them) are salient to you.
- The timeless folk wisdom of “sleep on it” may be helpful, particularly if you are tempted to make a decision while under the influence of powerful emotions. Make a tentative decision, give yourself some distance from the context in which you made the decision, and then evaluate your decision again. [8]
- Other people may have a different set of biases and blind spots than you have, so if you ask a variety of other people for their opinions about your dilemma, you may be able to broaden your possible alternative courses of action beyond what you would have come up with yourself. This can be especially useful if the people you consult have encountered similar situations to the one you are confronting and so can share how their decisions played out.
- Ask “[what would Brian Boitano do?](#)” Consider someone whose character or decision-making process you admire. Imagine how they would confront your decision in your place. This may help you break out of the status quo bias of how you “have always done” things of this sort.
- Imagine that instead of choosing a course of action for yourself, that you are advising a friend who is in your situation which action you would recommend they take. [9]

Measuring phrónēsis

In 2022, a team of researchers began to devise an assessment method designed to measure phrónēsis. [10] It subdivides phrónēsis into four “functions”:

1. *constitutive* (moral sensitivity)—“the ability to perceive the ethically salient elements of a situation and recognize the best response”
2. *integrative*—“allows one to adjudicate the cognitive and affective aspects of situations and choose the best action when conflicting demands arise”
3. *blueprint*—“overall understanding of how actions conduce to a flourishing life”
4. *emotional regulative*—“the ability to infuse one’s emotional experience with reason to appropriately shape those emotional responses”

The paper is paywalled, and my own practical wisdom ruled out sending \$35.95 to Elsevier for a peek, but judging from what is on this side of the paywall, the researchers were able to use some existing tests to approximate measures of these functions, the results of this testing were promising, and they hope this leads to a more precisely-targeted test for phrónēsis (that could then presumably help us to design interventions to improve it).



although deer prudence is celebrated in song, deer are notoriously incautious pedestrians

1. ^

see, for example, Kathryn Britton, [In Praise of Prudence](#) Positive Psychology News 12 March 2013

2. ^

Christopher Peterson & Martin E.P. Seligman, *Character Strengths and Virtues: A Handbook and Classification* (2004)

3. ^

Aristotle, *Nicomachean Ethics*, book VI, section 8

4. ^

Chip Heath & Dan Heath, *Decisive: How to Make Better Choices in Life and Work* (2013) is a good airport-bookstore-type overview of some of these techniques and how they have been put into practice.

Steven Johnson, *Farsighted: How We Make the Decisions That Matter the Most* (2018) is another.

5. ^

Steven Johnson *Farsighted* (2018) p. 67 (he refers here to the research of Paul C. Nutt on organizational decisions)

6. ^

Steven Johnson *Farsighted* (2018) p. 68

7. ^

Gary Klein *Streetlights and Shadows: Searching for the Keys to Adaptive Decision Making* (2009) pp. 235-36

Chip Heath & Dan Heath, *Decisive* (2013) pp. 202-03

Steven Johnson *Farsighted* (2018) p. 118

8. ^

e.g. Chip Heath & Dan Heath, *Decisive* (2013) p. 23 (and chapter 8), identify the main obstacle at the time-of-decision as “short-term emotion will often tempt you to make the wrong [choice]” and recommend that you therefore “Attain Distance Before Deciding.”

9. ^

Laura Kay & Richard Gonzalez “Weighting in Choice versus Advice: I’ll Do This, You Do That” *Journal of Behavioral Decision Making* (1999)

10. ^

Catherine Darnell, Blaine J. Fowers, & Kristján Kristjánsson [“A multifunction approach to assessing Aristotelian phronesis \(practical wisdom\)”](#) *Personality and Individual Differences* 196 (2022)

11. ^

[VIA Inventory of Strengths \(VIA-IS\).](#)

12. ^

[“Character Strengths: Prudence”](#) VIA Institute on Character website

13. ^

Flavio Gerbino [“Logical Fallacies when Assessing Risks”](#) *scip*

Notes on Know-how

This post examines the virtue of **know-how**. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is know-how?

“A human being should be able to change a diaper, plan an invasion, butcher a hog, conn a ship, design a building, write a sonnet, balance accounts, build a wall, set a bone, comfort the dying, take orders, give orders, cooperate, act alone, solve equations, analyze a new problem, pitch manure, program a computer, cook a tasty meal, fight efficiently, die gallantly. Specialization is for insects.” —Robert A. Heinlein^[1]

You have know-how to the extent that you have skills obtained from practice doing things in the real world.

Know-how is distinct from mere theoretical understanding. For example, the theoretical discipline of hydrodynamics may be fascinating and useful, but if you want to know how to swim you have to get wet.

Know-how as a virtue seems to mean having competence in a broad, well-chosen range of skills.

Related virtues

Know-how is related to “practical knowledge,” though practical knowledge can be knowledge *about* things as well as *how to do* things. There’s some overlap. If you’re building a cabinet, practical knowledge will help you know which type of wood is best for the job, while know-how will help you construct it skillfully.

Know-how is kind of like “craftsmanship,” but craftsmanship implies a more advanced skill in a particular specialty — connoisseurship, expertise, commitment to quality, attention to detail, efficiency — while know-how is more of a baseline proficiency in many areas.

Know-how has some connection with [prudence](#) (in the sense of “practical wisdom”), though prudence has more to do with having a well-honed process for making good [decisions](#) generally, while know-how is about experience-tested decisions in specific contexts.

How to know what you need to know

An American in 1900 who knew how to drive a car with a manual transmission was an early adopter. An American who knows how to drive such a car in 2050 may be something of an eccentric traditionalist. Somewhere in-between, that skill was considered an essential part of know-how. What skills are important change over time.

They also differ from culture to culture, and from class to class. Read one of the many classic novels about the lives of the European upper classes in the nineteenth century, and see how it's taken for granted that everybody knows the steps to various dances, the rudiments of horsemanship, at least some Latin and Greek, for example.

Modern societies will explicitly express some judgements about what skills are essential through what they teach in standard/compulsory childhood schooling. Some other elementary skills everyone is expected to just pick up as a matter of course but are not explicitly taught. People then also specialize professionally, and through their hobbies and lifestyle choices, and they learn additional skills appropriate to these areas.

What remains is a vast set of skills that some people pick up and others don't. The more of this set that is important by some measure and that you can acquire some baseline competence in, the more you have the virtue of know-how.

How do you know which skills are the important ones? "Skills everybody should have" is a popular listicle genre, but there doesn't seem to be much consensus from list to list. When I asked my Facebook Friends "What are some examples of know-how skills that you think a typical person ought to have in order to be a well-rounded person, but that people often neglect to pick up?" they gave a bunch of interesting answers, but nothing that sounded like a consensus set.

Is there a more reliable and methodical way to learn what we ought to learn? It seems to me that there ought to be some sort of rough function we could apply to a skill, using criteria like these:

- how important the skill is when it is deployed
- how likely it is that (or how frequently) the skill will be called for
- positive or negative side effects to having the skill (e.g. having the skill makes it easier to learn related skills, or, maybe you get unfortunately typecast as person-with-the-skill)
- the availability and suitability of alternatives for personally having the skill (e.g. how easy is it to find a professional specialist, or to just buy something instead of make it yourself)
- how likely it is that someone else will be available who has the skill when it is needed
- how likely it is that changes in technology, fashion, or economics will make this skill obsolete
- how personally rewarding it would be to practice the skill or to have the competence under your belt, along with the prestige factor; or, on the other hand, how silly you would feel if you found yourself helpless at a time the skill was called for
- how difficult it is to acquire enough proficiency in the skill or to keep it once you've acquired it
- the difference in quality between performing the skill skillfully vs. trying to muddle through unskillfully

So for a skill like CPR, for example, you would take into account that it's a very important, potentially life-saving skill; but you may go your whole life without ever having an opportunity to put it to use; there are professionals who know how to do it best as part of their job skills, but time is of the essence and they may take time to be summoned; rudimentary CPR skills are fairly widespread; being able to save someone's life is pretty hawt; you can probably pick up the needed skills in an hour or

two, and there are classes for that sort of thing; and doing a poor job of CPR is more likely to be useless or to break ribs than to be helpful, so it's very important to actually know what you're doing.

Could you list out a bunch of candidate skills you are lacking, rate them by criteria like these, rank them according to some rough formula, and then be able to see at a glance which new skills would be most worth picking up? Seems doable, though it's not immediately obvious how you would assign values to some of these criteria or what formula you would use to combine them. I don't think I've ever heard of anyone doing this.

Personally, I've been much more haphazard about how I've picked up how-to skills. I've picked up a lot of food preservation knowledge, for example — how to pressure can, dry herbs, brew beer, culture yogurt, cure meat, ferment pickles, and so forth — but although this is interesting and useful, I don't think it would have ranked very highly based on its day-to-day importance and on the availability of commercial alternatives. Would I have been wiser to have spent that time and mind-space learning something else?

How do you decide which skills are important to learn?

1. ^

Robert A. Heinlein, *Time Enough for Love* (1973), spoken by the character Lazarus Long

Notes on Honor

This post examines the virtue of **honor**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is honor?

"The one who is conscious of his soul's nobility will not endure a dishonorable life." —Sophocles

There are several different ways in which I see the concept of honor deployed, including:

1. **Honor as a package of other virtues.** An honorable knight, for example, is one who practices all of the various virtues in the code of chivalry. A person might dishonor themselves, or their family, or their profession, by flagrantly violating any one of a set of virtues. In some contexts, "honor" is more of a euphemism for one particular virtue: a young woman in an old novel who "defends her honor" is really defending her chastity; when you address a judge as "your honor" you're hoping to get the message through to the virtue of justice in particular. In other contexts, honor is an explicit *role-based* code-of-conduct such as Omertà or the Hippocratic Oath.
2. **Honor as reliability in the practice of virtues.** When a boy scout says, "and that's the truth: scout's honor" he is asserting that he takes his pledge of honesty more seriously than the typical person, because of an additional honor code he feels bound by. Sometimes this facet goes by the name "rectitude".
3. **Honor as extraordinary investment in one's character.** An honorable person may be defined as someone who strongly values his or her character, such that they will go to great lengths to avoid doing anything vicious or shameful (even if nobody else will ever know). Sometimes this facet is called "pride" (or in an inverted way, "[a sense of shame](#)"), "character," or "[dignity](#)."
4. **Honor as public standing or reputation.** There is also a sense of honor which means something like "unusually sensitive to one's social status, and prone to take exceptional offense to being insulted" — from which you get things like "honor culture," "honor killings," and the like. In this case, your character and dignity are your own, but your honor is determined by those around you, and you may be periodically called upon to prove it or to defend it against insults.

Something common to most of these is that honorable people tend to hold themselves to unusually high standards. Someone with a strong sense of honor is not satisfied with being "more or less as decent as the next guy" but instead judges him or herself in a more inflexible and exacting way.

"The man of honor thinks of his character, the inferior man of his position. The man of honor desires justice, the inferior man favor." —Confucius^[1]

This sometimes leads to an association between honor (and especially its more aristocratic cousin “nobility”) and arrogance or vanity. In this way, honor may be in tension with the virtues of humility or modesty. If a sense of honor is used as the excuse for conforming to some arbitrary fashion (“why, that simply isn’t done where I come from,” “I would not be seen in such a place”) honor can seem a fancy name for mere snobbery.

But if honor is genuine and wise, it can make for a firm foundation for the other virtues. If you hold your character at a high price, it will be that much harder for temptation to buy you off. If your sense of honor is what motivates you, you will conduct yourself honorably even when nobody is watching.

Honor can be a variety of self-esteem, or can be a way of earning one’s self-esteem (“I would think less of myself were I to behave dishonorably”). You might think of it as a standard that you hold up for yourself, and try to sculpt yourself into, in order to make yourself as admirable as you can be.

“Character — the willingness to accept responsibility for one’s own life — is the source from which self respect springs.” —Joan Didion^[2]

Dishonor usually connotes a failure of character rather than one of skill or luck. You can lose or fail honorably if you fought the good fight.

Philosopher Kwame Anthony Appiah, in his book *The Honor Code*, examined the sorts of moral revolutions that take place when widespread practices (like slavery, foot-binding, or dueling) come to be seen as reprehensible and fall out of favor. He claimed that evolving definitions of honor are what lie behind such changes, and explored how such evolution takes place.^[3]

Honor culture

In Tamler Sommers’s book *Why Honor Matters* he takes a sympathetic look at honor-based cultures.^[4] I’ll try to summarize his argument:

The way Sommers sees it, honor fills a gap in modern Western ethical philosophy. The “WEIRD” minority of humanity (“western, educated, industrialized, rich, and democratic”) have adopted an abstract, impartial, unemotional ideal of ethical evaluation and of the administration of justice that is foreign to most of humanity and doesn’t really harmonize well with human experience. (Sommers gives these the sobriquet of “dignity”-based in contrast to “honor”-based systems.)

One advantage to honor-based systems is that they have a built-in motivator for people to behave virtuously. Instead of offering people an abstract ethical system and *reasons* they should live up to it, honor-based systems offer people status and prestige: “They have rituals and traditions for bringing people together, for celebrating exceptional people and behavior, and for holding people accountable.”

There are a couple of varieties of honor: *horizontal* honor, which you gain (and must uphold) simply by being a member of a particular honor-based culture; and *vertical* honor, which you earn (or lose) by your deeds. This is also in contrast to dignity-based cultures in which everyone is supposed to have an equal worth without distinction.

In a dignity model, people discover themselves by factoring out all of their social roles to find the essential person underneath. In the honor model, people use their role and how they uphold their responsibility to it as core parts of their identity.

Dignity is your “human right,” allegedly, while honor is more fragile. You may have to regularly defend your honor against threats and insults. But this may make honor more worthwhile. Sommers compares dignity to a “participation trophy” and honor to the real thing.

One symptom of the decline of honor culture is a heightened concern for personal safety and more risk aversion. We value our lives more than our honor, and so become increasingly cowardly. Sommers ridicules our insistence on wearing bicycle helmets, for example, along with the usual helicopter parents and such.

Another symptom is isolation, hyperindividualism, lack of community, and our descent into a sort of Ayn Randian, contractarian abyss in which all of our intercourse is temporary and contingent on mutual gain, with no cooperation in the service of something bigger than ourselves.

In contrast, the better social cohesion of honor societies leads to better mental health (people need belonging) and lower crime rates (potential lawbreakers are deterred by social norms, or by fear of shaming themselves or their families).

There is also greater personal accountability (in honor cultures, people take responsibility for their actions whether or not they accept blame for them). Dignity-based cultures like ours, by contrast, are increasingly shameless. We have an attenuated sense of blameworthiness and so a large-scale refusal to take responsibility.

In the modern liberal justice system (“dignity”-based), the people who are most involved in resolving a dispute (lawyers, judges, and the like) are those with little personal involvement in it. Those with the most skin in the game (defendants, witnesses, victims) are given minor supporting roles at best. Because of this, people who go through this process tend not to feel like things have really been resolved satisfactorily. The law has been followed (more-or-less), but there’s little sense that justice has been done or that the conflict has been resolved.

The system even denies victimhood to the victim of a violent crime, saying that the case is between the offender and “the People” or “the State.” The victim’s desires, whether they be for revenge or for forgiveness, don’t count. Emotion, the *feeling* of being wronged, being victimized, being treated unjustly, is deliberately excluded from the deliberations. This is although emotions like these are key to why we consider something to be a criminal offense in the first place. In their place, the system has erected a sort of post-hoc scaffolding of rational-sounding, measured, consistent rules, but this both masks the ultimately irrational foundations for the rules and prevents them from operating in a way that brings catharsis to these emotions.

Honor societies, on the other hand, make no pretense of creating an objective system that treats all crimes the same and focuses on the blameworthiness of the offender without getting distracted by the feelings of the victim. Instead, their processes are victim-centered, emotionally validating, and seek a cathartic resolution that restores balance in the society. They more authentically reflect human psychology about justice.

Sommers also looks at some of the downsides of honor-based cultures (vendettas and feuds, honor being used to enforce or resist the reforming of reprehensible practices, higher levels of aggression and violence — though he notes that “dignity”-based cultures can have more *official* violence and repression that makes them only superficially less-violent). He suggests some ways to mitigate these problems, such as the cultivation of trusted mediators to dampen the escalation of violence in honor-prompted feedback loops.

Megalopsychia

In the *Nicomachean Ethics*, Aristotle introduces us to the *megalopsyche*, or great-soul — a sort of pinnacle of pride and self-regard, and a connoisseur of honor.^[5]

His portrayal of the great-souled man is slightly comical, even somewhat mocking. He skips opportunities to describe the great-souled man’s most attractive qualities, and lingers over his haughty unconcern and disdain and his presumption and self-regard and the way he works to dominate others and put them in his debt. I think Aristotle may be rubbing our noses in the fact that to him virtue is meant for the benefit of the virtuous person, not for the rest of us. We should not expect a great-souled person to be the sort of person we’d want as a best buddy, but as someone who is far above us and, probably, as a result fairly contemptuous of our affairs.

Among the traits of a great-souled man:

- He deserves and claims great things, but above all, honor.
- He is good in the highest degree, great in every virtue. You never see him behaving in a cowardly manner or wronging another person, because, loving honor above all, he has no motive to do such things.
- He will be moderately pleased at receiving great honors from good people, but just thinking these his due, in fact less than his due, but as the best honors perhaps that are available under the circumstances, he will make allowance. Casual honors from middling people, he will despise.
- He is indifferent to what fate brings him — “neither over-joyed by good fortune nor over-pained by evil” and cares not for power and wealth, except as a means to honor. Even honor, which he loves above all, he doesn’t make a big deal over.
- It doesn’t hurt if he’s rich, powerful, and well-born, though none of these things are sufficient.
- He doesn’t court danger, particularly since there’s not much he finds worth courting danger for. But when he encounters danger, he faces it “unsparing of his life, knowing that there are conditions on which life is not worth having.”
- He asks for nothing, but gives readily. He gives benefits and gifts, but hates to receive them, and hates to be in another’s debt, but will overpay a debt so as to turn the tables.
- Similarly, he remembers (and prefers to be reminded of) the services he has done for others, but not those he has received (for those things are reminders of having been in an inferior position, and the proud man prefers to be superior).
- He does not stoop but projects his dignity before people of high position and riches, but he behaves in an unassuming way towards ordinary folk, as it’s a vulgar thing to lord it over people below one’s station.
- He doesn’t exert himself for the sorts of honors most people strive for, but only for the best of the best. He’s a man of few deeds, but those few are fantastic.
- He’s a straight-talker. He respects truth more than people’s opinions of him, so he doesn’t hesitate to share his contempt and doesn’t waste time trying to be

- diplomatic. (This, amusingly, “except when he speaks in irony to the vulgar.”)
- He will not put himself in service to any so-called superior, but may choose to serve a friend.
 - He doesn’t much go in for admiring things, since to a great person like him, nothing else is particularly outstanding.
 - He doesn’t tend to bear grudges or remember wrongs against him.
 - He doesn’t gossip or praise or bad-talk others, mostly because he doesn’t much care about the things that typically motivate people to do these things.
 - He prefers to possess beautiful things of no particular use more than useful, profitable things.
 - He moves slowly and deliberately, not in a rush, and speaks in a deep, level voice.
 - He is, most assuredly, not he-or-she, though Aristotle doesn’t think he needs to point this out. The great-souled man is a great-souled *man*.

It’s almost like a James Bond-style action movie hero. And it reads more like a laundry list of what the great-souled man *would be* like than a description of what he *is* like. A fictional character, an avatar, The Übermensch.

1. ^

Analects of Confucius, IV.xi

2. ^

Joan Didion, [“Self-Respect: its source, its power”](#) (1961) *Vogue*

3. ^

Kwame Anthony Appiah, [The Honor Code](#) (2011)

4. ^

Tamler Sommers, *Why Honor Matters* (2018)

5. ^

Aristotle, *Nicomachean Ethics*, book IV, section 3

Notes on Moderation, Balance, & Harmony

This post examines the virtues of **moderation**, **balance**, and **harmony**. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

I also use this opportunity to go on a (related) tangent about the VIA Institute on Character, and its questionable focus on playing to character strengths over strengthening character weaknesses.

What are these virtues?

If you practice these virtues, you respect that each part of your life is meant to serve the whole of it, not the other way around.

Even with generally beneficial or neutral things, it's possible to take them too far. **Moderation** means you are on guard against this. Take [fitness](#), for example, the virtue I covered a few days ago. How can anyone complain about being more fit? But a compulsive focus on fitness as an end in itself can make you forget the whole point of being fit. "People who are always taking care of their health are like misers, who are hoarding a treasure which they have never spirit enough to enjoy."^[1]

If you are immoderate, you risk becoming a fanatic, an obsessive, a monomaniac. Such things in themselves might be mere eccentricities. But they are often accompanied by harmful neglect of other facets of life. **Balance** means that you have a good perspective on your life as a whole, and that you use this perspective to attend to the points that are most in need of attention, rather than letting habit or lack of moderation constrain your effort to where it's less helpful. Work/life balance, balance between focus on self and others, and balance between living-in-the-present, learning-from-the-past, and preparing-for-the-future are some common challenges.

Harmony is closely related to [integrity](#). Pursuing the musical analogy, when your life is harmonious, the different parts of it are attuned to each other; they are different instruments but playing the same symphony. Harmony is also not mere melody or monotone: being a "[Johnny One-Note](#)" isn't where it's at. [To everything there is a season](#), as the song goes.

Some related virtues include [temperance](#) (moderation in pursuit of pleasure), conciliation / cooperation / compromise (social harmony advanced through moderation in the pursuit of one's priorities), perspective, and flexibility.

Avoiding burnout by knowing how to pace yourself is one application of these virtues.

Aristotle made moderation a keystone of his virtue theory, in the form of the "[golden mean](#)." In his theory, *all* virtues are to be found at a moderate mid-point balanced between opposite extremes. For example, you can be anxiously over-sensitive to fear, or recklessly insensitive to it; the virtue of [courage](#) is the sweet spot moderately in-between.

These virtues apply in the intellectual realm as well. Moderation keeps you from becoming dogmatic, too attached to your theories. Balance keeps you from being close-minded and getting stuck in ideological bubbles; it encourages interdisciplinary thinking. The pursuit of harmony keeps you hunting for ways in which your ideas conflict, for the contradictions and paradoxes that are sometimes the first clues that you're mistaken about something.

How to strengthen these virtues

In investing, it's expected that periodically you should *rebalance your portfolio* so that your investments continue to match your risk tolerance and goals as they change and as you change. Do you ever rebalance your life with the same sort of deliberate, [rational](#) attention you would give to your retirement account?

Which hobbies still give you a good return on investment, and which just take up space in your closet? Are there parts of your life that seem at cross-purposes to each other? What are you currently neglecting that you would be wise to devote more attention to?

Where ought you to be investing your time and energy for the best return on investment? Where are you currently investing that time and energy? The first question is difficult to answer precisely and confidently, but the second should be easier: you just have to observe closely and take notes. Then you can ask: Is my answer to that second question a plausible answer to the first one?

What might you learn (and what might you change) if you periodically took the time to do an audit of your life?

The VIA Institute on Character

The “positive psychology” movement aims to build a version of psychological intervention that is designed less to fix broken people than to help ordinary people become more extraordinary. One way of thinking about it would be that remedial psychology is to positive psychology as physical therapy is to personal training.

The [VIA Institute on Character](#) uses a virtue-oriented approach to positive psychology. It calls the virtues “character strengths” and has identified the following set as those with good cross-cultural and -temporal support:

- **Wisdom**: including creativity, curiosity, [judgment](#), philomathy, and perspective
- **Courage**: including bravery, [honesty](#), perseverance, and zest
- **Humanity**: including [kindness](#), [love](#), and social intelligence
- **Justice**: including [fairness](#), leadership, and teamwork
- **Temperance**: including [forgiveness](#), [humility](#), [prudence](#), and [self control](#)
- **Transcendence**: including appreciation of beauty & excellence, [gratitude](#), [hope](#), humor, and [spirituality](#).

The Institute says that each person has certain “signature strengths” — a small set of key virtues that they are especially strong in and that they adopt as part of their identity. The Institute has created a personality test that’s supposed to tell you what your signature strengths are ([you can take the test on-line](#) and, if you fork over your

email address, you'll get a summary of your strengths & weaknesses along with an offer to buy a more complete results report).

Their work in trying to make character strengths and character strength-based interventions more rigorous and measurable is at the core of a lot of ongoing scientific research into the virtues. Bully for that.

But here's why I'm mentioning them today, in this post about harmony and balance: Having identified your signature strengths (and weaknesses), the VIA Institute then counsels exclusively that you play to your strengths. Here's a quote from a book associated with the Institute:

"The newest research is showing that techniques for helping people boost their strengths can have important advantages over techniques that focus on correcting their deficits."^[2]

No footnote, though, so I'm left wondering what this "newest research" might be. Some of the citations that I sometimes see mentioned in support of this idea are:

- A study of workers in New Zealand found those who gave more positive answers to questions like "I know my strengths well" and "I always try to use my strengths" were more likely to be "flourishing" (using various measures of things related to lifestyle, health, psychology, and employment).^[3]
- In *The Elements of Great Managing*, the authors say that you'll be more effective as a manager if you assign employees to tasks based on the employees' existing strengths, rather than trying to mentor their weaknesses.^[4]
- "The use of coaching principles to foster employee engagement" builds on that to assert that if an employee has a weakness in a certain area, rather than coaching that employee to become better there, you should encourage them to repurpose one of their existing character strengths to cover the weakness: "For example, if an individual has a strength in collaboration, but not in resilience, then the collaboration strength could be used to manage their lack of resilience by talking through their issues with their colleagues (ideally ones who have a strength in resilience)."^[5]

The examples I usually see deployed to support this idea tend to have to do with employee motivation, satisfaction, and engagement. (This may just reflect where most of the research is being done; practitioners of positive psychology struggle a bit to find a lucrative niche in which to practice their craft, and management/employee motivation seems to be one.) It makes sense that if your job tasks match the character strengths you feel the most competence in, you will have more engagement and satisfaction at your job.

I take issue, though, with the VIA Institute's seeming extrapolation of these results from the workplace to life in general. In life, you don't have as much opportunity to specialize as you do on-the-job. In life, all of the virtues are important, not just the subset in your job description. Trying to patch a virtue you don't have by repurposing one you do have can be an inefficient stop-gap solution, not a good long-term strategy.

There have been a few studies that directly compare people who try to extend the use of their existing strengths with people who try to strengthen their weaknesses, but the ones I have seen don't support the VIA Institute's claim that focusing on existing strengths is clearly superior. For example:

- One study randomly assigned 76 students to groups in which they would either work on expanding the use of their existing “signature strengths” or would work on a combination of using existing strengths and strengthening existing weaknesses. The authors say that both groups reported higher satisfaction-with-life scores following the work they did, but there was no significant difference between the two groups. They conclude that “focusing on relative character weaknesses (along with strengths) does not diminish—and may assist in increasing—life satisfaction.”^[6]
- A larger and better-controlled study (375 people assigned to either signature-strength work, strengthening-weakness work, or a placebo) also found that both styles of strengths-improving interventions were roughly equally effective (in this case, at increasing life satisfaction over a six-month follow-up period).^[7]

Another VIA-associated book, [*Character Strengths Interventions: A Field Guide for Practitioners*](#), has an overview of the literature on character strengths that is more nuanced in this respect than what you’ll see on their website. (It’s a little pricey, so I don’t own a copy yet, but I’ve seen some excerpts.)

Future research may change my mind about this, and the published research I’ve seen so far doesn’t strike me as definitive, but my current best guess is that having a full, broad set of virtues is important to human flourishing, and that it’s short-sighted to concentrate on those virtues you’re already competent in while avoiding work on the ones that need improvement.

For example, if you’re not very courageous, you may have a strong sense of [caution](#) or [prudence](#) that you have already been using to avoid frightening situations. Further relying on your caution or prudence to avoid what frightens you, rather than working on your courage, is a way of playing to your strengths, sure, but it’s also a way of cementing your weaknesses. You would be more capable, and would be able to use your prudence and caution in more valuable (not merely compensatory) ways, if you tackled your courageousness directly.

1. ^

Richard Griffith (impersonating Laurence Sterne), *The Koran* (1798)

2. ^

Ryan M. Niemiec, Robert E. McGrath, *The Power of Character Strengths: Appreciate and Ignite Your Positive Personality* (2019), p. 18

3. ^

Lucy C. Hone, Aaron Jarden, Scott Duncan, & Grant M. Schofield, [“Flourishing in New Zealand Workers: Associations With Lifestyle Behaviors, Physical Health, Psychosocial, and Work-Related Indicators”](#) *Journal of Occupational and Environmental Medicine* (Sept. 2015)

4. ^

Rodd Wagner & James Harter, *The Elements of Great Managing* (2006)

5. ^

Shane Crabb ["The use of coaching principles to foster employee engagement"](#)
The Coaching Psychologist (June 2011)

6. ^

Teri Rust, Rhett Diessner, & Lindsay Reade, "Strengths Only or Strengths and Relative Weaknesses? A Preliminary Study" *The Journal of Psychology* (2009)

7. ^

René T. Proyer, Fabian Gander, Sara Wellenzohn, & Willibald Ruch, ["Strengths-based positive psychology interventions: a randomized placebo-controlled online trial on long-term effects for a signature strengths- vs. a lesser strengths-intervention"](#) *Frontiers in Psychology* (2015)

Notes on Patience & Forbearance

This post examines the virtues of **patience** and **forbearance**. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What is this virtue?

If you have patience and forbearance you can gracefully put up with the usual slings and arrows that beset a human life. You can show patience in the face of boredom, suffering, difficulty, pain, wait, and annoyance. Forbearance usually means, more specifically, restraint in the face of difficult *people*. Someone with forbearance suffers fools gladly, because, well, why let their foolishness stop your gladness? A person with forbearance is slow to anger, and tolerant of insults and idiocy.

It is related to [self-control](#) / serenity / [good temper](#) (you don't lose your cool), perspective and [humility](#) (it's not all about you), charity / [forgiveness](#) / [mercy](#) / [clemency](#) / tolerance (you give other people more slack), [dignity](#) and grace (patience is a good look), [endurance](#), and stoic acceptance.

The virtue of *kṣānti* (Sanskrit) seems to cover both patience and forbearance.

What good is it?

Levius fit patientia, quicquid corrigere est nefas
("What cannot be quite cured, is made easier by patience")

—Horace^[1]

Patience makes it easier to acquire skills and to accomplish things that are difficult or time-consuming.

Impatience can make a bad situation worse, both by being an unpleasant thing in itself and by prompting us to make suboptimal [decisions](#).

Patience is also a factor in the “attention span” that everybody seems to be complaining about in this day and age. By helping us to resist the eyeball-capturing techniques of clickbait instant gratification and peripheral ²⁰notifications, patience helps us more [rationally](#) prioritize our time.

Patience makes it easier to avoid jumping to premature conclusions, and so it can help you to consider nuance, to change your mind, and to develop greater understanding of the unfamiliar.

(The Marshmallow Test is meant to measure a young child's patience or willpower in one particular way, and was once thought to have extraordinary predictive power

about that child's later quality of life. It apparently is yet another victim of the social science [replication crisis](#), however, so interpret it with caution.)^[2]

How to strengthen it

The theory of "[ego depletion](#)" (that people have a limited reserve to draw on for tasks like willpower, patience, and self-control, and that this reserve is depleted in various ways) is also disputed and remains under investigation. There does seem to be a folk intuition that patience is a depletable resource, as reflected in phrases like "I'm running out of patience." I find it easier to be patient with something when I am well-rested, well-fed, unwearied, and not beset by distractions and stress from other quarters. Attention to environmental factors like these might be a useful way you can indirectly improve your patience.

Thubten Zopa recommended a method of training in forbearance, when you encounter a difficult person, that doubles as a way of immediately bolstering your patience. Instead of seeing the difficult person as someone who is being a jerk and trying to ruin your day, see them as someone who is sacrificing their own mental stability in order to provide you with exercises to strengthen your patience:

Ask yourself, "Where did I learn this patience that I practice? I learned it from those who have been angry at me... Therefore, all the peace and happiness that I enjoy in this and future lives as a result of my practice of patience has come from the angry person... How kind this person is! How much benefit this person has given me!"^[3]

This reminds me a bit of the advice from Marcus Aurelius: "Say to yourself in the early morning: today I shall meet meddling, ungrateful, violent, treacherous, envious, and uncharitable men."^[4] You are less likely to react with impatience if you habitually factor in a certain expectation of people being difficult sometimes.

Long lines at the post office, medical appointments that leave you sitting in the waiting room long past the scheduled time, traffic jams... you can recharacterize all of these as opportunities to strengthen patience and forbearance rather than as excuses to throw a tantrum.

Any task you want to accomplish or skill you want to learn that takes time and that involves some frustration and difficulty can be also an opportunity to strengthen your patience. If you look upon such tasks and skill-building in this way, this may increase the value you get from them (not only am I accomplishing X, but I'm also building patience!).

Meditation in particular removes other distractions so that you and your impatience can meet face to face. If you stare down your impatience, boredom, and fidgetyness in a meditative context, you may find that these things are more paper tigers than they had first appeared.

1. ^

Horace, *Odes* I.24 ("To Virgil on the Death of Quintilius")

2. ^

Brian Resnick, ["The 'marshmallow test' said patience was a key to success. A new replication tells us s'more"](#) Vox 6 June 2018

3. ^

Thubten Zopa, *Virtue and Reality* (1998)

4. ^

Marcus Aurelius, *Meditations*, book II

Notes on Care

This post examines the virtue of **care**. It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What do I mean by care?

Care is another one of those words that has a wide range of common uses:

- You can care *about* something (be curious or concerned about it) or care *for* something (be personally invested in it or engaged in promoting its well-being).
- You can care for things (for example, doing a job *carefully* or *with care*, or being the *caretaker* for a building) as well as e.g. people, but these have different connotations.
- Care can be an affectionate sentiment ("I really care for you, Jane"), or assistive actions ("I'm caring for Jane as she recovers from the accident").
- Care can sometimes be a synonym for [caution](#) (be *careful*, *take care*).
- "Care ethics" grew out of a feminist critique of traditional justice-oriented ethics schemes (see appendix below).

The sense of care that I want to explore in this post is the sort of care that is directed toward people (or animals-as-pets, etc., but not mere things or animals-as-livestock) and that takes the form of actions that are meant to promote their well-being. The other meanings of care I've either already covered in my posts on [compassion](#) and [prudence](#), or may get to later if I get around to writing up virtues like concern, conscientiousness, affection, and [kindness](#).

When I decided to limit the definition of care I was using in this way, I worried at first that I had defined it so narrowly that it was no longer a "virtue" — a habit characteristic of a flourishing human being — and more of a "skill." I think what I have in mind as a virtue is something like "caring," as in "Of course Pat pulled the car over and got out to make sure everyone was okay. Pat's such a caring person." To have the virtue of caring is to habitually have the inclination to give care to those who need it, the discernment to know when and what care is called for, and the skill to offer care competently.

There's a lot of overlap between that and [compassion](#) as I described it. The way I imagine it, these two virtues blend into each other a bit. Compassion has more to do with feeling moved by another person's plight; care has more to do with acting to relieve it. But compassion that does not at least sometimes culminate in care can be just a sort of pity, not really compassion. And often the motivation for offering care is a compassionate one, though it's also possible to offer care dispassionately, without being moved by compassion but by something else, such as duty, filial piety, or professional responsibility. (Part of the key to what makes the television character [Dr. Gregory House](#) successfully eyeball-catching is his remarkable cleverness in giving care to the sick combined with his jarring absence of compassion.)

As a further refinement, I suggest that you are offering *care* when you give people help that is essential on a basic level (definitions of “essential” and “basic” left as an exercise for the reader), and when you help them in a way that they cannot easily do themselves. So it may be part of care to help a very young child tie his shoelaces, or to help a person on crutches to bring in the groceries; but it’s probably something else to help someone organize their Magic cards, or to help someone load the dishwasher who otherwise would do it themselves just fine. If you serve someone a meal who cannot otherwise afford one, you are offering *care*; if you serve someone a meal because that’s what they ordered off the menu at your restaurant, you’re offering *service*. Care might in this way be considered a specialized version of “helpfulness” that applies in cases of more urgent humanitarian need. But that’s not to say that it’s only used to respond to crisis, debility, and other things-going-wrong. You can also exercise care in a nurturing way, as when you teach a child to read, or, maybe, if you support a hesitant friend as they quit their dead-end job to follow a dream.

Care also often includes some responsibility or stake in the well-being of whoever you are helping. This distinguishes it from “kindness” which includes more ephemeral, less invested interventions. You can be kind to someone by picking up something they dropped and handing it to them, or by telling them they left their headlights on. “Care” is more involved than that, more of a commitment. Care often includes giving convincing reassurance to whomever you’re caring for that you’re prepared to stick by them until help is no longer needed.

Sometimes a person may simply volunteer to care for someone else who needs it, just because. Other times, a person may take on a professional caring role like that of a paramedic, nurse, hospice worker, or teacher, and be caring as part of their professional obligations. Other times, a person may exhibit care because they are in a personal relationship of a sort in which care is an essential part: the care of parents for young children, or children for old parents for instance. Each of these scenarios calls for the elements of the virtue of care.

Okay, deep breath; that was a lot of defining.

How to be caring

I come from a family of caregivers. My father and brother are both social workers; my mother taught learning-disabled children. But I... became a software engineer and then a technical writer. So I had to hunt around for good remedial advice on how to be caring.

Much of what I found is specific to the caring *professions* and is meant as vocational training. I also found a useful set of short videos [at this YouTube channel](#) that teach specific care-giving skills for lay people who need to take care of someone who is sick or disabled or who is losing capabilities due to aging or dementia.

I did not find much about how to develop a caring inclination. Some of the advice on how to develop [compassion](#) is probably also applicable to this aspect of caring. To care for someone else can be a big investment of attention and time and sometimes of material resources; you are unlikely to do something like that unless you have a strong motive — either you find caring itself to be rewarding or to be a crucial part of your self-image, you value the well-being of whomever you are caring for, or the care you give is packaged up in something else you value like your profession.

Having that inclination is one part of caring. The two other parts are knowing when care is called for, and giving care competently. For the first of these, being [mindfully observant](#) is helpful. The widely-discussed [Wang Yue](#) case in China, in which a child was struck by a car and her dying body was ignored by multiple passers-by, may have been an example of the [bystander effect](#) causing people to blind themselves to an urgent need for care. Being caring may require being bold and discerning enough to be that one person who breaks through the bystander inertia. [Empathy](#) and good communication / listening skills also come to your assistance here, both to determine what sort of assistance might be needed, and what would be respectful to offer. For example, fairly or unfairly, in the wake of a suicide people will sometimes wish they had been astute enough to recognize something the victim did ahead of time as a disguised “call for help.”

The third part of caring is to competently give care. For this you need confidence. Part of that confidence is general self-esteem (not thinking of yourself as someone who always mucks things up), and part comes from practical skills that cover the sort of care required. The more caregiving skills and resources you have, the more opportunities you will have to offer care competently, and the more confidence you will have when offering it. You also need [prudence](#)—both in the sense of having good [decision-making](#) skills, and in being cautious. You can err on the side of over-confidence as well: for example, injuring someone by thinking you know how and when to give CPR because you saw it on TV once.

Another skill that caregivers report as being important is responsibility / dependability / reliability. When you step up to give care, to do so responsibly means to see it through until the person either no longer needs your help or until someone else can take your place as caregiver. For this reason also, [patience](#) can be important, and the flexibility to be able to adjust to someone else’s needs. Another is competence in offering comfort, consolation, encouragement, and reassurance. Being calm and composed can be important for assessing the need for care, for offering it competently, and for being a reassuring presence to the person you are caring for.

So this is another virtue with a lot of moving parts. Any time I have an opportunity to offer care, I am likely to do so with a mixture of relative competence and incompetence. I would be wise to review how I did, and try to identify those areas in which I was weakest so I can work on them. Was I slow to recognize the need? reticent to offer help? did I get flustered or impatient? did I miss a clue I should have noticed?

“Self-care”

Is there such a thing as caring for yourself—seeing yourself as a target of your own nurturance, etc.—or is that just speaking metaphorically?

Being in the habit of taking care of yourself rather than “letting yourself go” is certainly important to living a flourishing life, so it’s a virtue of some sort. The question is whether it belongs under the umbrella of “care” or if it’s its own thing that needs its own investigation. I can see arguments either way.

Being-cared-for can be seen as a supplement for taking-care-of-yourself when that fails for whatever reason or has yet to maturely develop. Seen that way, caring is *by default* taking-care-of-yourself, and taking care of others is what you do when you have a surplus of that ability and they have a deficit.

A surprising amount of the advice for caregivers that I read or viewed had to do with self-care. Apparently it is an occupational hazard for caregivers to neglect their own care while offering care to others, and so they need to remember to give this extra attention. This might suggest that the same mental faculties are used for self-care and care-for-others and that caring for others can absorb or fatigue those faculties to the extent that self-care suffers.

Appendix: “Care ethics”

Some feminist ethicists have suggested that because the discipline of ethical philosophy was male-dominated, this unbalanced the ethical systems developed within it to prioritize impartial, impersonal, justice-based ethics that met the sort of ethical challenges that came up in male-dominated fora (e.g. politics, law, war). To restore the balance and raise the priority of virtues that are more important in traditionally female-dominated spaces, a discipline of [“care ethics”](#) was developed.

To oversimplify: Care ethics is more contextual and situation-specific, and less generalizable and abstract than many other varieties of ethical philosophy. In care ethics, people are not seen as interchangeable, essentially equal, and autonomous, but as unique, as validly special in different ways to different people, and as necessarily enmeshed in a variety of crucial relationships.

Notes on Attention

This post examines the virtue of **attention** and its facets like awareness, discernment, mindfulness, presence, focus, and concentration.

It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is this virtue?

This virtue has to do with having command over your attention, and using that command skillfully to attend to the right things in the best manner. Different terms associated with varieties of attention are used in a variety of ways by different authors; I'm going to use the following terms to differentiate them:

- **Mindfulness** (presence, being observant) is the ability to consciously perceive the full breadth of what is currently available to your senses.
- **Awareness** (understanding, comprehension) is the ability to contextualize what you are experiencing, including in that context things that are relevant but not currently available to your senses.
- **Discernment** (orientation), in this context, is the ability to direct your attention to specific things in your field of awareness, and to identify the most important or relevant things.
- **Focus** (absorption, concentration) is the ability to narrow the attention so as to apply it in a more detailed and penetrating way for sustained periods of time on some chosen part of your present experience.

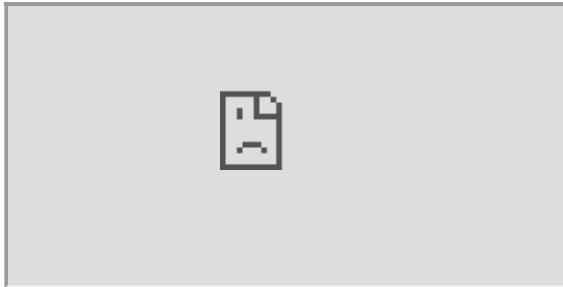
There is also a sort of meta-skill (I don't know a good name for it) that covers the ability to transit in a timely way between these various modes of attention: to be mindful when it's time to be mindful, to incorporate awareness when awareness is called for, to notice when it's time to orient yourself toward something in particular, and then to apply your focus where it is needed for as long as it is needed.

The virtue of attention includes all of these skills as components, and you have that virtue if you characteristically practice these skills proficiently.

Some related virtues or skills include savoring (attending to delight), elevation/awe/wonder (attending to the sublime), perspective (a facet of awareness), curiosity and intellectual humility (so you're not closed to mindfulness), acceptance/surrender (willingness to see reality as it is), vigilance/alertness (discernment of potential dangers in particular), and self control (which helps you maintain focus). Buddhist traditions are especially attuned to the nuances of attention, and have fine-grained terms of their own to describe aspects of it (e.g. mushin, jhāna) that are difficult to map to common English terms.

Iris Murdoch made a case for attention being central to morality. On the one hand, you can train yourself to attend to the morally relevant details of whatever you are confronted with in your day-to-day life, and this attention focuses and orients you appropriately so that the correct action naturally follows.^[1] On the other, you can attend periodically to "things which are valuable: virtuous people, great art, perhaps... the idea of goodness itself" and this can displace your attachment to less-valuable things and make you more able "to act well 'when the time comes'".^[2]

Narrow focus and broad mindfulness are mutually-exclusive, on opposite ends of a range. Opposed to both of these is “daydreaming”—a state in which you are neither very aware of what’s going on around you, nor focused on anything in particular. In some interpretations, this state of mind is the brain’s “idling” somewhat-lower-energy state that it returns to by default when there are no pressing mental demands. The [“default mode network”](#) in the brain is active when you are idly daydreaming or recalling memories: in other words, when your mind is retracing familiar paths rather than taking in new information. However, this mode of thinking may be important for incorporating and retaining previously-acquired information and insights—so maybe rather than thinking of it as a *failure* of mindfulness and focus, it would be better to think of it as an additional *variety* of “attention” that is healthy to exercise in moderation.^[3]



Mindfulness

“Mindfulness is nonconceptual awareness. Another English term for *sati* is ‘bare attention.’ It is not thinking. It does not get involved with thought or concepts. It does not get hung up on ideas or opinions or memories. It just looks. Mindfulness registers experiences, but it does not compare them. It does not label them or categorize them. It just observes everything as if it was occurring for the first time. It is not analysis that is based on reflection and memory. It is, rather, the direct and immediate experiencing of whatever is happening, without the medium of thought. It comes before thought in the perceptual process.” —Henepola Gunaratana^[4]

Mindfulness is the attempt to take in as much as possible of your raw immediate experience. “Raw” and “immediate” are exaggerations, as our experience is mediated by our senses, which do a great deal of processing to our perceptions before they become available to consciousness. But the point of the mindful state is to approach as near as we can the raw material of our current experience, as it is before we begin to try to make sense of it all or fit it into our agendas.

If you’ve ever had the experience of, say, looking all over the place for your scissors only to discover that they were right there on the table in front of you the whole time, you’ll know that our preconceptions about what is relevant or true can hide from us what actually is. When you are being mindful, such preconceptions have fallen away so the *is* shines through.

(There’s unfortunately a lot of ambiguity in how the term “mindfulness” gets used—for example, some examples of what people call “mindfulness meditation” have more to do with what I’m calling “focus.”)

There is a difference of opinion, or maybe a difference of perspective, about the extent to which you really can take in a broad spectrum of sense input simultaneously. Some people describe mindfulness in a way that suggests that you’re really hitting your attention with a firehose of everything all at once: sight, sound, feeling, and so forth. Others ([Daniel Ingram](#) for example) say that you really can only attend to a single sensation at a time, so that each individual sensation excludes the others at the moment it is being perceived attentively, and

mindfulness of this sort is better understood as attention to a rapid *series* of specific sense impressions than as a *panorama* that includes them all.

Sometimes the phrase “beginner’s mind” is used to describe this sort of attention. The idea here is that when you see something for the first time, you don’t know what are the most important, relevant, or atypical aspects of it: Everything is equally novel and seems potentially important. As you become more experienced with some category of things, you come to attend to only the most relevant, tractable, or off-kilter aspects about them. If you can reactivate the “beginner’s mind” you will notice things that you might otherwise ignore.

Three major obstacles to mindfulness of this sort are 1) a rush to reabsorb your raw experience into the familiar and understood by identifying it, categorizing it, telling yourself stories about it, or interpreting it in terms of your agenda, 2) focusing-in on some interesting detail or thought so as to shut out the rest of experience, and 3) slipping into daydreaming and letting your experiences pass by unattended-to.

Another obstacle to mindfulness is that it can be difficult to come up with a good reason to practice it. If you need to focus, or need to be vigilant, there is usually some specific articulable reason (e.g. I need to study for an exam, or I need to watch out for road hazards). Mindfulness is typically more open-ended and not obviously practical. You don’t know ahead of time quite what you’ll find there, and that’s part of the point. This can make it more difficult to justify or to find time to practice.

Awareness

“He sacrificed to his forefathers as if they were present; he sacrificed to the gods as if the gods were present. The Master said: ‘For me not to be present at a sacrifice is as if I did not sacrifice.’ ” —Confucius^[5]

Awareness is like mindfulness in that it has a broad range of attention. It differs in that it interprets what is being perceived through the filter of concepts, categories, interests, agendas, and so forth. It looks at the world not only in terms of what it is, but what it *means*. It (or sometimes, focus) is usually what people mean by “paying attention.”

The absent-minded professor, or the person who walks into a utility pole while looking at their phone, are examples of people who are lacking in awareness.

Awareness includes an awareness of context, which requires memory skills and also sustained attention. *Social* awareness is a valuable skill with lots of nuance, and often requires you to [empathically](#) model other people, which brings in yet another set of virtues and skills.

If you are aware, you see the big picture, and also are more apt to notice any anomalies — things that don’t seem to fit or that require explanation. Vigilance is a variety of sustained awareness in which you attempt to be constantly alert for particular sorts of anomalies. It is related to “alertness,” though that sometimes has connotations more in line with more general “wakefulness.” People tend to have difficulty maintaining [vigilance](#), and in experimental conditions the ability to notice unusual signals in an otherwise uninteresting background declines quickly.

Discernment

“If you cut up a large diamond into little bits, it will entirely lose the value it had as a whole; and an army divided up into small bodies of soldiers, loses all its strength. So a great intellect sinks to the level of an ordinary one, as soon as it is interrupted and disturbed, its attention distracted and drawn off from the matter in hand; for its

superiority depends upon its power of concentration—of bringing all its strength to bear upon one theme.” —Arthur Schopenhauer^[6]

Discernment is the ability to pick, from the multitude of perceptions in your awareness, those that are most potentially interesting or relevant. It is “to notice” what is notable. Orientation is to train the attention on whatever it is that you have singled out as interesting. Orientation is what you do when, for example, you choose one of many conversations in the room to follow.

You may orient your attention deliberately and consciously, or more reflexively and unconsciously. Stage magicians master the art of manipulating attention reflexes such that their audience orients away from the trickery to notice instead the deceptively salient red herring.

Focus

“No man is in any degree fit for either business or conversation, who does not command his attention to the present object, be it what it will. When I see a man absent in mind, I choose to be absent in body; for it is almost impossible for me to stay in the room, as I cannot stand inattention and awkwardness.” —Lord Chesterfield^[7]

The aspect of attention I most often hear people complain about lacking is focus or concentration.

Focus can be deliberate and effortful, or it can be automatic and effortless. Something that is inherently fascinating, outrageously shocking, pruriently intriguing, or potentially threatening can cause us to focus on it without us consciously deciding to do so. More ordinary things require us to make a deliberate effort if we want to concentrate on them, and more effort still to sustain that concentration.

The amount of time you can effortfully focus on something before becoming distracted (by something else or by mere boredom) is your “attention span.” Characteristics of the thing or task being focused on, and of your previous experience with it, can change your attention span, as can factors like fatigue, hunger, environmental distractions, whether you are being observed, and stress. (This suggests that things like fitness and tranquility may be helpful to attention.)

People who have extreme difficulty in the ability to focus, or whose difficulty causes problems in particular areas of life (e.g. schooling) may be diagnosed with attention deficit hyperactivity disorder, for which a variety of treatments are available.

The “flow” state is an elusive Shangri-La in which the activity we feel we ought to be attending to is also one that holds our effortless focus, such that we can devote our sustained effort to the activity itself rather than to maintaining our concentration.

“Hyperfocus” happens when something has so captured our attention that we tune out all else, to an extreme level—for example, someone playing a video game who does not notice when someone else in the room calls their name repeatedly or when the pizza deliverer rings the doorbell.

Being able to shift focus gracefully from one thing to another, and to multitask in such a way that you prioritize your focus well between things, are also important parts of the skill of focus.

How can you improve your attention?

"Every time that a human being succeeds in making an effort of attention with the sole idea of increasing his grasp of truth, he acquires a greater aptitude for grasping it, even if his effort produces no visible fruit." —Simone Weil^[8]

The previous sections of this page establish some of the terminology and concepts we can use to describe facets of attention. There is a lot of interesting scientific research out there about the neural correlates of these facets, the course of their development in children, how to measure them, and how various pathologies interfere with them. I found much less of such research about how a typical person can improve their attention control in general.

You can typically attend to things better when you are in a state of arousal. As Samuel Johnson noted, "when a man knows he is to be hanged in a fortnight, it concentrates his mind wonderfully."^[9] Athletes will sometimes use rousing music (or something like a [haka](#)) to pump themselves up before an athletic performance in order to sharpen the mind.

[Amphetamines](#), which artificially generate arousal, are the go-to drug for treating attention deficits, and also for enhancing vigilance, for example in soldiers in battle. The milder stimulant caffeine is used more casually to aid alertness, and eugeroics (e.g. modafinil/provigil) to improve wakefulness. Folk interventions like "pinching yourself to stay awake" are also in this category.

Psychedelic drugs can heighten and discombobulate the senses to the extent that everything takes on an appearance of unfamiliarity, and a sort of starry-eyed "beginner's mind" results that can clear out the cobwebs in fascinating ways ("have you ever *actually looked at your hand, man?*").

Environmental changes can make it easier to attend effectively to the right things. Removing clutter and other distractions can make attention less difficult, for which the virtues of orderliness and [simplicity](#) can help.

User interface designers take pains to make important parts of interfaces more easily discernible with bright colors, blinking lights, arresting sounds, and so forth. On a more potentially sinister note (as highlighted in e.g. [The Social Dilemma](#)), media companies that commodify human attention are evolving their platforms in sophisticated and automated ways to make them more captivating.

We can also manipulate the user interface of our own surroundings to make more important things more prominent: For example, store your resistance bands or barbells out where you can see them, or on top of your favorite lounging spot, so as to more easily remind you to do your exercises.

But these things have a gimmicky, stop-gap feel to them. What would be nice would be if we could improve attention in a sustained way, such that we become characteristically skillfully attentive.

Meditation

There are meditation practices that are designed to improve attention or to expand its capabilities in various ways. There are forms of meditation that exercise mindfulness, others that exercise awareness, others that exercise concentration.^[10]

The point of attention-related meditation practices is often not to strengthen attention for its own sake, but so that you can then use attention for the purpose of developing insight into reality or for the purpose of entering into a variety of interesting altered states of consciousness. (Though more recent developments like Mindfulness-Based Stress Reduction also harness mindfulness meditation for gently therapeutic relaxation.^[11]) That said, I know of no other discipline that has put so much effort into mapping out practical interventions

that ordinary people can use to strengthen and expand attention, so even if you have no interest in seeing the sights around nirvana, you might want to investigate.

One of the meditation practices I do from time to time is a five-senses scan. For example, I attend for several minutes at a time to the experiences of seeing, then hearing, then tactile feeling, then tasting, then smelling. I sometimes try to imagine when I do this how I would feel if the sense I am attending to had just switched on for the first time: for example, if I'd never tasted anything before and suddenly my mind began to experience taste sensations. What would that be like? How would I distinguish these new sensations from other ones? How would I discern the contours and boundaries of this new sense? This perspective helps me to drop some of my preconceptions about what the sense is—what it means, where it comes from, how it works, what it suggests—and helps me to attend as closely as I can to its uninterpreted original essence.

A skill that is useful for concentration is the ability to notice when you have lost your focus so that you can restore it. If your mind begins to wander off, it typically does so without announcing its intentions first, and in a devil-may-care way that may leave few clues that anything has gone awry. It can be tricky to notice that this has happened so that you can correct for it. This is something for which concentration meditation practices may help. Such practices allow you to catch your mind wandering off again and again, and bring it back to whatever you are trying to focus it on again and again, until the state of wandering becomes more salient and the process of return more automatic.



Agnes Callard
@AgnesCallard

Texting with a friend, explaining to him how I'm more observant than he is, all the while unbeknownst to myself I'm traveling the wrong way on the metro. I travel several stops, vehemently texting about my magnificent powers of observation, before I realize.

4:16 PM · Jun 18, 2022 · Twitter for iPhone

[If you're not being observant, would you even notice?](#)

You can also incorporate some of the same techniques that you would use in meditation into the more mundane activities of your day to day life. [\[12\]](#)

Active, reflective attention

[Active listening](#) and particularly [reflective listening](#) (listening to someone speaking with the intention of being able to reconstruct and regenerate what the other person has said) are ways of improving attention to spoken communication.

William James suggested that this sort of approach might have broader application: One form of attention he identified was “the reproduction of the sensation from within.” He recommended that “the habit of reading not merely with the eye, and of listening not merely with the ear, but of articulating to one's self the words seen or heard, ought to deepen one's attention to the latter.” And he put this into practice successfully in his own life:

"I can keep my wandering mind a great deal more closely upon a conversation or a lecture if I actively re-echo to myself the words than if I simply hear them; and I find a number of my students who report benefit from voluntarily adopting a similar course."^[13]

A use of this technique in meditation is called "[noting](#)" in which as each new perception arises, you also consciously make note of it in an objective way ("seeing... itching... hearing... remembering..."). This is meant to help you keep your focus on whatever is currently present to the mind, as well as to become more aware of your mental activity.

Related to active listening is curiosity or inquisitiveness. The more you want to know, the more eagerly you will listen. This too applies to more than conversation. If you can develop more curiosity about something, you will find it easier to attend to it. William James again:

"Try to attend steadfastly to a dot on the paper or on the wall. You presently find that one or the other of two things has happened: either your field of vision has become blurred, so that you now see nothing distinct at all, or else you have involuntarily ceased to look at the dot in question, and are looking at something else. But, if you ask yourself successive questions about the dot, — how big it is, how far, of what shape, what shade of color, etc.; in other words, if you turn it over, if you think of it in various ways, and along with various kinds of associates, — you can keep your mind on it for a comparatively long time. This is what the genius does, in whose hands a given topic coruscates and grows."^[13]

Other techniques

Some people find that making explicit lists or plans helps them to focus. If you've already gone to the trouble to add something to a list as a focal item, you don't later have to keep reminding yourself that it's important. If you have already made a list of the things that need doing, you can concentrate on the current thing on the list, not distracted by anxiety about what else you maybe ought to be doing instead.

Short "centering" rituals can help to focus the attention. A friend of mine said he swears by "close eyes, two quick inhales, one long exhale." In baseball, it's common to see a batter approach the plate, tap their bat in a particular place, adjust their wristband just so, and go through a routine of such odd tics before settling in to concentrate on the pitch. Athletes will also sometimes use various relaxation techniques (progressive muscle relaxation, breathing exercises, meditation) before a game or performance if stress or anxiety interferes with their ability to focus.

["Attention restoration theory"](#) holds that periodic exposure to nature in an unstressful, undemanding way can restore attention capacity. An attention-restorative environment is one that is inherently fascinating or awe-inspiring, away-from-it-all (from life's cares in particular), extended in space (you can keep moving through it or scanning along it), and comfortable. ([W](#))

Simply taking breaks to stare out the window and give your mind time to go wandering can help to rejuvenate it for more concentration. Some [brief, vigorous exercise](#) can also cause enough arousal to improve attention. Sometimes people get up from what they are doing to take a walk or stretch, or occupy their mind in some more meandering task like journaling. A friend told me:

I only ever fill my water cup at work halfway full, so I'm constantly getting up to get a drink. That allows a mental pause and prevents a physical slump every time I get up to have more water. I don't get bored as easily and I don't have computer shoulder slump. When my brain feels stuck my body Pavlovs to go get some water.

If some thought — like a worry or an impending task or something you don't want to forget — is interrupting your concentration, writing that thought down somewhere where you know you'll see it when you have time to deal with it can make it easier to put it aside for the purposes of present concentration. Similarly, you can set an alarm to remind you of something so you don't have to use mental energy to remember it. A friend of mine said she has an alarm to tell her when it's lunch time: "on a low motivation day it stops me from constantly checking the clock to decide when lunch is, because the alarm tells me."

Similarly, when I meditate I use a timer that has chimes that go off a few times during the meditation session. I know that one of the chimes will tell me unambiguously when I've been on the cushion as long as I intended to be, so I don't need to be distracted by watching the clock. The occasional other chimes are not so distracting as to interrupt my meditation if I'm "in the zone" but are distracting enough to interrupt my wandering mind if it's off in the bushes somewhere making mischief.

When I asked friends for their concentration tips, I heard from one person that they concentrate better when "overstimulated with media" — this person likes to keep music and television going, and to always have a social media tab ready to jump to. It may sound paradoxical, but when he figured all this out, he says, his productivity went up. He told me, "It's like an anxiety relief thing, where the anxiety is, I guess, that I'm going to get bored while I'm waiting two seconds for the work program to load..."

Coming up with artificial rewards (gold stars, treats, checking things off your list) can help improve focus both by giving you something to shoot for and by giving you positive reinforcement for maintaining attention. A friend of mine who is in grad school tells me she repurposed a lapping stopwatch to help her concentrate on her academic reading: "A page is a lap. I've timed the approximate wpm that I read for textbooks, and at the beginning of an assigned reading I will write the approximate time needed for me to read it. Then when I read, I set my phone next to me with the stopwatch running."



Zach Weinersmith @ZachWeiner

...

I'm sure this violates quite a few pedagogical standards, but lately my daughter insists that I tell her she has no chance of getting her math homework done. When I do, she literally gets through it three times as fast with higher accuracy, while shouting "We'll see about that!"

6:16 AM · Jan 3, 2021 · Twitter Web App

As the above-pictured example demonstrates, you can make a dull task more interesting by adding competition, camaraderie, or some arbitrary private goal.

Something that is unfamiliar can be difficult to attend to (once the initial novelty wears off) just because it seems confusing or meaningless. For example, a foreign language just seems like a tangled mess of syllables at first: it's hard to attend to it because any segment of it seems interchangeable with any other. But as you become more familiar with something, you learn how to pick out what is notable about it. This suggests that becoming more knowledgeable about something may help you to be attentive toward it.

For the particular variety of attention involved in *savoring*, Fred B. Bryant & Joseph Veroff recommend these five techniques: 1) sharing your experience verbally with others, 2) deliberately building memories from your experience, 3) congratulating yourself, 4) attending to the details, and 5) attempting absorptive full-attention in what it is you intend to savor.^[14]

Being well-rested and unfatigued seems to make a big difference in the ability to maintain attention. This suggests both that you can improve your general attentiveness by attending to the quality and quantity of your sleep, and that it might be worthwhile to consider the schedule on which you do things: moving things that require more attention to the part of the day when you are most “fresh.”

There are [apps](#) that are meant to make your various beeping buzzing devices somewhat less distracting by temporarily blocking your access to certain sites (or allowing you access for only a limited time), emitting neutral noises to drown out distracting sounds, and so forth. You can also spend some time adjusting the notifications settings of your various applications, and adding filters to your inbox, so they don’t poke at you so frequently. There are always the blunt-force tactics of closing the laptop, or muting the phone and putting it face-down on the table. My own favorite technique is to “forget” my phone in the other room.

There is some, so far inconclusive, evidence that modest increases in CO₂ can impair cognitive performance. It probably wouldn’t hurt to open a window.

With all that is known about the brain networks involved in attention and in mind-wandering, this seems to be something that would be ripe for biofeedback techniques. I was able to find promising examples of studies of biofeedback to treat ADHD in children, but didn’t immediately find examples of adults using biofeedback to improve attention.^[15]

Having stronger working memory helps you keep your goal in mind as you tick off the tasks that help you accomplish the goal, and thereby reduces the influence of distractions. There was some initial excitement about enhancing attentiveness by [improving working memory via computer-assisted training](#) but this didn’t seem to hold up.^[16]

It can be hard to attend to something if you find it repulsive, frightening to contemplate, or painful. Sometimes this can be unfortunately distracting, causing you to avert your attention from something you need to attend to. In such a case, it can sometimes be helpful to try to devote some time when you can deliberately, and in a safe environment, focus on just the aversion-causing thing (or on the feeling of repulsion/fear/pain itself). Sometimes when you look at such a thing directly, “staring it down” as it were, you can defang it so that it causes you less distress. When you’re sitting in meditation, for example, when you get bored, or irritated, or frustrated, or your back starts to hurt, or your nose itches, you’re encouraged to incorporate those things into your meditation by making them the focus. This has a way of making them seem less like imperatives and more like ordinary phenomena.

Learning how to be comfortable being assertive can help your concentration. Just knowing that it’s okay to say “not now; I’m busy” can be enough to free up the mental space you need to concentrate on what you’re doing.

If you have difficulty with attention because you are distracted by persistent unwanted thoughts, traumatic memories, hallucinatory voices, or things of that nature, professional psychological or psychiatric evaluation is probably your best bet.

In a recent essay, chess grandmaster Jonathan Rowson explored how chess is “a socially permissible pretext to concentrate for several hours at a time” that “can teach us more about what concentration really means.” So chess may be another sort of concentration-strengthening exercise worth investigating. Rowson concludes:

"Our problem today is not that we don't or can't pay attention, but that the systems and structures of society oblige us to pay attention so frequently and fleetingly that we cannot in fact concentrate. Lacking an ability to concentrate, it's a struggle to construct and maintain a coherent and autonomous sense of self, which leaves us at the mercy of digital, commercial and political puppeteers. Without concentration, we are not free."^[17]

1. ^

Iris Murdoch, "The Idea of Perfection" (1962) in *Existentialists and Mystics*, pp. 299-336. ("[I]f we consider what the work of attention is like, how continuously it goes on, and how imperceptibly it builds up structures of value round about us, we shall not be surprised that at crucial moments of choice most of the business of choosing is already over. This does not imply that we are not free, certainly not. But it implies that the exercise of our freedom is a piecemeal business which goes on all the time and not a grandiose leaping about unimpeded at important moments. The moral life, on this view, is something that goes on continually...")

2. ^

Iris Murdoch, "On 'God' and 'Good'" (1969) in *Existentialists and Mystics*, pp. 337-362.

3. ^

Markham Heid, "[Why Your Brain Needs Idle Time: Some vital brain functions demand downtime](#)" *Elemental* (Feb. 14, 2019)

4. ^

Henepola Gunaratana, *Mindfulness in Plain English* (2011), p. 134

5. ^

Analects of Confucius, III.xii

6. ^

Arthur Schopenhauer, *On Noise* (1851)

7. ^

Lord Chesterfield, *Letters to His Son*, 22 September 1749

8. ^

Simone Weil, "Reflections on the Right Use of School Studies with a View to the Love of God" (1942)

9. ^

James Boswell, *The Life of Samuel Johnson* (1791)

10. ^

For example, two popular families of meditation techniques are *samatha* or "focused-attention" and *vipassana* or "open-monitoring." The former is more oriented toward concentration/focus while the latter is more oriented toward awareness/mindfulness/vigilance. See A. Lutz, et al. "Attention regulation and monitoring in meditation" *Trends in Cognitive Sciences* (2008) pp. 163-169

11. ^

Jon Kabat-Zinn "An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results" *General Hospital Psychiatry* (1982) pp. 33-47

12. ^

Scott Crabtree and Chris Wilson "[Mindful for a Moment : Integrating Attention into a Busy Day](#)" *Positive Psychology News* (June 6, 2012)

13. ^

William James, *The Principles of Psychology* (1890)

14. ^

Fred B. Bryant & Joseph Veroff, *Savoring: A New Model of Positive Experience* (2007)

15. ^

Vincent J. Monastra, Steven Lynn, Michael Linden, Joel F. Lubar, John Gruzelier, & Theodore J. LaVaque "[Electroencephalographic Biofeedback in the Treatment of Attention-Deficit/Hyperactivity Disorder](#)" *Applied Psychophysiology and Biofeedback* (June 2005)

16. ^

Torkel Klingberg, Hans Forssberg, & Helena Westerberg, "[Training for working memory in children with ADHD](#)" *Journal of Clinical and Experimental Neuropsychology* (2002)

Charles Hulme & Monica Melby-Lervåg, "[Current evidence does not support the claims made for CogMed working memory training](#)" *Journal of Applied Research in Memory and Cognition* (2012)

17. ^

Jonathan Rowson, "[Concentrate! \(Playing Chess is an Essential Life Lesson in Concentration\)](#)" *Aeon* (6 January 2020)

Notes on Rationality

“The knowledge of the theory of logic has no tendency whatever to make men good reasoners.” —Thomas Macaulay^[1]

This post examines the virtue of **rationality**. I’ve been dreading this one.

I have been writing a sequence of posts about virtues, and strategies for strengthening them in ourselves. My [notes on notes on virtues](#) post explains in more detail what I’m trying to accomplish and why.

I’m following the lead of virtue-oriented traditions and philosophers of the past as I choose which virtues to highlight. Many of these traditions included rationality (or some closely-related virtue like reason or love-of-truth). I agree that it’s a virtue (or at least that it can be useful to consider rationality in the form of a virtue) and a valuable one. But what can I add to the rich discussion about rationality that already exists here on *LessWrong*?

I decided that the best way I could advance the conversation would be by, instead of discussing *rationality* as a virtue, discussing rationality *as a virtue*.

What is a virtue?

“He who knows the truth is not equal to him who loves it, and he who loves it is not equal to him who delights in it.” —Confucius^[2]

I’m using the word “virtue” in the way I’m most familiar with it from the virtue ethics tradition.

“Virtue” can have implications that I don’t intend: In popular use, virtue is morally correct behavior, something urged on by the angel on your shoulder, in opposition to a devil’s temptingly delicious vice. You practice this sort of virtue because you’re supposed to, or because God is watching, or because it is your unfortunate duty. Popularly, “virtue” can be prim, naïve, old-fashioned, maybe a little ostentatiously holier-than-thou. It also is more often singular than plural (do you exhibit virtue?) or is conceived of as something of a continuum (a person might be more-or-less virtuous).

By contrast, in the virtue ethics tradition I’m most familiar with, “virtues” are a variety of character traits. Those character traits that tend to help you to succeed at living an excellent human life (or that are themselves ways of living excellently) are virtues; those that interfere with this are vices; any others are just part of life’s rich variety. To consider rationality as a virtue is to consider it as one of the human excellences that individuals can strive to practice characteristically.

Other things you can probably say about virtues:

- They may be directly under rational, deliberate control, or they may be more subconscious inclinations driven by procedural memory, but they are all at least somewhat voluntary and malleable with effort. So something like “height,” for example, is not a virtue, even if it turns out that certain heights are better than others for human flourishing.

- A virtue is a habit of *choosing* to do something (or to do something in a certain way). Habitually doing something *unchosen* (e.g. because it's some unconscious process like digestion, or because you have a gun to your head) is also not an example of a characteristic habit that might be a virtue.
- A virtue is in your own best interests by definition. For this reason, if you're wise and on top of things, you will be virtuous self-interestedly, and not in a spirit of sacrifice or self-denial.
- That said, there are many ways we can go astray and pick up habitual vices that stunt our potential. And it's not always intuitively obvious which traits are virtues, and to what extent. So we have to put effort into getting this right.

The definition I'm using here is not gospel. Different people define virtues differently depending on what work they hope to accomplish with their definition. For example, while I was composing this post, I saw a series of tweets from philosopher @AgnesCallard in which she contrasted rationality as a *virtue* with rationality as a *skill*.^[3] If I understand her right, she's saying that someone who tries to be rational because that's what the situation calls for, when there are no strong temptations against rationality, is merely being more-or-less skillful at rationality. To be *virtuous* at rationality, on the other hand, means to exhibit rationality when it is costly to do so, or when there are strong temptations to do otherwise. By contrast, in the system I'm using, rationality can be a virtue if it is a characteristic trait, whether or not that trait is presently being exercised in trying circumstances (though if rationality deserts you in such circumstances, this suggests that your character trait is not well-established and your virtue is perhaps shallow or weak).

What is rationality?

“[A]n aim of philosophy is patiently and unremittingly to sustain the vigilance of reason in the presence of failure and in the presence of that which seems alien to it.” —Karl Jaspers^[4]

Rationality is also a word that people define differently in different contexts. On LessWrong I often see “instrumental rationality” (making effective and efficient decisions when pursuing goals) and “epistemic rationality” (having processes that reliably lead you to adopt more accurate beliefs) joined under the rationality banner. I've already covered some aspects of instrumental rationality in my post on the virtue of [prudence / practical wisdom](#), so I will focus more on epistemic rationality here.

Epistemic rationality remains part of instrumental rationality (it is difficult to make good decisions if you begin with bad assumptions or faulty data). But it is also something else. Some assert (to be less coy, *I assert*) that to have beliefs that more accurately represent reality is itself valuable, [even if it has no instrumental value](#) beyond that. This makes epistemic rationality not just a means to an end, but the means to one of *the* ends.

[The LessWrong summary of rationality](#) notes that:

...rationality is both a science and an art. There's study of the iron-clad laws of reasoning and mechanics of the human mind, but there's also the general training to be the kind of person who reasons well.

“To be the kind of person who reasons well” is another way of saying “to have the virtue of rationality.” Aristotle, in his *Nicomachean Ethics* (the foundational work of the

virtue ethics tradition), wrote that the ultimate goal of ethical philosophy ought not to be to understand theoretically what goodness essentially is, but to understand practically how we are to become good people.^[5] So the virtue ethics approach may be useful to us as we consider how “to be the kind of person who reasons well” on top of our more theoretical understanding of good and bad reasoning patterns in the abstract.

Rationality vs. Rationalization

“[T]he majority of men do not think in order to know the truth, but in order to assure themselves that the life which they lead, and which is agreeable and habitual to them, is the one which coincides with the truth.” —Tolstoy^[6]

“So convenient a thing to be a *reasonable creature*, since it enables one to find or make a reason for every thing one has a mind to do.” —Benjamin Franklin^[7]

One problem with trying to restrict yourself to instrumental rationality is that some irrational antipatterns are hard to avoid without epistemic rationality as a back-up. For example, if you are being rational only in order to meet, and only to the extent that you meet certain instrumental goals, you may find that you can efficiently cheat by being less-than-rational in how you evaluate whether those goals have been met.

If you do not hold tight to epistemic rationality, you may lose your grip on instrumental rationality as well. By what standard will you judge whether your instrumentally rational decisions are actually rational ones? If you don't actually love truth and reason themselves enough that you would spit out any counterfeits with disgust, it's very easy to convince yourself that whatever way you arrive at attractive conclusions is also the rational way.

Rationality and Rationalities

Assuming you decide you want to be rational, what sort of rational will you be? Philosophers have noted that there are competing systems of rationality that are internally consistent but incompatible with each other.^[8] Each system is more rational than the others in the terms of rationality it itself considers valid. No universal system of rationality exists by means of which we can objectively adjudicate between them.

The solution to this problem that seems most promising to me is to see reality itself as the final adjudicator. This is the approach favored by for instance positivists, scientists, and pragmatists. [Make your beliefs pay rent \(in anticipated experiences\)](#). and then contrast your beliefs with your experiences.

This approach has some drawbacks, however: For one thing, interpreting reality to compare it with a system is a technique the details of which are sometimes specific to the system that does the interpreting, so competing systems may have competing ways of deciding whether revealed reality conforms to past predictions.^[9] For another, this approach suggests that you can only be certainly rational about things that can be actually revealed and to the extent that they become actually revealed, which limits the application of rationality more than we might like (how are we to approach questions about other sorts of things? are there no more or less rational ways to do

so?). For another, experiences can be deceiving: they may happen to conform to irrational expectations for unsuspected reasons.

Problems with the rationality-as-a-virtue approach

There are a few possible snags with considering rationality as a virtue.

Is rationality a virtue?

“Rationality” is complex enough that maybe it ought to be considered not as a virtue but as an umbrella term covering several virtues. For example, Eliezer Yudkowsky considered [twelve virtues of rationality](#): “curiosity,^[10] relinquishment, lightness, evenness, argument, empiricism, [simplicity](#), [humility](#), perfectionism, precision, scholarship, and the void.”

Aristotle listed his own set of intellectual virtues: art (knowing how to manipulate the world into desired forms), science (deriving conclusions through reliable methods), [wisdom](#) (choosing the right means for wise ends), philosophy (wrestling with the big questions competently), and intuition (knowing sensible first principles to bootstrap from).

Some others that you arguably could add to the package include: imagination, creativity, foresight, inventiveness, originality, resourcefulness, adaptability, inquisitiveness, open-mindedness, philomathy, skepticism, [attention](#), [awareness](#), [mindfulness](#) / [presence](#), [focus](#), [observation](#), [heedfulness](#), [vigilance](#), [discernment](#), sensitivity, wonder, [reverence](#), [faith](#), awe, elevation, taste (aesthetic appreciation), [know how](#) / [practical knowledge](#) / [craft](#) / [skillfulness](#), [practical wisdom](#) / [decision theory](#), intelligence, factual knowledge, devotion to the truth / good faith reasoning / careful evaluation of evidence, seeking out good advice, [judgment](#), ethics, perspective, righteousness, insight, emotional intelligence, self awareness, and intellectual autonomy / independent thinking.

If rationality is complex in this way, it might not be helpful to consider it as a virtue. It may be that each of its facets requires its own attention and can be best developed in its own specific way. (Another option would be to narrow “rationality” so that it means more specifically *epistemic* rationality, and maybe yet more narrowly the correct use of deductive and inductive reasoning and related skills.)

Is rationality a virtue?

Is rationality really a keystone of human flourishing, or is it just the eccentric enthusiasm of the rationalists and Aristotle-heads I hang out with on-line? (And how would I know if I had the correct answer to that question?)

Occasionally you will see the argument that rationality itself is not a component of a flourishing human life but indeed can interfere with human flourishing. If true, you should only be rational to the extent that rationality helps you meet other ends of human flourishing, and then stop. If it costs more effort to learn truth X (or rational

method X) than the benefit you receive from knowing X rather than some placeholder irrationality Ψ , stick with Ψ . Rationality and truth themselves are, in this telling, orthogonal to human flourishing, and it's a mistake to confuse one with the other.

The [steadfast](#) pursuit of truth and reason [comes with no guarantee of leading to a better life](#) unless it turns out that the steadfast pursuit of truth and reason is itself part of a better life. In other words: If rationality is not a virtue, [it might turn out to be a poor use of your time](#).

Here is one way this argument can play out: you aren't here to be rational but to *live*. Living includes being irrational, perhaps even flamboyantly, egregiously so. Life is not a problem to be solved by reducing it to its lowest common denominator, but a drama that requires from you a whole-hearted immersion and [a necessarily irrational suspension of disbelief](#). Being devoted to non-instrumental rationality is like being so devoted to literary criticism that you can no longer enjoy a story, or like spending your time on a roller coaster carefully examining each of the rail welds as they go by rather than enjoying the thrill of the ride.

On the other hand, even a potato can *live*. To reason, on the other hand, is something it seems you have to be a person to really appreciate. Reasoning is not something we do in-between our moments of living, or in contrast to them, but is part of *how* we live as humans. In the virtue ethics approach, it also comes into play in all of the other virtues that represent human excellences: if a virtue is a variety of characteristic *choice*, reason is key to the careful discrimination with which we make those choices and is thereby an ingredient in most if not all virtues.

That doesn't exactly contradict the objection that reason need be carried on only so far as it has practical results. But it suggests that what is practical about reason may cover a broader range of human life than it might at first appear.

Is human flourishing a coherent concept?

One of the biggest vulnerabilities of the virtue ethics approach is its appeal to human thriving or flourishing — to the pursuit of *eudaimonia* — as the basis of ethics:

Student: How do we know what human flourishing is?

Professor: <desperate handwaving> Eh? [You know it when you see it.](#)

Will Wilkinson argued that the idea of human flourishing as a criterion for ethics hopelessly fails: "there is no non-stupid *natural* fact of the matter about what it would mean for you to realize or fulfill your potential, or to function most excellently as the kind of thing you are."^[11] To the extent that humans can be said to have a natural, essential *telos*, and to more-or-less flourish when measured against it, this *telos* has seemingly turned out to be a disappointingly inane and pointless running in circles: to preserve and propagate our genes so that the next generation might do the same, *ad nauseum*.

I don't think this is as fatal a flaw as Wilkinson does. For one thing, it is notoriously difficult to defend a good bedrock foundation for ethics in *any* system, so it would be no great embarrassment to find that virtue ethics suffers from this too. Granted that "human flourishing" is indeed a little handwavey, is it really any worse than the alternatives?

Also, in spite of Wilkinson's objection, people seem to be comfortable making at least some confident judgements about human flourishing. We call things like blindness, deafness, aphasia, paralysis, etc. "disabilities," "handicaps," "afflictions," or what-have-you, because we have a common-sense idea of human flourishing that includes things like sight, hearing, language, locomotion, etc. Virtue ethics asks us to use this same intuition to consider courage, industriousness, patience, rationality, and so forth — in other words, to see our "organs of character" as more or less capable, more or less healthy, more or less conducive to our success at humaning exceptionally well.

For all its fuzziness, one of the advantages of "human flourishing" as a criterion — when compared with other popular criteria used in ethical theories, such as "[reducing] suffering" or "happiness" — is that it is more resistant to certain strange dead ends like [wireheading](#), [the repugnant conclusion](#), [the experience machine](#), etc.

Among the disadvantages: Human flourishing is difficult to define precisely, and it's hard to aim at a target you can't precisely locate. How are we supposed to distinguish what makes us flourish from what is fashionable, what is *de rigueur* for our class, what is dogmatically insisted upon by our culture, what habits we unthinkingly picked up as kids, what we have become inadvertently dependent upon, what helps us flourish locally but prevents us from reaching a zone of maximal flourishing, etc.? It's easier to imagine being incorrect about your flourishing than about whether you are "suffering" or "happy".

Although virtue ethics scholars love to wring their hands worriedly about objections like these, the core of virtue ethics remains mostly easy to swallow. In short, if you believe ① a human life can be a better or worse one to live, ② some significant part of what determines the quality of a human life is the choices that human makes, ③ the better choices are not wholly arbitrary, but have regularities such that choices of-certain-sorts more reliably characterize better lives, and ④ choices of-certain-sorts can become learned habits through deliberate effort, then you implicitly believe in some sort of virtue ethics.

How can you develop the virtue of rationality?

When Eliezer Yudkowsky was assembling his sequences of essays about rationality, he would occasionally pause to wonder whether we could be doing something more deliberate and methodical (or at least more effective) to promote rational thinking:

- "[Why aren't there dojos that teach rationality?](#)"
- "[We practice our skills, we do, in the ad-hoc ways we taught ourselves; but that practice probably doesn't compare to the training regimen an Olympic runner goes through, or maybe even an ordinary professional tennis player. And the root of this problem, I do suspect, is that we haven't really gotten together and systematized our skills.](#)"
- "[\[T\]here ought to be some discipline of cognition, some art of thinking, the studying of which would make its students visibly more competent, more formidable](#)"

In response to longings like these, the [Center for Applied Rationality](#) formed and [evolved an approach to training in rationality](#). You can find [a lot of discussion about the Center's theory & practice](#) here on *LessWrong*, and you can [look over the group's](#)

[handbook](#) if you dare. But if you want to attend one of their workshops and get personally-guided, hands-on direction... you may be out of luck. So far as I can see, the Center went into hibernation during the covid pandemic and hasn't yet recovered.

LessWrong includes an impressive catalog of methods to improve your epistemic rationality. (For this reason, I limit myself to a thumbnail sketch here.) The trick is to put these methods into practice in such a way that they shape how you *characteristically* come to conclusions. This may mean that you must behave in ways that seem instrumentally suboptimal in the short term because you have in mind not just whatever immediate application of rationality you are engaged in, but the long-term goal of becoming a more rational person through practice.

It is difficult to be textbook-rational in real time, about things whose domains are unclearly bounded, while using squishy hardware. Alas, this describes most of our questions for which rationality would be helpful. Alan Watts put it this way:[\[12\]](#)

If we were rigorously “scientific” in collecting information for our decisions, it would take us so long to collect the data that the time for action would have passed long before the work had been completed. So how do we know when we have enough? Does the information itself tell us that it is enough? On the contrary, we go through the motions of gathering the necessary information in a rational way, and then, just because of a hunch, or because we are tired of thinking, or because the time has come to decide, we act....

In other words, the “rigorously scientific” method of predicting the future can be applied only in special cases — where prompt action is not urgent, where the factors involved are largely mechanical, or in circumstances so restricted as to be trivial. By far the greater part of our important decisions depend upon “hunch” — in other words, upon the “peripheral vision” of the mind. Thus the reliability of our decisions rests ultimately upon our ability to “feel” the situation, upon the degree to which this “peripheral vision” has been developed.

So part of the trick to becoming characteristically rational is to shape our “peripheral vision of the mind” so that it more closely approximates the results we would (if we could) arrive at through more rigorously rational methods.

Here is one way we can do this: Our peripheral vision of the mind is distorted by a variety of biases. As we learn about these biases we can craft a set of “corrective lenses of the mind.” For example, if we know that we share a common human bias to selectively seek evidence and arguments that would confirm a theory we favor or refute a theory we dislike, we can apply corrective lenses by also [earnestly seeking inconvenient evidence and arguments](#). If we *habitually, regularly* wear corrective lenses like these, we sharpen our peripheral vision, and so strengthen our virtue of rationality.

Another way is to recalibrate the heuristics of our peripheral vision of the mind by contrasting their results with either [painstakingly-reasoned or -researched answers](#) (when time allows) or with [eventually-revealed reality](#) (in the case of predictions).

1. ^

Thomas Macaulay “Lord Bacon” (1837)

2. ^

Analects of Confucius, VI.xviii

3. ^

Agnes Callard ["Thread on rationality"](#) Twitter 2 March 2021. Lightly detwitted:

I've been thinking about a conversation with [@TheZvi](#) about whether rationality is a skill or a virtue. I'm starting to think it depends on one's "position" in relation to the question. Imagine two positions you might occupy at a criminal trial:

1 juror

2 defendant's mother

Supposing the juror starts with few assumptions about the case, they will find their mental states (roughly) tracking the evidence. For the mother, rationality is more expensive, because it comes at the cost of psychological pain (acknowledging the possible guilt of her child).

The juror may come to an irrational decision due to failures in cognitive processing—these would be signs of a lack of rationality as skill—but (unless there is some way that the case is personal for him) rationality as virtue is not so much on the table for him.

The mother, by contrast, has an opportunity to showcase extreme—one might even call it heroic—rationality in the virtue sense. (This is somehow similar to Socrates saying in the *Laches* that people who know how to dive into wells—experts—do not count as courageous for doing so)

If correct, this argument shows that you could be very skilled in rationality while lacking the virtue, and thus "when push comes to shove"—when there's a psychological cost—you'll find it just as hard to be rational as anyone else; your rationality could fail you when you need it most.

4. ^

Karl Jaspers, *Way to Wisdom* (1950)

5. ^

Aristotle, [Nicomachean Ethics](#), book II, chapter 2

6. ^

Leo Tolstoy, [The Kingdom of God Is Within You](#) (1894) chapter 6

7. ^

Benjamin Franklin, *Autobiography* (1791)

8. ^

And I'm not just talking about the usual postmodernist suspects. See e.g. Robert Nozick, [The Nature of Rationality](#), (1993); Alasdair MacIntyre, [Whose Justice?](#)

Which Rationality?, (1988).

9. ^

Thomas S. Kuhn, The Structure of Scientific Revolutions (1962)

10. ^

Curiously, he suggests that curiosity is the fundamental virtue, from which epistemic rationality is derived. And furthermore, that curiosity is an emotion.

11. ^

Will Wilkinson “Eudaimonism is False” BigThink (7 February 2012)

12. ^

Alan Watts, The Way of Zen (1957)

Notes on Amiability

This post examines the virtue of **amiability** (and closely-related or synonymous virtues like friendliness, geniality, agreeableness, conviviality, affability, niceness, affection, and warmth). It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations.

I wrote this not as an expert on the topic, but as someone who wants to learn more about it and to become better at it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What are these virtues?

These virtues have to do with being pleasant to be around in casual social settings. If you exhibit these virtues, people feel at ease either initiating interactions with you, engaging with you, or simply being around you. You signal that you have benign and respectful intent, in a way that is legible to those around you (you are not a grouch, abrasive or obnoxious, or socially awkward in a way that is off-putting or hard to negotiate). You harmonize well with your social environment (you are not contentious or a shit-stirrer). You welcome friendly interactions with others (you are not stand-offish, cold, brusque). You tone down or repress any inclinations to ratchet up social tensions (you are not ill-natured, querulous, snappy, abrasive, hostile, disputatious).

If you go overboard, being insincere or over-the-top in the way you try to butter up those around you, you might be accused of being a flatterer or being fawning, obsequious, unctuous, oleaginous. The old-fashioned term “man-pleaser” is sometimes deployed in this context. Someone who is so friendly that you’re sure they’re about to pitch you Amway or Krishna Consciousness puts you on edge rather than at ease with their amiability.

Many of the social virtues can play a role in assisting amiability. Some related virtues include hospitality, graciousness, connection, goodwill, [courtesy](#), [kindness](#), sympathy, gentleness/tenderness, tolerance, tact, civility, cheer, warmheartedness, sympathy, concern, and consideration. Friendliness is distinct from friendship, as the latter involves the skills of properly maintaining a more durable, less-superficial relationship, while the former concerns how you interact with people in general, including strangers and casual acquaintances. That said, affection and warmth are also important ingredients of more intimate friendships, romantic relationships, and family relationships.

Affection and touch

When I looked around for resources about “affection” in particular, I mostly found resources about affection in the context of romantic (or sometimes parental) relationships — particularly when it comes to how to deploy physical affection / [touch](#) in a graceful and welcome way. But I think that’s just a specific case of a more general virtue. I’ve had times when someone has placed their hand on my arm and looked me in the eye in a gently encouraging way that was very effectively affectionate without being either romantic or parentalish. I can’t put my finger on what qualities exactly made it work where in other contexts it might have been awkward or counterproductive. Touch is difficult: it can be a good way of expressing affection/warmth, but it can also be misinterpreted as a romantic overture or a predatory gambit; sometimes it is also seen as condescending.

When going along to get along is a bad strategy

At first, amiability looks like a sort of common-sense “things I learned in kindergarten” sort of virtue. But it has a common and challenging element attached to it: An example of a situation in which we struggle to find the Golden Mean of this virtue would be one in which we are in a group of casual acquaintances and one of them tells a joke that depends for its humor on the shared assumption of an offensive racial stereotype. Do we laugh in order to be agreeable and just try to move on, or do we signal our disapproval? When does our obligation to be agreeable and tolerant get eclipsed by our obligation to insist on better standards of behavior or our [disgrace](#) at being associated with shameful behavior? “Go along to get along” is a real problem, and it comes from being inattentive to the balancing act this virtue requires.

Other ways niceness can go awry

“The true gentleman is friendly but not familiar; the inferior man is familiar but not friendly.” —Confucius^[1]

Some people prioritize niceness at the expense of other virtues. Niceness can be cloying if it seems forced or insincere (or overwhelming or presumptuous).

If you presume more intimacy than you have earned — by sharing or demanding personal information, or by assuming you have permission to touch someone affectionately for instance — you may be overstepping your bounds in a way that comes across as more threatening than friendly, whatever your intentions.

Charming sociopaths

Geniality can be a thin social virtue. Sociopaths are sometimes very charming, but also very self-serving: buttering you up to see what they can get out of you. (But do sociopaths perhaps get an unfairly bad name: tarred by the brush of the more sadistic among them? After all, when it comes down to it, [we love sociopaths](#).)

The virtue of being disagreeable

Is there a virtue to be found in being disagreeable? Maybe there is a case to be made for the virtue of being a cantankerous grouch instead. Different people shine in different contexts and in different ways.



Spencer Greenberg @SpencrGreenberg · 11h

As a high agreeableness person (~77th percentile) I want to give a shout-out to those low agreeableness folks who do valuable things it's harder for others to do, especially 1. publicly fighting back against bullies, 2. publicly resisting bad (yet very popular) ideas. Thanks! :)

How to improve at amiability

With most virtues, the key to getting better is to practice. You start off more-or-less clumsy, then you put in effort at the margin of your current ability, and over time you become more capable. With social virtues, the early, clumsy stage of this process can be embarrassing. You have to put yourself out in front of other people, deliberately doing things beyond your current skill level.

If you find social embarrassment intolerably uncomfortable or frightening, you will have difficulty with this. You somehow need to be able to say “I’m definitely going to screw up from time to time because I’m pushing myself beyond my current comfort zone, but *that’s okay* — I’ll just brush that off and move on, because I know that’s what it takes.” Easier said than done, I know. Maybe some preliminary work on the virtue of [courage](#) would help.

Different people have different sorts of deficits in amiability, with different roots. Some people want to be agreeable and just don’t have a good idea what kind of vibe they’re putting out (e.g. the awkward). Other people developed disagreeableness as a strategy for keeping people at bay (e.g. the gruff). Other people like drama and find other people more interesting when they’re uncomfortable (e.g. the shit-stirrer). Others are unfriendly because they think they’ve got more important things on their agenda than being pleasant (e.g. the jerk). With such variety (and this is just off the top of my head), there will probably also be a variety of strategies to pursue in the course of becoming more agreeable. It may require a lot of work just to identify what’s causing your deficits in the first place before you start working on them.

The difficulty of getting reliable feedback

With amiability there is an additional challenge: it can be difficult to get good feedback on how well you are doing. Let’s say you find yourself sitting at a bus stop with some random stranger, and you think to yourself — “aha! I have an opportunity to practice my amiability.” You notice they are wearing an unusually interesting sweater, and decide to compliment them on it as an opening conversational gambit.

Imagine that you do this in a fabulously competent, suave, utterly disarming way, and the stranger replies by grunting, looking down at their shoes, and inching away from you on the bench. Maybe they’re having a bad day, or they aren’t very good at friendliness themselves, or they’re hard of hearing and are embarrassed to confess they didn’t understand what you said. Any number of things might have happened, but your feedback is: “boy howdy, that sure flopped.”

Or on the other hand, maybe you compliment them on their sweater but do so in an incompetent way that makes you sound like you’re being sarcastic, or are making an inappropriate sexual overture, or something like that. But they overlook that and smile and tell you the story of how they got the sweater and then ask you about your jacket, and you hit it off grandly. Maybe they’re just especially fond of conversation, or they’re charitable about the foibles of the people around them, or maybe *they* misheard you. You may never know. But your feedback now is: “nailed it!”

It may take a lot of data before reliable patterns show up. If you have friends you [trust](#) to be [frank](#) with you, you can ask them for feedback on how you’re doing and how you might improve.

Become a brilliant conversationalist by letting them talk

I was lucky enough to have [a good friend who was extraordinarily good at this virtue](#). And the way he described it, it was definitely an acquired skill and not something that just came naturally to him. So that (and my own experience at just becoming more middlingly competent) makes me more confident in saying that this virtue is something that is learnable and can be improved with practice.

Greg, my friend, had an incredible knack for turning a stranger into a friend in minutes. I tried to study and learn from his techniques, but I think I was only perceptive enough to pick

up some of the rudimentary stuff.

One thing I noticed was that he was very skillful at quickly turning the conversation to whatever it was that the other person was most interested in talking about. Just about everybody has some thing or things that they're passionate about. Sometimes they're a little reluctant to start, though, because they don't want to get typecast or to come off as a monomaniac. But Greg would somehow manage to steer the conversation until it became about the other person's favorite thing, and then would be full of questions. Before long, the other person was loquacious, comfortable, and fully convinced that Greg was a man of excellent taste whom he or she was lucky to have met. Meanwhile Greg was learning all about some new niche subject directly from an expert, while also making a new friend.

I don't have anywhere near the knack for this that Greg did, but I've tried to learn from his technique. Now I tend to spend most of my casual conversations with people asking them questions about things they have already expressed enthusiasm about. I learn a lot that way, and I think I come across much better in conversation than when I used to spend most of my half of the conversation saying things I thought were interesting or important or impressive.

1. ^

Analects of Confucius, XIII.xxiii

Notes on Simplicity

This post examines the virtue of **simplicity**. It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is simplicity?

Simplicity is often recommended, but it means different things to different people. That shouldn't be too surprising, as you can be more simple or more complex on many different dimensions. The puzzle, then, from the point of view of simplicity-as-a-virtue, is to discover whether there is some underlying benefit to all of these different forms of simplicity, or whether each one needs to be appreciated (or not) on its own merits.

- Simple can mean unadorned, as in a simple tan T-shirt vs. a blouse with ruffles, embroidery, and a glittery appliqué.
- Simple can mean uncomplicated, not having a lot of components; in this sense your life is simple perhaps if your motives and needs are few.
- Simple can mean comprehensible, legible, easy to understand; in this sense someone may be more simple if their life is an open book and they wear their heart on their sleeve.
- Simple can mean easy to accomplish; a simple life in this sense may be one that sticks to the basics and does not have extraordinary [ambitions](#).
- Simple can mean unburdensome, as in "I simplified my life by replacing the lawn I had to mow every week with a shuffleboard court."
- Simple can mean direct and to-the-point, without unnecessary detours; in this sense simplicity may be a type of [straightforwardness](#).
- Simple can mean unsophisticated, as in someone who has simple tastes.
- Simple can mean unextravagant, down-to-earth, unostentatious ("I'm just a simple farmer").
- Simple can mean "reduced to the essentials"; simplicity in this sense might mean finding out what really matters and maintaining your focus there.

Simplicity as an aesthetic

Simplicity is also a popular aesthetic. The google.com front page with its logo, one-field form, and two buttons, centered on a plain white background, is one example. The consistent Steve Jobs / Mark Zuckerberg uniform of black (or grey) shirt, blue jeans, and sneakers is another.

The household interiors in [Real Simple](#) magazine prioritize function over decoration, and feature subdued colors, absence of clutter, natural materials, and few electronic gadgets and logos. Simplicity here seems to imply calm, reduced distraction, and insulation from the blinking beeping jarring noise of technology and commerce. The LessWrong site, with its bountiful whitespace, subdued color scheme, and absence of

distracting pop-ups, ads, and animated doo-dads, is an on-line example of a similar aesthetic.

The [Marie Kondo](#) phenomenon reabsorbed the aesthetics of simplicity into a more comprehensive “life changing” program that promises to “bring joy into your life” through tidiness, order, and simplicity.

On the other hand, the simple aesthetic can look ridiculous when pushed to extremes: when tidiness has less to do with creating an orderly space for life’s work, and more to do with creating a sterile display that no work is allowed to pollute. Epictetus counseled: “If you have chosen a simple life, don’t make a show of it. If you want to practice simplicity, do so quietly and for yourself, not for others.”^[1]

Simple as a pejorative

“Simple” is sometimes also deployed pejoratively. Calling someone simple can be a way of saying they’re not too bright. Someone who is simplistic sees the world in black-and-white, or forces complex matters into simple categories by recklessly discarding nuance. You don’t want to “over-simplify matters.” This suggests that simplicity is another of those “golden mean”-style virtues of the Goldilocksian too-little, too-much, just-right variety.

Some related virtues

Some virtues that are in the same ballpark as simplicity include [moderation](#) (balance, not taking things to extremes), [temperance](#) (having well-regulated desires), efficiency (avoiding wasted effort), and orderliness (being uncluttered, everything in its place). Coping with complexity well — prioritizing, maintaining perspective, transiting gracefully between the big picture and the gory details, being able to switch focus rapidly without getting flustered — is also important, and maybe we should beware of embracing simplicity if we might be doing so in order to avoid working on skills like those.

Simplicity and [rationality](#)

Eliezer Yudkowsky made simplicity the seventh of his [twelve virtues of rationality](#), and described it this way:

Antoine de Saint-Exupéry said: “Perfection is achieved not when there is nothing left to add, but when there is nothing left to take away.” Simplicity is virtuous in belief, design, planning, and justification. When you profess a huge belief with many details, each additional detail is another chance for the belief to be wrong. Each specification adds to your burden; if you can lighten your burden you must do so. There is no straw that lacks the power to break your back. Of artifacts it is said: The most reliable gear is the one that is designed out of the machine. Of plans: A tangled web breaks. A chain of a thousand links will arrive at a correct conclusion if every step is correct, but if one step is wrong it may carry you anywhere. In mathematics a mountain of good deeds cannot atone for a single sin. Therefore, be careful on every step.

[Occam's Razor](#) is one classic example of how simplicity can come to the aid of rationality.

The story of the progress of science is often told as a series of simplifications and consolidations, as when the various complex methods of predicting the mysterious motions of the heavenly bodies were subsumed under a single explanation that also explained the motion of more mundane bodies close at hand.

Simplicity and prioritization

“Our life is frittered away by detail. Simplify, simplify, simplify! I say, let your affairs be as two or three, and not a hundred or a thousand; instead of a million count half a dozen, and keep your accounts on your thumb-nail.” —Thoreau^[2]

Simplicity is one antidote for distraction. If you are pulled this way and that by a thousand unimportant demands, you may lack the focus you need to make any headway in more crucial areas.

Thoreau describes his experiment in *Walden* in this way: “I went to the woods because I wished to live deliberately, to front only the essential facts of life, and see if I could not learn what it had to teach, and not, when I came to die, discover that I had not lived.”^[2]

Simplicity in this sense is a sort of reassessment and retrenching: What’s in that sack I’m lugging around with me, and what do I fill my days with? Is some of that just accumulated junk I’d be better off without? What’s the essence of what I’m about and what’s superfluous? If I take that rubbish out to the curb, can I make more room in my life for more important things?

Marie Kondo invented a charming ritual to accompany this process as it applies to things we own. As part of the decluttering process of getting rid of an item that no longer “sparks joy,” she recommends thanking the item for whatever service it gave: “People may feel guilty when letting go of items. By expressing gratitude toward the items you let go, it will lessen the feeling of guilt.” I can imagine something like this also making it easier to move on from non-material things, like hobbies or causes or social media accounts, that are offering diminishing returns. This way you don’t feel you have to harshly repudiate who you were in order to become who you’d like to be.

Simplicity and the rat race

“From time to time I meet people who live among riches I cannot even imagine. I still have to make an effort to realize that others can feel envious of such wealth. A long time ago, I once lived a whole week luxuriating in all the goods of this world: we slept without a roof, on a beach, I lived on fruit, and spent half my days alone in the water. I learned something then that has always made me react to the signs of comfort or of a well-appointed house with irony, impatience, and sometimes anger. Although I live without worrying about tomorrow now, and therefore count myself among the privileged, I don’t know how to own things. What I do have, which always comes to me without my asking for it, I can’t seem to keep. Less from extravagance, I think, than from another kind of parsimony: I

cling like a miser to the freedom that disappears as soon as there is an excess of things.” —Albert Camus^[3]

Advocates of simplicity often point out the advantages of being *unencumbered* by many material things. Simplicity in this sense is a prerequisite for freedom. The more needs you have, the more things you are the caretaker of, the larger your footprint, the heavier your knapsack, the more restrained are your choices and the more limited your range. Our possessions confine us: “Chains of gold are stronger than chains of iron.”

Conspicuous consumption and social aspiration often take the blame for these encumbrances. We buy a clever new time-saving kitchen appliance, and forget that we’ll still have to clean it, find a place on our crowded counters for it, and try to remember where we put its warranty when it breaks. We try to keep up with the Joneses with a house as big as theirs, a car as nice as theirs, a vacation no less exotic and adventurous, and so forth, and what we get in return are obligations that bind us to the treadmill.

Living simply usually means living less expensively. This can help you have more options in your professional life: you are freer to choose a more intrinsically rewarding job even if it pays less, and you can squirrel away that treasured “fuck you money” that allows you to walk away from an unethical or demeaning job rather than compromise to pay the bills.

Simplicity and contentment

The more complicated and sophisticated our needs, the more difficulty we have in meeting them, and the more dissatisfaction we can expect. Are you sure that’s a price worth paying for refined and fashionable tastes? “The luxurious receive no greater pleasure from their dainties than the peasant does from his bread and cheese,” wrote William Paley, “but the peasant whenever he goes abroad finds a feast, whereas the epicure must be well entertained to escape disgust.”^[4]

That slur against the epicure aside, [Epicurus](#) himself mostly agreed with Paley about this:

We believe that self-sufficiency is a great good, not in order that we might make do with few things under all circumstances, but so that if we do not have a lot we can make do with few, being genuinely convinced that those who least need extravagance enjoy it most; and that everything natural is easy to obtain and whatever is groundless is hard to obtain; and that simple flavors provide a pleasure equal to that of an extravagant lifestyle when all pain from want is removed, and barley cakes and water provide the highest pleasure when someone in want takes them. Therefore, becoming accustomed to simple, not extravagant, ways of life makes one completely healthy, makes man unhesitant in the face of life’s necessary duties, puts us in a better condition for the times of extravagance which occasionally come along, and makes us fearless in the face of chance.^[5]

Sometimes people will also suggest that extravagance on your part probably means deprivation for someone else. “Live simply that others may simply live,” reads the bumper sticker. Nowadays, simplicity is often measured in part by one’s carbon footprint or, more generally, by how much non-renewable resources one consumes:

"How many planets would we need to satisfy the demand if everyone used as many natural resources as me?" Simplicity in this sense sometimes finds itself under a conservationist umbrella with car-free living, vegetarianism, reduce/reuse/recycle, going off-the-grid, and things of that sort.

Simplicity as a virtue

To express simplicity as a virtue is to understand the value of simplicity and to incorporate ways of achieving it into your life.

Complexity itself has costs. It makes life harder to manage, reduces our degrees of freedom, and so forth. Often people do not factor those costs into their decisions as they incrementally and inattentively complexify their lives. A person with the virtue of simplicity asks, of any decision they make, "does this make my life more complex, and if so is that worth it?"

If you value simplicity you will also reassess your current possessions and pastimes using simplicity as a metric, while on guard against things like the [sunk cost fallacy](#) and [status quo bias](#) that might tempt you to keep that spoiled milk well past its expiration date.

1. ^

Epictetus, *Enchiridion*

2. ^

Henry David Thoreau, *Walden* (1854) chapter 2

3. ^

Albert Camus, "The Wrong Side and the Right Side (Preface)" (1937)

4. ^

William Paley, *The Principles of Moral and Political Philosophy* (1785)

5. ^

Epicurus, *Letter to Menoeceus*

Notes on Forgiveness

This post examines the virtues of **forgiveness**, **mercy**, **clemency**, and **epieikeia**. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

Forgiveness in particular has become kind of Oprahfied as a self-help cure-all in recent years — it is not just good for your soul but “can reap huge rewards for your health, lowering the risk of heart attack; improving cholesterol levels and sleep; and reducing pain, blood pressure, and levels of anxiety, depression and stress.”^[1] This trend has gone so far that a backlash began and now there is also a self-help cure-all centered on cultivating and nurturing your grudges.

Different people mean very different things by “forgiveness” so for that virtue I want to spend some time exploring why that is, what the different facets of forgiveness are, and how some people emphasize some over others.

What are these virtues?

“Mistrust all in whom the impulse to punish is powerful.” —Nietzsche^[2]

You exercise **mercy** when you have the power and inclination to harm or punish someone else, and you refrain from doing so. Sometimes mercy is a way of tempering **justice** (I would be in my rights to have you arrested, but I won’t), but sometimes it has nothing to do with justice or can even temper injustice (kill all the men, but spare the women & children). Mercy isn’t necessarily even all that nice. For example, it might be called an act of mercy to warn people in a village to evacuate before you bombard it, even though by bombarding it you are harming it. Or, it might be called an act of mercy to give a condemned prisoner a last cigarette or to offer them a blindfold, even though you ultimately mean them harm.

Clemency typically applies only to judgement or punishment (not, in other words, to more arbitrary harm). You offer clemency to someone when you cut them slack over their wrongdoing: judging them less severely and inflicting less severe consequences, letting it slide. This term is used in a legal context to describe reduced legal punishments or the executive pardon power. Sometimes “leniency” and “indulgence” are also used in this sense. A pardon of many people, or of unknown/anonymous offenders is sometimes called an “amnesty.”

Epieikeia is something like clemency, but it can go in a more-severe direction as well as a less-severe one. It is a way of wisely adhering to the spirit of the law when the letter of the law doesn’t quite match the nuances of a particular case. Aristotle described *epieikeia* this way: “*Epieikeia* is to pardon human failings, and to look to the [intentions of the] lawgiver and not to the law; to the spirit and not to the letter; to the intention and not to the action; to the whole and not to the part; to the character of the actor in the long run and not in the present moment; to remember the good rather than evil, and good that one has received, rather than good that one has done; to

bear being injured; to wish to settle a matter by words rather than by deeds; lastly, to prefer arbitration to judgment, for the arbitrator sees what is equitable, but the judge only the law, and for this an arbitrator was first appointed, in order that *epieikeia* might flourish.”^[3]

Epieikeia respects justice. It looks at some naively-just outcome, notices that it is actually unjust when you look at the big picture, and corrects that outcome to be more thoroughly just. Clemency or pardon, on the other hand, can (and often do) operate by bypassing justice: They look at some just outcome and decide to evade it in the interests of mercy, or as a favor.

The virtue of **forgiveness** is much trickier to define, so I'll postpone that definition for now.

Related virtues and concepts

Some other virtues in the same ballpark include [good temper](#), flexibility, pity, placability, and reconciliation.

If you squint your eyes and look at “tolerance” just right, it can look like a sort of preemptive forgiveness. For example, if you tolerate people who are noisy about the way they eat their popcorn at the movie theater now, you don’t have to bother about forgiving them for it later.

The flip-side of these virtues includes what you can do if you are the wrongdoer, in order to make amends: things like accepting fault, atonement, apology, restitution, reform, [shame](#), remorse. I hope to cover some of those things in a later post.

There is some tension between these virtues and things like justice, impartiality, and righteous anger—in some views they work at cross-purposes; in others they are complementary. (*Epieikeia*, for example, is said to “crown” justice, although to people without a sense of *epieikeia*, it can seem to violate justice.)

The vices of deficiency associated with these virtues go by names like legalism, ruthlessness, grudge-holding, vengefulness, vindictiveness, harshness, mercilessness, implacability, and resentment. The vices of excess are sometimes called leniency, indulgence, being a pushover or schnook, letting people walk all over you (being a doormat), and things of that sort. If you are too apt to let your children evade judgement and punishment you will be said to be spoiling them.

Mercy/clemency and forgiveness are distinct from justification, excuse, condoning, or pardon, with which they are sometimes confused. When mercy/clemency are applied to the way justice is meted out to a wrongdoer, or a wrongdoer is forgiven, this does not imply that the wrongdoer’s action has been excused, justified, or condoned. The mercy/clemency/forgiveness is *in spite of* the blameworthiness of the wrongdoer, not a way of erasing it.

The “[restorative justice](#)” movement is attempting to formalize something like a forgiveness/atonement/reconciliation stage into the way society handles criminal justice.

In some interpretations of [satyagraha](#), this form of nonviolent resistance to injustice operates in part through a sort of preemptive forgiveness in which the *satyagrahi*

provokes injustice towards herself or himself, absorbs it, and forgives it as a method of trying to end it.

Taking a closer look at forgiveness

Forgiveness turns out to be difficult to define, and various philosophers and researchers have come at it from different angles and have come up with very different ideas of what it consists of and accomplishes. There is, for example, fundamental disagreement about whether forgiveness is a sort of transaction between people or an internal change-of-heart within a single person. Some people say it's one, some the other; some say it's primarily one but sometimes also includes the other; some say both things are important, and they often come together, but that they are distinct and it's unfortunate that we use the same word for both; some say forgiveness always includes both in tandem.

In this section, I'll describe some of the possibilities and open questions involved in defining forgiveness.

Forgiveness applies when you have been wronged by someone, and you resent that person having done that wrong. The forgiveness somehow acts on that resentment: removing it, tempering it, or changing its effects in some way. Forgiveness may (or in some versions *must*) also change the relationship between the victim and the wrongdoer: reconciling them to some extent. Forgiveness may (or must) also change what counts as a just outcome: for example, perhaps the wrongdoer no longer needs to make restitution or do penance if they have been forgiven. Forgiveness may (or must) also change the forgiver's attitude toward the wrongdoing (e.g. from intolerable to tolerated, from something that must be addressed to something that can be left alone, from something that demands reparations to something that does not).

Clemency, *epeikeia*, and mercy typically assume that the person practicing them has some power or authority over someone else that they can exercise in a more or less severe way. Forgiveness is different in that regard: it can be exercised by someone with or without such power or authority.

Usually forgiveness is applied to a specific act or series of acts ("I forgive you *for* this, that, and the other"). Every once in a while you see something that looks like blanket forgiveness ("I forgive you for being such an asshole"). On close inspection, some of those things look more like excusing than forgiving ("I see now that you are that way because you were raised by wolves").

Forgiveness is prototypically represented as a discrete action: the forgiver decides to forgive someone for something, and may mark that decision by making a declaration or [resolution](#) like "I forgive so-and-so for such-and-such." But forgiveness might also be the result of a more-or-less conscious long-term process, or even just the slow, gentle erosion of a grudge by the sands of time. Maybe you just realize one day that you have forgiven someone, without ever having formally decided to do so.

Occasionally you see forgiveness defined as merely the *moderation* of resentment into [reasonable](#) bounds. If, when you are wronged, you are tempted to go beyond the bounds of justice and to attempt for a more sadistic catharsis — e.g. pounding your hurt out of the hide of the wrongdoer — then that sort of vengefulness can be a vice, and "forgiveness" of a sort can be a way of avoiding falling into that vice.

Forgiveness might be considered a “pink” (gendered feminine) virtue in American culture. One forgiveness study I looked at noted in passing that it had used “[a]lternative framing” in order to recruit more male participants — calling its forgiveness training intervention “grudge-management training” in outreach material in the hopes that this would seem more inviting to men.^[4]

Forgiveness has been analyzed in a game-theory context. In a [prisoner's dilemma](#) game, for example, in which something like tit-for-tat is a good strategy, it can be easy for players to get locked into a repeating mutual-defect mode that is suboptimal for both parties. If somebody is able to reset the protocol by “forgiving” their partner/adversary, they can both benefit. On the other hand, if you forgive too readily or too often, you can end up losing to your adversary’s benefit.

What does forgiveness do?

Forgiveness usually does not mean that the forgiver changes their mind about whether they were wronged. The wrong still happened, and it was still wrong: it’s just a forgiven wrong now. If you do change your mind about whether you were wronged, what you’re doing is not forgiving so much as reassessing. Forgiveness also usually does not mean that the forgiver changes their mind about whether the wrongdoer was blameworthy. If you change your mind in that way, you’re probably excusing, condoning, or justifying rather than forgiving.

Forgiveness is often represented in such a way as to imply that it is a clean break with the way things were before: once you have forgiven, there’s no going back — the hatchet is buried for good. But there is difference of opinion about this, too. Maybe just as you changed your mind to forgive someone, you can change it back again to deforgive them. If you know you can always deforgive someone later if forgiving turns out to have been a bad idea, that might make it easier to practice forgiveness, which could be a good thing. On the other hand, maybe a retractable-forgiveness isn’t the real deal, and doesn’t provide the same benefits as a final-forgiveness would.

In some definitions of forgiveness, in order to forgive someone you must renounce certain things. You may have felt that your anger justifies hostility, unkindness, contempt, or scorn towards the wrongdoer, and maybe in order to forgive you have to lose that justification. You may have felt the need to take revenge, to adopt a stance of righteous indignation, or to demand satisfaction or recompense, and maybe in order to forgive you have to drop some or all of that. When you “forgive a debt,” for example, you resolve to leave it uncollected, and you no longer hold it against the debtor: you wipe it off the books. But there is a lot of difference of opinion as to what things (if anything) you necessarily renounce as a condition of forgiving someone. The [VIA Institute's page on forgiveness](#) says that “[i]t means to let go,” for example of “frustration, disappointment, resentment, or other painful feelings associated with an offense.” But do you have to (or should you) disavow your *feelings* or merely your sense that you are justified in acting on those feelings in certain ways?

Those who focus on forgiveness as an internal thing represent forgiveness as the successful culmination of the process of coping with being wronged: If you get interrupted or frustrated and cannot reach a state of forgiveness, in this telling, something has gone wrong, and the result will be that you churn uselessly in painful resentment of your own making. Those who focus on forgiveness as a transactional thing, on the other hand, are more apt to see forgiveness as something conditional that may (or ought) to be withheld — indefinitely if need be — from a wrongdoer who does not make an effort to earn it.

One possible way of looking at forgiveness is that it seals off a past wrongdoing and makes it fully-past.^[5] Until you forgive some wrongdoing, that wrongdoing continues to be a source of injury in the present. When you forgive it, it's a way of declaring that the harms done by that wrongdoing were done but are no longer *being* done. Because of this, you can still blame the wrongdoer for the harm they did, but you no longer have to adopt a defensive stance against the wrongdoer for the harm they continue to do.

People who focus on the internal aspect of forgiveness describe it as laying down a burden. If you don't forgive, you have unfinished business with the wrongdoer that nags at you. Forgiveness allows you to check it off the list and finally be done with it. You release your grip on the hatchet so you can bury it.

There are many sorts of negative feelings that may come from being wronged; the ones that form your resentment to the wrongdoer are only a subset of them. Forgiveness focuses on that subset, and sometimes is criticized for prioritizing those negative feelings that target the wrongdoer over the others. But one way people may keep their resentment toward a wrongdoer hot is by blowing on the coals of their other hurts; so if you give up on the resentment this may help you let the other hurts cool off as well.

Forgiveness is sometimes said to operate on *meaning*. It changes the frame or story that the forgiving person uses when describing the events of the past, so as to give those events a different interpretation or to refit them into the evolving context of the forgiving person's life. For instance, event X might go from being "X is why I do not trust Y" to "X is something I had to forgive in order to again trust Y." Forgiveness can be a variety of the alchemy that turns misfortunes in our past from calamities-that-afflicted-me into challenges-I-overcame and helps us envision ourselves as more empowered, active agents.

A possible stumbling block to forgiveness-as-meaning-transformation is if you believe that by admitting that forgiveness is the right thing to do now you are also admitting that resenting the injury was the wrong thing to do before now. It may be easier to forgive if you can tell a story in which you were right to feel hurt, right to feel resentful, but then *also* right to eventually forgive — you don't have to renounce your former feelings to adopt new ones.

Forgiveness can be an attempt by the victim to reconcile with the wrongdoer — to repair the rupture in the relationship that was caused by the wrongdoing or by the victim's reaction to it. But this isn't always the case. You might forgive a person who has died, or forgive a stranger you'll never meet again, for instance. But you might also (maybe?) forgive someone without any desire to be reconciled with them. Or you might forgive some things a person did without forgiving others, and so still be on the outs with them for that reason. Sometimes forgiveness implies that you will give the forgiven person another chance, sometimes not.

Another way forgiveness might operate on my relationship with the wrongdoer is just to change the priority of the wrongdoing in how I evaluate them. Instead of them *being* the Wrongdoer, they become somebody who, among other things, did me wrong. It demotes the wrongdoing from being the most salient fact about that person, and the overriding consideration in my evaluations of them, to just being another fact about that person. This can allow me to see them more clearly and in less of a caricature.

Just as you might forgive someone internally (by no longer dwelling on their wrong) without forgiving them externally/socially (by reconciling with them or by telling them they are forgiven), you might also do the reverse. This is sometimes called an insincere or “hollow” act of forgiveness, and is meant to reap the social benefits of forgiveness without having to renounce the grudge.

Who can be forgiven, and by whom?

“We must develop and maintain the capacity to forgive. He who is devoid of the power to forgive is devoid of the power to [love](#). There is some good in the worst of us and some evil in the best of us. When we discover this, we are less prone to hate our enemies.” —MLK, Jr.^[6]

It's almost always inappropriate, or inapplicable, to forgive someone for something blameworthy they did to hurt someone else; only the direct victim can do the forgiving. As a third party you can offer mercy or pardon — you might say “well, I won't hold that against you” but to say “I forgive you” is usurping authority you don't have. Vicarious, third-party forgiveness doesn't work except maybe in some unusual edge cases (or arguably in certain religious contexts).

People who emphasize the internal-transformation aspect of forgiveness note that you can forgive someone even if they have passed out of your life and you will never see them again. People who emphasize the transactional aspect of forgiveness have a harder time explaining such cases (they may suggest that this is only a quasi-forgiveness that might be better termed “letting go” or something like that).

Does it make sense to talk about forgiving an institution (“I don't know if I can forgive the Army for what they did to my boy”)?

Can you forgive yourself? Is there a sense in which you can be resentful of yourself? It seems like this is more metaphorical. But to the extent that it makes sense to say things like “I let myself down” / “how could I have let myself do that” / “I was disappointed in myself” I suppose you might also say “I forgive myself for that.” Sometimes people feel a lot of guilt for things that only they are privy to — secret feelings or fantasies or things they have done in private — and maybe something akin to forgiveness is a way to relieve that.

Offering someone forgiveness for something can be a backhanded way of reminding them that you find them blameworthy. It can be insulting (maybe intentionally so) to offer forgiveness to someone who does not believe they have done anything blameworthy. In this way explicitly offering forgiveness can sometimes backfire as a method of prompting reconciliation.

Sometimes people will forgive in a way that implies that they are acknowledging that they are just as flawed as the offender (“I've probably unthinkingly cut someone off in traffic myself at some point”). And occasionally forgiveness is described as though this were an important key to it: we're all sinners, we all have our good and bad sides. But sometimes you resent an act of wrongdoing precisely because it's something you consider completely beyond the pale, the sort of thing you would never do — something that cannot be explained by carelessness or ordinary human foibles. For example, I would be astonished to hear someone forgive someone who raped them by noting “after all, if things had gone a little differently, I might have raped you.”

Maybe there is a sort of forgivability-ladder: On the lowest rung are inconsiderate but common things; they can be forgiven easily and without a lot of fuss because

they're considered ordinary human failings. Above that are things that are exceptionally unkind or maybe even monstrous, but where the offender has recognized this and has corrected for it (e.g. by vowing to become better, by making amends); these things are perhaps forgivable-with-effort. And above that are exceptionally awful things for which the offender has shown no remorse; such things might remain unforgivable.

There are also "micro-forgivenesses" that are part of day-to-day [courtesy](#): "Pardon me"; "no problem" / "de nada."

Is forgiveness conditional?

There is a difference of opinion also about the extent to which forgiveness is entirely up to the discretion of the person who was wronged. Does that person need to have reasons to forgive, or can they just decide to forgive without needing to justify that decision? If forgiveness is all about laying down the burden of unpleasant internal thoughts and feelings, that may at first make it seem entirely up to you whether that's worth doing. But even in such a case, you might be criticized for being unwise in forgiving someone prematurely, or forgiving someone who is unrepentant, or if forgiving someone leads you to drop your guard against future harm caused by them. If you are too eager and willing to forgive you might be criticized for being a doormat — letting people walk all over you. Such a thing might be a symptom of a poor sense of self-worth or a lack of [dignity](#).

If forgiveness is transactional and more social, some argue that you have an *obligation* not to forgive someone until they have repented, atoned, made reparations, or some-such: that it is one of the unfortunate burdens of being victimized that you must continue to stand as an accusing witness against your victimizer, perhaps for the benefit of others in the community. If this is true, you might be behaving in a socially irresponsible way if you forgive someone from the selfish motive of getting the burden of resentment off of your shoulders. It might be *unfair* that if someone does you wrong they also burden you with the obligation to hold it against them, but so it goes.

Punishment of wrongdoers has multiple purposes — such as to discourage other wrongdoers, to prevent them from doing further harm, for restitution, to chastise or provoke penance, and for the satisfaction of their victims. Maybe the victim can renounce only the last of these in the course of their forgiveness, while society remains on the hook for enforcing the remainder.

If you refuse to forgive a wrongdoer until the wrongdoer meets some condition (apologizes, repents, makes amends, does penance), this seems to give the wrongdoer power over you: they can force you to hold onto your resentment by withholding these things. If forgiveness is mostly about laying down a burden, then this seems unwise. But if forgiveness is more like an offer (the "give" in "forgive"), then it makes more sense to make it conditional. That can lead to a stalemate, though, in which the wrongdoer sees no value in the victim's offer, and so no value in meeting the conditions, and so forgiveness never comes to pass. Depending on your perspective, that can be considered an additional and partially self-inflicted harm to the victim of the wrongdoing. On the other hand, offering conditional forgiveness can be a way of testing the waters to see if forgiveness would be wise; if your offer is rejected because the wrongdoer doesn't give a fig about healing the breach between you, that might be a good sign that you should take forgiveness off the table for now.

Forgiveness *can* be a way to signal that an apology has been accepted, that amends have been made, that atonement is complete. This can be a way of encouraging

apology and so forth by holding out the promise of some sort of token that finally lets the wrongdoer off the hook. A public declaration of forgiveness can also relieve others of any obligation they have taken on in sympathy to the victim to be vengeful to the wrongdoer on behalf of the victim.

Even if forgiveness is mostly at the discretion of the person forgiving, that person can still exercise that discretion in a vicious way (e.g. "I can forgive just about anyone except a damn dirty Jew"), and so can be blameworthy for some ways in which they exercise that discretion. If, as in that example, you forgive people or don't forgive people based on irrational/irrelevant considerations, you can inadvertently (or maybe not so inadvertently) cement irrational/irrelevant prejudices ("why are all the people I have grudges against Jewish... there must be something about them!").

What's good about forgiveness (and mercy and such)?

A virtue is a trait that characterizes someone who is living a flourishing human life. How do forgiveness/mercy/clemency help you to flourish? Here are some of the ways:

1. Forgiveness can be a relief; it can feel good directly. It has been compared to laying down a burden.
2. The Stoics would argue that it is irrational to continue to experience torment about something that is in the past and that you cannot change. You should save your emotional and intellectual labor to expend on things that are in your power to change, if you want to live the most flourishing life. Forgiveness is a way to stop this unproductive emotional churning.
3. Simmering anger about the past can sometimes make you behave poorly in the present. For example, you might lash out impatiently at some innocent person because your temper is frayed.
4. Mercy/clemency can put another person in your debt, which may be a resource you can call on when you need it. "If I am even with my enemy, the debt is paid; but if I forgive it, I oblige him for ever." —William Penn [\[7\]](#)
5. Forgiveness/mercy/clemency can contribute to social harmony and reconciliation, which can make your path through life easier.
6. Forgiveness can help you to maintain and restore personal relationships, which can be important parts of a flourishing life.
7. Forgiveness can encourage other people to be forgiving towards you, which you're bound to need eventually.
8. Forgiveness/mercy/clemency can be a mechanism to practice certain other virtues, like generosity, charity, tolerance, or magnificence.
9. Forgiveness/mercy/clemency/epikieia are themselves admirable characteristics, aside from their instrumental value.
10. Forgiveness/mercy/clemency can be a way of signalling your strength and power. If you shrug off wrongs committed against you, this can demonstrate confidence. Shakespeare wrote, "Thinkst thou it honourable for a noble man / still to remember wrongs?", [\[8\]](#) and, of mercy:

'Tis mightiest in the mightiest. It becomes
The thronèd monarch better than his crown.
His scepter shows the force of temporal power,
The attribute to awe and majesty

Wherein doth sit the dread and fear of kings,
But mercy is above this sceptered sway.
It is enthronèd in the hearts of kings.
It is an attribute to God himself.
And earthly power doth then show likkest God's
When mercy seasons justice.^[9]

The Christian perspective

"Blessed are the merciful: for they shall obtain mercy." —Jesus^[10]

Christianity strongly emphasizes forgiveness. The short prayer Jesus taught his followers to recite begs God for three things, the second of which is: "Forgive us our sins as we forgive those who sin against us."^[11] When Peter asked Jesus when enough is enough when it comes to forgiving someone, Jesus said "not seven times, but seventy-seven times"^[12] and he warned that if you ask God for forgiveness and yet you have not practiced forgiveness towards others, you'll be out of luck.^[13] The whole story of Jesus is in part the story of humans becoming reconciled with God through forgiveness of sins, as mediated through Jesus's sacrifice. In one of the gospels, Jesus during his crucifixion asks God to forgive those who have crucified him.^[14]

A big selling point of Christianity is its promise that once you've been washed in the blood of the Lamb, your sins are forgiven and you get to start with a clean slate. This suggests that there is a lot of hunger for forgiveness that Christianity helps to satisfy.

Critics of the modern self-help cult of forgiveness sometimes accuse it of being essentially old Christian assumptions dressed up in new psychological trappings.^[15]

What's good about holding a grudge, though?

It's easy to find reasons why holding a grudge is bad. The very word "grudge" sounds like a name you'd give to something bad-to-hold. In [my post on the virtue of good temper](#) I compared holding a grudge to "dragg[ing] around the rotting carcass of a grievance." Yet hold them we do. Maybe there's a good reason. Do we get anything valuable in compensation for the trouble we take to hold a grudge?

A couple of years back, the self-help pendulum began to swing back the other way, and the *New York Times* declared "[Grudges can be good!](#)", highlighting the work of Sophie Hannah and her book [How to Hold a Grudge: From Resentment to Contentment — The Power of Grudges to Transform Your Life](#). (The pendulum swung back a few months later and the *Times* reminded us: "[Let Go of Your Grudges. They're Doing You No Good.](#)") Hannah's theory, in a nutshell:

- A grudge is essentially a story, one that helps us remember how something went wrong and warns us when we see a similar pattern again. Nursing a grievance may be worth the sacrifice in present comfort if it fortifies us against future wrongs. It's a variety of mnemonic or heuristic (e.g. "this person doesn't play fair") that helps us make good decisions.
- Having a grudge is a way of asserting that the wrongdoing towards you *mattered*. In this way, it is one way the virtues of [justice](#) and [dignity](#) can play out in your life.

- Grudges can make us more aware of our values (in this case, what we detest) such that we are less apt to do contemptible things.

Hannah recommends that you try to make your grudges more conscious and deliberate. Write down your grudge stories — make them explicit — as this will help you analyze them more rationally. This is also cathartic: writing down our grudge stories gets them “out of ourselves so we’re not stuck in a feeling.” Once you have done this, go through your story step by step and ask “what could I have done differently?” This changes the focus from what was done to you in the past to how you can defend yourself in the future. This helps restore your agency and also makes your grudge more useful to you as a source of advice for how to go through life better-fortified.

Other criticisms of forgiveness

Other forgiveness-skeptics note that if we praise forgiveness, and express our admiration for people who forgive those who have wronged them, this can act as an implicit judgement against people who do not forgive for whatever reason. This can have the effect of revictimizing those people — unfairly punishing them for continuing to resent their injuries. Praising forgiveness can also be a way of prolonging continuous victimization in, for example, domestic violence situations, or when a politician trots out a variation on the old familiar “mistakes were made; let’s not focus on the past” gambit in order to avoid accountability.

People who make decisive declarations of forgiveness may also be claiming to have more conscious control over their emotions than they really do. If forgiveness is supposed to mean that you no longer have hard feelings about some past injustice, maybe this is something you can *notice* but not something you can force by diktat.

To the extent that forgiveness is supposed to erase blame, it has been criticized for being false. Philosopher John Kekes wrote, “when blaming wrongdoers is reasonable, there is no reason to forgive them; and when blaming them is unreasonable, there is nothing to forgive.”^[16]

The victim “forgiveness” anti-pattern

There is an anti-pattern of thought and behavior seen in some chronic victims, in which when they are victimized their first inclination is to wonder what they did to deserve it. Because of this, rather than feeling righteous anger towards those who victimize them, they instead may feel ironically apologetic and as though it is they who ought to make amends or attempt to mend fences. These efforts may appear superficially like a form of forgiveness (and observers may interpret them that way).

Forgiveness and status

There is some trickiness in negotiating forgiveness between people of different social status.

If a person of low status forgives a person of high status, this may be interpreted more as a display of *submission* than of forgiveness. (I’m reminded of the anti-hero Widmerpool in Anthony Powell’s *A Dance to the Music of Time*, who, as a schoolboy, on being accidentally smacked in the face with a banana by a higher-status boy, forgave instantly, in a way that his fellows explicitly interpreted as “slavish” and repulsive.)

When a high-status person forgives a low-status person, it can be viewed as condescending: sending a message of either “you are so insignificant that I can’t be expected to be overly concerned with any harm you’ve done me” or “your inconsiderate action is about par for the course of what I expected from the likes of you.”

In such circumstances of status difference, forgiveness may require nuanced attention to technique in order to avoid such undesirable side effects.

How does one forgive?

There have been some attempts to break down the process of forgiveness into parts that can be addressed individually. I was tempted to try to write up a sort of [RFC](#) (“Human Resentment/Forgiveness Protocol”) with a [UML state diagram](#) and everything, but decided that would probably be more cute than helpful.

Robert Enright (founder of the [“International Forgiveness Institute”](#)) developed a twenty-step forgiveness pathway^[17] that has roughly the following stages:

1. Take a close look at the wrong you suffered, who caused it, and the context in which it happened.
2. Examine the anger you feel about it, any shame or guilt associated with it, and how the wrong and your reaction to it have affected you since.
3. Decide whether you want to move past how you are currently dealing with the wrong and into an attitude of forgiveness instead. If so...
4. Work on understanding, compassion, and acceptance, and make a gesture of reconciliation to the offender.
5. Reformulate the way you remember your experience of being wronged and working toward forgiveness in various healthy and forward-looking ways.

Everett Worthington (director of the “A Campaign for Forgiveness Research” project) developed another pathway,^[18] roughly:

1. Acknowledge the wrong, but in a way that leaves the way open for forgiveness and affirms the worth of the offender.
2. Express your grievance to the offender *in absentia*; then turn this around and imagine yourself as the offender and try to see their perspective.
3. Forgive the offender in the form of a gift (i.e. not from selfish motives, or conditionally).
4. Formalize your forgiveness by writing it down somewhere (for yourself).
5. Revisit that note to remind yourself that you have forgiven, if your forgiveness lapses.

Worthington has [a set of do-it-yourself workbooks](#) on his website that claim to help people become more forgiving.

Frederic Luskin (author, *Forgive for Good*) has a nine-step path^[19] that, in contrast to the previous two paths, is much more focused on the internal benefits of forgiveness than on reconciliation with the offender (my paraphrase):

1. Articulate your feelings about what happened to you, and tell someone you trust.

2. Commit to doing what it takes to feel better. “Forgiveness is for you and not for anyone else.”
3. Come to an understanding of what forgiveness is and what it’s for.
4. Understand that your current distress isn’t being caused by what happened to you then, but how you are dealing with that now.
5. Use a stress management technique when your anger becomes acute. Get some distance between the event that hurt you and the response you are going to have to it.
6. Take control of what you can control; drop expectations for things and people that you cannot control.
7. Look for other ways to meet the goals that you are currently trying to meet through holding your grudge.
8. “[A] life well lived is your best revenge” so do that.
9. Change the story you tell about your grievance to highlight your process of forgiveness.

Each of these pathways makes certain assumptions about what forgiveness is and what it is meant to accomplish, however, and those things are still very much up for debate.

In 2004 a meta-analysis of several controlled studies of forgiveness-oriented psychological interventions was published.^[20] It tried to determine whether certain classes of intervention helped people to forgive, and also whether this helped their emotional health in general. It found strong support on both counts for process-based forgiveness interventions (ones that helped people go through a multi-step process like one of those described above), but no support for decision-based forgiveness interventions (ones that helped people decide to forgive). However, the process-based interventions tended to be longer in duration, so the results might be explained by the quantity instead of the quality of the two approaches.

Another meta-analysis of group-based forgiveness interventions to determine how well they work to increase self-reported forgiveness (or decrease “unforgiveness”) concluded that “The data appear to speak clearly: Forgiveness interventions are effective.”^[21] That’s not too shocking — forgiveness interventions improve self-reported forgiveness — but it does mean you can headline: “Science Discovers How To Forgive.”

For what it’s worth, another meta-analysis also looked into how forgiveness interventions affected depression, anxiety, and hopelessness, and concluded “interventions designed to promote forgiveness are more effective at helping participants achieve forgiveness and hope and reduce depression and anxiety than either no treatment or alternative treatments.”^[22]

How does one become forgiving, merciful, etc.?

“Forgiveness is not an occasional act; it is a constant attitude.” —MLK, Jr.^[23]

To rise to the level of a virtue, your forgiveness or mercy or clemency needs to be something that has become a part of your character: you need to have become a forgiving, merciful, clement person. This doesn’t mean that you are *maximally*

forgiving, merciful, and clement (which would put you into doormat territory) but that you are skilled at practicing forgiveness, mercy, and clemency in the right amounts, in regards to the right people and actions, at the proper time, and in the best way.

Usually forming a virtue like this is a matter of practice and habit. In the cases of these virtues in particular, it may require that you give yourself a bit of a nudge. “I have an old grievance I’ve become kind of fond of chewing on. Might now be a good time to practice how to give that up?” or “I would be entirely within my rights to demand my pound of flesh from this so-and-so, but might now be a good time to practice mercy?” If you want to be a merciful, forgiving person you need to be on the lookout for opportunities, which may happen when you are least in the mood for them.

Forgiveness and mercy may largely depend on having a certain disposition, but clemency and *epieikeia* also require a well-tempered sense of justice and practical wisdom. A good cook can safely improvise to improve a recipe, but a bad cook had better stick with what’s printed in the book. Similarly, if you aren’t pretty sure you know what you’re doing, your attempts at clemency and *epieikeia* might just be indulgence and bias dressed up in fancy clothes. We may be able to learn how to practice those virtues well by carefully examining precedents or by learning from people who are especially Solomonic.

1. ^

[“Forgiveness: Your Health Depends on It”](#) Johns Hopkins Medicine: Wellness and Prevention

2. ^

Friedrich Nietzsche, *Thus Spoke Zarathustra* (1883) II “On the Tarantulas”

3. ^

Aristotle, *Rhetoric*, book I, chapter 13

4. ^

Alex H.S. Harris, Frederic Luskin, Sonya B. Norman, Sam Standard, Jennifer Bruning, Stephanie Evans, & Carl E. Thoresen, [“Effects of a Group Forgiveness Intervention on Forgiveness, Perceived Stress, and Trait-Anger”](#) *Journal of Clinical Psychology* (2006)

5. ^

P. Hieronymi [“Articulating an uncompromising forgiveness”](#) *Philosophy and Phenomenological Research* (2001)

6. ^

Martin Luther King, Jr., “Loving Your Enemies” *Strength To Love* (1963)

7. ^

William Penn, *Some Fruits of Solitude in Reflections and Maxims* (1693)

8. ^

William Shakespeare, *Coriolanus*, spoken by Volumnia

9. ^

William Shakespeare, *Merchant of Venice*, spoken by Portia

10. ^

Matthew 5:7

11. ^

Matthew 6:12

12. ^

Matthew 18:22

13. ^

Matthew 6:15

14. ^

Luke 23:34

15. ^

See, for example, [Murali's comment on "Forgiveness"](#) at *Crooked Timber*, 24 May 2019: "Forgiveness fetishism, as I have observed it, seems to be a very western/white person thing. It seems like a holdover from a christian past and people who advocate it seem to be finding any sort of excuse to preach what they learned when they were young even if they currently profess to no longer being christian."

16. ^

John Kekes, ["Blame vs. Forgiveness"](#) *The Monist* (2009)

17. ^

[Enright Forgiveness Process Model](#)

18. ^

[REACH Forgiveness of Others](#)

19. ^

[Forgive for Good: 9 Steps](#)

20. ^

Thomas W. Baskin & Robert D. Enright, ["Intervention Studies on Forgiveness: A Meta-Analysis"](#) *Journal of Counseling & Development* (2004)

21. ^

Nathaniel Wade, Everett Worthington, & Julia Meyer, "But Do They Work? A Meta-Analysis of Group Interventions to Promote Forgiveness" *Handbook of Forgiveness* (2005)

22. ^

Nathaniel Wade, William Hoyt, Julia Kidwell, & Everett Worthington, "Efficacy of Psychotherapeutic Interventions to Promote Forgiveness: A Meta-Analysis" *Journal of consulting and clinical psychology* (2013)

23. ^

Martin Luther King, Jr., "Love In Action" (~1963)

Notes on Integrity

This post examines the virtue of **integrity**. It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is integrity?

“[The nobler type of man] first practices what he preaches and afterwards preaches according to his practice.” —Confucius^[1]

Integrity is strangely slippery to define. It seems to have both a narrow and a broad definition, and it's not always clear which of these a person has in mind when they use the word:

- **Narrowly:** “integrity” follows its etymological connection with the Latin *integer*, meaning intact, undivided, whole, coherent. In this sense, you have integrity if your words match your actions and you are not at cross-purposes with yourself. You are not two-faced or hypocritical. You don't tie yourself in knots trying to be all things to all people.
- **Broadly:** if you look up definitions of integrity, you'll often see many other virtues bundled in beside this narrow meaning, including authenticity, [honesty](#), rectitude, incorruptibility, reputability, [fairness](#), harmony, consistency, reliability, accountability, and decency. Sometimes “integrity” is used as a synonym for [honor](#). Occasionally someone will just gather a bunch of their favorite, seemingly unrelated, virtues together and cover the whole set with the “integrity” [halo](#).^[2]

In this post, I'm going to stick with the narrower of these definitions.

How integrity helps you thrive

The most obvious way that integrity helps you to thrive is that it keeps you from acting at cross-purposes: all of your oars go in the water at the same time.

Integrity can also be socially desirable. Someone with integrity is more of a predictable known quantity. You know where they stand. This makes it easier to cooperate with them in shared projects.

Integrity also tends to travel alongside nice things like self-knowledge, a sense of purpose, and being comfortable in your own skin, though integrity may grow out of such things as much as it contributes to them.

Hazards of pseudo-integrity

"A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines." —Ralph Waldo Emerson^[3]

Sometimes people misfire in their attempts at achieving integrity. Someone may become simply stubborn — unwilling to give up on a bad idea because they've already made up their mind. Integrity, to such a person, means sticking to your guns no matter what.

Or someone may stumble upon some attractive and simple principle or theory (falsifiability! objectivism!) and make an *idée fixe* out of it — trying to cram all of complex reality into its narrow bounds: achieving a sort of counterfeit integrity by rejecting any evidence that does not conform to the theory.

Integrity does not require uniformity or rigidity. If in some contexts you are more loose and others more formal, if you [switch dialects](#) depending on whom you are speaking with, if you sometimes speak "as your doctor" and sometimes "as your friend," this may just mean that you are flexible, not that you lack integrity. There is an art to discerning when you've gone too far and put on so much makeup that you're disguising yourself.

Hazards of a fetish for integrity itself

Practicing what you preach is, all else being equal, superior to hypocrisy. But if either what you preach or what you practice is harmful nonsense, making sure your preaching and practicing are in alignment is the least of your troubles.

A disproportionate number of the moral arguments we hear these days have to do with hypocrisy. Public Figure says X but their actions prove they believe $\neg X$, or they said Y yesterday but say $\neg Y$ today. There is something suspiciously *lazy* about this sort of critique. Alasdair MacIntyre argued that this fetish with what he called "unmasking" has emerged because we have lost the ability to [rationally](#) argue about ethics, in a sort of Tower of Babel-like disaster.^[4] We don't know how to confidently distinguish whether X is right or wrong, but we can still notice that $X \wedge \neg X$ is false, and so we lean into that with all we've got.

"No one is any longer carried away by the desire for the good to perform great things, no one is precipitated by evil into atrocious sins, and so there is nothing for either the good or the bad to talk about, and yet for that very reason people gossip all the more, since ambiguity is tremendously stimulating and much more verbose than rejoicing over goodness or repentance over evil." —Søren Kierkegaard^[5]

Integrity as a process

Do I contradict myself?
Very well then I contradict myself,
(I am large, I contain multitudes.)

—Walt Whitman^[6]

Integrity, in the narrow sense, is usually described as a quality — an achievement or accomplishment maybe — something you *have* as opposed to something you *do*. But it seems to me that integrity can be better understood as an ongoing process:

1. Notice some contradiction in yourself: some way in which [your beliefs, your words, and your actions don't correspond](#). (A form of “notice your confusion.”)
2. Analyze the contradiction and carefully try to understand the basis of the conflict, then either remove the contradiction by making your beliefs, words, and actions conform, or find some way to synthesize the apparent contradiction harmoniously.

The first of these steps requires self-awareness and attention. In my experience, the more self-awareness and attention I learn to apply, the more contradictions I find; I've yet to discover a fundament of pure integrity. This step also is aided by [humility](#): you can look most penetratingly at your flaws if you don't reflexively explain them away to defend an inflated self-image.

I'm going to go out on a limb and suggest that the second step requires [empathy](#). Here's why: When you notice a contradiction in yourself, this contradiction will typically be between different modes in which you think and operate. It may be (for example) a conflict between your conscious motives and more subconscious ones, or between your public persona and your private self-image, or between your emotions and your reason, or between how you approach a problem in [far mode vs. near mode](#) or while thinking [fast vs. slow](#). Once you have noticed this contradiction and you try to analyze it, the process of analyzing will itself belong to one of these modes (e.g. conscious, private, reasoning, slow), and will therefore be initially biased towards the conclusion of that mode. If you then quickly move to try to resolve the contradiction, you will be in effect trying to overpower the other modes with the modes you are doing the analysis with — trying to persuade the non-analytical mind with rhetoric the analytical mind finds persuasive. This won't lead to *integrity* but at best to a hyperanalytical imbalance (more often, it just leaves the initial contradiction intact, as the non-analytical modes remain as unpersuaded as they were before the analysis). So as you analyze your contradiction, you need to empathize with non-analytical modes of thinking, in order to give those modes their day in court and allow those modes to participate in forming a harmonious synthesis. The same sorts of skills that allow you to empathize with different people who have different values and worldviews and experiences and so forth also allow you to empathize with the different parts of yourself that are in tension with each other. To the extent that you can promote these empathic skills in the analytical part of yourself, the analytical part of yourself will be more successful in coordinating your integration, which is to say, promoting your integrity.

(After writing this, I read the description of [Internal Double Crux](#) which also suggests an empathetic approach to resolving internal contradictions.)

For example: effective altruism

Effective altruism is an example of this integrity-building process. For example, let's say I look at my charitable giving and notice that while I think of myself as being motivated by an altruistic desire to relieve suffering, I do not give to charity in a way that is rationally crafted to do so most effectively. I have noticed a contradiction between my professed motives and my actions. I analyze this contradiction in myself, looking at the various ways I have given to charity and what other motives were at

play. Maybe I gave to one charity because a friend of mine was involved in it and I wanted to show my support for their efforts. Maybe I gave to another because I was emotionally touched by prominent news stories of an acute emergency like a tsunami or refugee crisis. Maybe I gave to a third because it made me feel solidarity to some social cause I'm sympathetic to.

In this way I come to realize that I have more motives than I'd given myself credit for. Now I can make better decisions about which of these motives I want to respect, and how I want to prioritize them. For example, maybe I don't think "being emotionally touched by an acute emergency" is an especially good criteria to use when picking a charity to give to. But maybe I can compromise with the part of me that feels cathartic relief from giving in such circumstances. I might do so by taking some time to form a more emotional reaction to a chronic emergency that my donation could address more effectively (e.g. malaria). I could read up on the subject, view a documentary, or envision something concrete like "x dollars will save one life." I might also look for a way to combine my giving with some sort of social display, so that the part of me that wants to get kudos from friends or to demonstrate solidarity gets some attention.

Role-play as a danger to integrity

"To have integrity is to refuse to be, to have educated oneself so that one is no longer able to be, one kind of person in one social context, while quite another in other contexts. It is to have set inflexible limits to one's adaptability to the roles that one may be called upon to play." —Alasdair MacIntyre^[7]

A common threat to integrity comes when we allow ourselves to be overcome by the roles we are sometimes called upon to play. In such scenarios, you compartmentalize yourself as me-playing-a-role and separate that from me-being-me. Having done that, you then can apply a different set of rules to your behavior, attributing those rules to the role and not to yourself. This lets you get into all sorts of mischief (and that is part of what makes it tempting).

For instance, you might do something at the behest of your employer that you would be ashamed to do otherwise, but then excuse yourself by thinking "don't blame me; that's just my job." The "I was only following orders" excuse is a more notorious variety of this.

The way to respond to this threat/temptation is not to try to give up on playing roles altogether, but to refuse to compartmentalize and separate me-playing-a-role from me-being-me. Remember instead that me-playing-a-role is *an example of* me being me, and that a single set of rules apply, and you'll be fine, integrity-wise.

1. ^

Analects of Confucius, II.xiii

2. ^

For example, "[Integrity: Definition and Examples](#)" at the *Indeed Career Guide*

3. ^

Ralph Waldo Emerson, "Self-Reliance" in *Essays: First Series* (1841)

4. ^

Alasdair MacIntyre, *After Virtue* (1981)

5. ^

Søren Kierkegaard, *The Present Age* (1846)

6. ^

Walt Whitman, *Song of Myself*, section 51

7. ^

Alasdair MacIntyre, "Social Structures and their Threats to Moral Agency"
Philosophy (1999)

Notes on Humility

This post examines the virtue of **humility**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is humility?

Humble people are aware of their fallibility and imperfection, and of the smallness and brevity of their lives, and act accordingly. They avoid flattering themselves, and are less vulnerable to the flattery of others. They know that while they may be particularly special to themselves, the rest of the universe does not have to go along with that assessment. They have the courage to occasionally empathize with the universal, objective point of view in which they are an ephemeral spark doomed to be extinguished into an eternal obscurity.

A humble person does not mistake confidence for accuracy, or the limits of what they know with the limits of what is knowable. Humble people do not become defensive or flustered when they discover they were mistaken, but take this as a matter of course, and adjust course accordingly. They are not so proud of their self image and opinions that they will refuse to trade them for better ones.

Humility is one of [Eliezer Yudkowsky's "Twelve Virtues of Rationality"](#) — described there in this way:

To be humble is to take specific actions in anticipation of your own errors. To confess your fallibility and then do nothing about it is not humble; it is boasting of your modesty. Who are most humble? Those who most skillfully prepare for the deepest and most catastrophic errors in their own beliefs and plans. Because this world contains many whose grasp of [rationality](#) is abysmal, beginning students of rationality win arguments and acquire an exaggerated view of their own abilities. But it is useless to be superior: Life is not graded on a curve. The best physicist in ancient Greece could not calculate the path of a falling apple. There is no guarantee that adequacy is possible given your hardest effort; therefore spare no thought for whether others are doing worse. If you compare yourself to others you will not see the biases that all humans share. To be human is to make ten thousand errors. No one in this world achieves perfection.

Humility seems at first to conflict with the virtue of pride (or related self-aggrandizing virtues like [honor](#), magnificence, boldness, or [ambition](#)). There is disagreement about whether pride even is a virtue; in the Christian tradition it is a vice. Maybe pride is best thought of as being at an Aristotelian golden mean between arrogance and poor self-esteem. You would tell an arrogant person to work on their humility, and an abased person to work on their pride — and these wouldn't be contradictory, but would be context-appropriate ways of telling two differently-oriented people to aim for the same virtuous mean.

Character Strengths and Virtues: A Handbook and Classification suggests that the opposite of humility is not pride, but *narcissism*: a defensive obsession with an unrealistically grandiose self image.^[1]

Humility is a component of some other virtues, quasi-virtues, and supposed virtues. For example, **modesty** is in part the proper social expression of humility. If you lack humility to the extent that you cannot acknowledge your own faults, you will also lack the quasi-virtue of **shame**. Humility can help you to decenter or deemphasize the ego, and can thereby help you with **selflessness** and **empathy**. If you don't think of yourself as the measure of all things, you will be more likely to exhibit **submission** or **obedience** when these things are called for. And measuring yourself appropriately infinitesimally in reference to the divine can be a crucial ingredient of **piety**.

What humility isn't

"No doubt, when modesty was made a virtue, it was a very advantageous thing for the fools; for everybody is expected to speak of himself as if he were one."

—Schopenhauer^[2]

There are a number of non-virtuous things that sometimes get mistaken for the virtue of humility, such as:

- **false modesty** and **self-deprecation** — insincerely speaking of yourself as being lowlier, worse, less able, etc. than you actually feel yourself to be (sometimes this is a culturally-learned convention of politeness, other times it is more of an unskillful variety of modesty)
- **poor self esteem** or feelings of inferiority — sincerely feeling yourself to be lowlier, worse, less able, etc. than you are, and unable to improve or unworthy of being any better
- **obsequiousness** — over-willingness to cede to another person's opinions or agenda
- **sheepishness** — a sort of chronic, cringing shame or embarrassment
- **humiliation/debasedness** — behaving in a way that seemingly is intended to demonstrate one's lack of dignity and self-worth

What is it good for?

"Perfection is impossible without humility. 'Why should I strive for perfection, if I am already good enough?' " —Tolstoy^[3]

Humility is an important ingredient in human thriving in part because it helps you gain a better understanding of what *human* thriving is. If you in your heart of hearts think of yourself as something of a god or as the crown of creation, you may try to thrive in a way appropriate to such an Olympian instead. If you acknowledge yourself to be the brief bag of bones that you are, your aspirations will better match your true condition.

Humility aids self improvement. A big disadvantage of a lack of humility is in thinking that you've already got all the answers: You fail to learn from others (what do *they* know?), you don't try harder because you think you've already become as good as can be, and you don't recognize your mistakes and so cannot learn from them. A humble person takes penetrating looks at the weakest points their theories; a proud

person gazes fondly at the strongest and most impressive parts and sweeps the flaws under the rug.

A humble person doesn't embarrass so easily. If they screw up or are mistaken, they shrug and think "well, no surprise I'm not perfect," and calmly learn from their mistake. Without humility, they might instead get flustered, or make excuses, or deny that anything went wrong in order to defend a faulty self-image.

A humble person knows that common human cognitive biases, foibles, and bad judgement calls aren't just mistakes that other people make. Humble people seek accurate information about themselves; proud people seek information that confirms their greatness. The paper "Humility: Theology Meets Psychology" by David G. Myers has an amusing run-down of the many ways in which self-flattery gives people a ridiculous view of the world and their place in it.^[4] For example: "In one College Entrance Examination Board survey of 829,000 high school seniors, 0 percent rated themselves below average in 'ability to get along with others'..."

The name "LessWrong" is a clever reminder to be intellectually humble as you try to improve your thinking. A similar approach might wisely be taken to improve your development in the other virtues: consider being "[less bad](#)", "[less dishonest](#)", or "[less lazy](#)" for example.

How can you develop humility?

"Think of the whole universe of matter and how small your share. Think about the expanse of time and how brief—almost momentary—the part marked for you. Think of the workings of fate and how infinitesimal your role." —Marcus Aurelius^[5]

How do you bootstrap humility? If you begin by being ridiculously impressed with yourself, you may believe [you have already reached the pinnacle of humility](#). There's something of a paradox: the more you lack humility, the less you will be conscious of that lack or motivated to do anything about it.

Benjamin Franklin (see below) was astute enough to listen undefensively to a friend who told him that he was thought of as being conceited and overly concerned with winning arguments. The fact that he was able to listen to this criticism and then consider it gracefully and learn from it was key to his deciding to work on becoming more humble (or at least more modest).

Franklin's experience suggests that confidence and self-esteem, which both make self-criticism less threatening, are important in this bootstrapping process. Grandiosity, braggadocio, and self-obsession seem often to be compensatory reactions to insecurity and poor self worth: You brag about yourself constantly because you're trying to convince yourself; you don't believe in yourself so you try to believe in your press releases instead.

People sometimes describe as "humbling" the experience of considering the brevity of human life in relation to the vastness of time, the smallness of the Earth's thin ecoshell in the vast empty chasm of space, the meagerness of our knowledge in the face of all that remains unknown, and so forth. The poem [Ozymandias](#) is a fine meditation on the absurdity of delusions of grandiosity in these contexts.

Sometimes people will deliberately take on lowly tasks (e.g. cleaning latrines) as a way of chastening excessive pride. The Christian practice of washing another person's feet (e.g. by the Supreme Pontiff of the Universal Church himself^[6]) is an example of this.

Practicing [gratitude](#) can be a way of sharing credit with others that lessens the temptation to attribute one's good fortune to one's own marvelousness.

Ben Franklin's experience

Early in his life, Benjamin Franklin launched a personal project of methodical improvement in the virtues. He picked a set of virtues that he thought were particularly important, and concentrated on each one in turn, doing a daily accounting of each virtue he was practicing. He created a notebook with a table for each week. The table had one column for each day of the week, and one row for each of his virtues. Each time he failed to fulfill a particular virtue on a certain day, he marked the table cell for that virtue/day with "a little black spot" (or more than one if he screwed up multiple times). The plan was that when he achieved a week in which he successfully kept the row for Temperance blank, he would move on to concentrating on Silence (attending to Temperance as well). When he managed to keep both of those rows clear for a week, he would move on to Order, and so on.

"I was surpris'd to find myself so much fuller of faults than I had imagined; but I had the satisfaction of seeing them diminish." He carried his book around for several years. "[T]ho' I never arrived at the perfection I had been so ambitious of obtaining, but fell far short of it, yet I was, by the endeavour, a better and a happier man than I otherwise should have been..."

When Franklin first chose his virtues, he apparently bragged about his plan to a friend, for as he says in his autobiography:

My list of virtues continued at first but twelve; but a Quaker friend having kindly informed me that I was generally thought proud, that my pride showed itself frequently in conversation, that I was not content with being in the right when discussing any point, but was overbearing and rather insolent, of which he convinced me by mentioning several instances, I [determined](#) endeavoring to cure myself, if I could, of this vice or folly among the rest, and I added Humility to my list, giving an extensive meaning to the word.

I cannot boast of much success in acquiring the *reality* of this virtue, but I had a good deal with regard to the *appearance* of it. I made it a rule to forbear all direct contradiction to the sentiments of others, and all positive assertion of my own. I even forbid myself, agreeably to the old laws of our [Junto](#), the use of every word or expression in the language that imported a fixed opinion, such as *certainly*, *undoubtedly*, etc., and I adopted, instead of them, *I conceive*, *I apprehend*, or *I imagine* a thing to be so or so, or it so *appears* to me at *present*. When another asserted something that I thought an error, I denied myself the pleasure of contradicting him abruptly and of showing immediately some absurdity in his proposition; and in answering, I began by observing that in certain cases or circumstances his opinion would be right, but in the present case there *appeared* or *seemed* to me some difference, etc. I soon found the advantage of this change in my manner; the conversations I engaged in went on more pleasantly.

The modest way in which I proposed my opinions procured them a readier reception and less contradiction; I had less mortification when I was found to be in the wrong, and I more easily prevailed with other to give up their mistakes and join with me when I happened to be in the right.

And this mode, which I at first put on with some violence to natural inclination, became at length so easy, and so habitual to me, that perhaps for these fifty years past no one has ever heard a dogmatical expression escape me. And to this habit (after my character of integrity) I think it principally owing that I had early so much weight with my fellow-citizens when I proposed new institutions, or alterations in the old, and so much influence in public councils when I became a member; for I was but a bad speaker, never eloquent, subject to much hesitation in my choice of words, hardly correct in language, and yet I generally carried my points.

In reality, there is, perhaps, no one of our natural passions so hard to subdue as *pride*. Disguise it, struggle with it, beat it down, stifle it, mortify it as much as one pleases, it is still alive, and will every now and then peep out and show itself; you will see it, perhaps, often in this history; for, even if I could conceive that I had completely overcome it, I should probably be proud of my humility.

1. ^

Christopher Peterson & Martin Seligman, *Character Strengths and Virtues: A Handbook and Classification* (2004), chapter 20

2. ^

Arthur Schopenhauer, *The Wisdom of Life (Essays)* chapter IV, section 2 “Pride”

3. ^

Leo Tolstoy, *A Calendar of Wisdom* (1910), January 11

4. ^

David G. Myers, “Humility: Theology Meets Psychology” *Reformed Review* (1995)

5. ^

Marcus Aurelius, *Meditations* V.24

6. ^

Devin Watkins [“Pope washes inmates' feet at Mass of Lord's Supper”](#) Vatican News 18 April 2019

Notes on Optimism, Hope, and Trust

This post examines the virtues of **hope**, **optimism**, and **trust**. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What are these virtues?

These virtues have in common a sort of “look on the bright side” / “expect the best” approach to life. But there are a number of ways to interpret this, and if we are looking for a virtue — that is, a characteristic disposition that promotes human flourishing — we would be wise to be precise and careful.

There is a naive version of optimism that I would not want to try to defend on LessWrong:

the belief that the probability of an outcome is increased by one's positive disposition toward it

So let me say at the start that this is just one possible way the optimistic outlook can be defined, and that others can bear more [rational](#) weight.

There is little controversy about hope, optimism, and trust being parts of a flourishing life. However there is controversy about whether encouraging a hopeful, optimistic, trustful outlook puts the cart before the horse. Are hope, optimism, and trust *ingredients* of a life well lived, or are they *results* of a life well lived (or perhaps of good fortune)?

Related virtues and vices

These virtues are closely related to some others: Cheer and joy, for instance, are easier to maintain when one is optimistic. Confidence, boldness, and [courage](#) are aided by the hope of triumph. Imagination can help you to discover possible good outcomes to be hopeful about. Intimacy, openness, and vulnerability can be boosted by trust. Solidarity can include an extension or expectation of trust. Trust also complements trustworthiness. Richard Rorty thought shared hope was at the base of civility.^[1]

There is some tension between optimism and [caution](#), if optimism causes you to underweight the possibility of bad outcomes that you ought to prepare for.^[2] If optimism becomes an excuse to discount evidence that disagrees with a positive outlook, it can [undercut rationality](#) and skepticism.

The vices associated with a lack of optimism / hope / trust include cynicism, distrust, doubt, paranoia, suspicion, pessimism, and despair. Sometimes lack of hope in particular becomes fatalism, a feeling of lack of agency (“nothing I do matters”), or, philosophically, a loss of faith in free will. There are also vices associated with an excess of optimism / hope / trust, like unpreparedness, gullibility, or bliss-bunnyishness. [“Williams syndrome”](#) is a genetic disorder that includes a dangerous overabundance of trust among its symptoms.

The catalog *Character Strengths and Virtues* combines hope/optimism with “future-mindedness” or “future orientation”—which is to say that someone with this virtue has evaluated possible futures and judged which would be the better ones, and has some sort of plan for getting to one of those.^[3] This makes the virtue less passive than hope/optimism might imply.

Hope

The *Catholic Encyclopedia* defines hope as “the desire of something together with the expectation of obtaining it.”^[4] However, in common use people often hope for things they do not expect or even think likely (e.g. to win the lottery). Hope does seem to at least require some possibility that the hoped-for thing could come to pass, but also some possibility that it might not: you cannot “hope” for something that is a certainty (though you might have a pleasant expectation of it) or an impossibility (though you might wish it could be otherwise). Hope (secular hope, anyway) isn’t the *expectation* that the hoped-for thing will come to pass but includes the tension of some fear that it will not.

Usually, hope refers to the future. Occasionally, however, people seem to express hopes for past events (“I hope I remembered to turn off the oven”). Some people interpret this as a sort of roundabout way of expressing a future hope (“I hope I find the oven is off when I return home”).

It probably goes without saying, but a thing hoped for is always a positive thing (for the hoper, anyway). You can have expectations of positive or negative things, or fear / dread / foreboding about negative things, but hope always has positive content.

Hope can be motivating: it allows you to imagine the possibility of a future state and its benefits, and so can help drive you to do what needs doing to get there. A sort of mundane hope underlies most all purposeful activity: you hope that by doing something you will get some result. Hope can also be motivating in the way it helps you to set your sights on long-term goals. For example, the hope of what you will accomplish as a doctor can help you endure your long hours of residency. Hope can in this way be active, more like aspiring than passively anticipating.

Hope might also be thought of as a sort of imagination that helps to prepare us for the future. In this it is a complement to fear: fear encourages us to prepare for a future where fears come to pass; hope prepares us for a future where hopes come to pass. (But we don’t call “fear” a virtue; instead we talk about courage or caution or prudence or preparedness. Maybe we need something similar for hope.) [Utopian](#) and dystopian writings are elaborate sorts of hopes and fears that ask us to imagine possible futures in a way that invites us to prepare for them or be vigilant about them.

Hope is inherently valuable in that imagining positive things is a pleasant sort of daydreaming. Hoping also helps you to discern and make salient your values: you can learn what you value in part by paying attention to what you hope for.

Intransitive, superstitious, and self-fulfilling hopes

What I have described so far is “hope for” something. There is also a sort of intransitive hope — hope in general — that is more like what I’ll cover in the “Optimism” section below. Hope as a virtue — habitually, characteristically adopting a hopeful stance — resembles this intransitive sort of hope, though it may exhibit itself through specific acts of transitive hope.

There is also the superstitious hope that I alluded to earlier: “wishful thinking” in which you have the irrational belief that your hope will influence future events to align with the contents of your hope. There is all sorts of common magical thinking surrounding this kind of hope: totems and rituals and beliefs that if you just hope strongly enough or sincerely enough you can materialize your hopes just like that. In its worst form, hope of this sort can displace the kind of practical action that might actually help make hopes come to pass.

That said, there are such things as self-fulfilling prophesies in the realm of hope (see the discussion of William James below for more on this). By anticipating fortunate opportunities, we may better prepare ourselves to take advantage of them if they arrive, so in this way hope can indeed help good things come to pass. If we believe superstitiously that

"everything happens for a reason," this may prompt us to look more confidently for the silver lining in the cloud, which may help us find it where we otherwise would have missed it. And sometimes "hope" is the name we use to describe confidence that comes from careful preparation and experience.

There is also such a thing as an irrational *lack* of hope: a superstitious pessimism ("I've always had bad luck") or paranoia. In such a case, something like "hope" might be recommended as a corrective to help you evaluate reality more rationally. It may be that the virtue is called "hope" not so much because it is a good thing to be more hopeful than reason permits, but because on the continuum of hopeful-to-hopeless the hopeless side is more harmful and so it is wise to err on the hopeful side, all else being equal.

Hope as an intellectual virtue

Hope is sometimes described as an intellectual virtue. It can come to the assistance of intellectual, rational pursuits. William James (see below) noted, for example, that while science declares allegiance to dispassionate evaluation of facts, the history of science shows that it has often been the passionate pursuit of hopes that has propelled it forward: scientists who believed in a hypothesis before there was sufficient evidence for it, and whose hopes that such evidence could be found motivated their researches. Nancy Snow says that hope works as an intellectual virtue in three ways: "(1) hope that knowledge/truth can be obtained furnishes a motivation for its pursuit; (2) hope imparts qualities to its possessor, such as resilience, perseverance, flexibility, and openness, that aid in the pursuit of knowledge/truth; and (3) hope, through imparting such qualities to its possessor, functions as a kind of method in the pursuit of knowledge/truth."^[5]

Hope in a Christian context

Faith, hope, and charity (or [love](#)) are [the traditional Christian virtues](#). In a Christian context, hope is the confidence that God has your back, and will extend His help to you in your efforts to "reach eternal felicity."^[4] Hope is said to be an "infused virtue"^[4] (one that is implanted in us by God, not one that we develop ourselves through habit). There ultimately is no hope to be had in the material world and our mortal lives; if we try to put our hope there, we will end in despair (the opposite of hope). Hope is a matter of *will*, not of *intellect*: hope "elevate[s] and strengthen[s] our wills"^[4] as we work toward unity with God. If we reject God, we end up losing hope ("[Abandon Hope All Ye Who Enter Here](#)" is legendarily inscribed on the gates to hell).

William James

The philosopher William James believed he had discovered rational grounds for making our opinions about the truth depend to some extent and in some circumstances on our emotional disposition towards it.^[6]

In short, James said that when

1. you have [to decide](#) between hypotheses and cannot just remain comfortably in doubt,
2. you are unable to rationally and scientifically decide between hypotheses (no time, no data, competing hypotheses match the data equally well), and
3. it would be better for you if some particular hypothesis were the truer one

then you are justified in choosing — indeed you ought to choose — the preferred hypothesis as the one to provisionally believe.

For one thing, he notes that for some hypotheses, belief in them can have causal influence on their being true. For instance, if you believe that someone is your friend, and you therefore behave in a friendly way toward them, this may influence them to become your friend if they were previously on the fence about it. If you do not believe that anyone can be

trusted, you will never extend trust to anyone, and sure enough you will inhabit a world in which nobody behaves in trustworthy way towards you. There are many things like this in life, James says, where if you treat them as “live hypotheses” you can in fact bring them about, just by orienting your life in such a way as to make a space for them to occupy.

James applied this idea to a sort of generic religious hypothesis (that the most perfect and most eternal things are the important ones, and that this religious outlook underlies the best sort of life). This is an unproven hypothesis, and one that James asserts you have to accept or reject: there’s no middle ground. If you wait until you have air-tight proof either way, you’ll be waiting your whole life — and this amounts to the same thing in practical terms as to reject the hypothesis dogmatically. The risk of accepting the hypothesis if false is the fear that you will have been duped by a fairy tale; the hope of accepting the hypothesis if true is that you don’t waste your “sole chance in life of getting upon the winning side.” Given this balance: “what proof is there that dupery through hope is so much worse than dupery through fear?” (This of course resembles [Pascal’s Wager](#).)

James goes further and says that the religious outlook tends to personalize the perfect/eternal in such a way that — as with our beliefs in trust or friendship — our religious beliefs can have causal influence on the truth value of those beliefs. In order to live in a god-filled, hopeful universe, James suspects, you must meet the religious hypothesis half-way: you must extend some belief in its direction in order to receive the flow of evidence that supports the belief in return. Stubborn skepticism and despair becomes a self-fulfilling prophecy (in the same sense as with the person who does not believe people can be trusted), and, according to James, is thereby irrational: “A rule of thinking which would absolutely prevent me from acknowledging certain kinds of truth if those kinds of truth were really there, would be an irrational rule.” [Take the hypnotoad pill, man.](#)^[7]

James ends his essay by quoting Fitz James Stephen:

We stand on a mountain pass in the midst of whirling snow and blinding mist through which we get glimpses now and then of paths which may be deceptive. If we stand still we shall be frozen to death. If we take the wrong road we shall be dashed to pieces. We do not certainly know whether there is any right one. What must we do? “Be strong and of a good courage.” Act for the best, hope for the best, and take what comes... If death ends all, we cannot meet death better.^[8]

Optimism

Optimism is the belief that things are good, or at least are heading in that direction. It sometimes also includes the belief that we ourselves are doing well or getting better, and that we have some control over our fate. [John Dewey](#) preferred the term “meliorism” which emphasized this active element: a meliorist is optimistic that our efforts can make things better.^[9]

Is the universe benign, malign, or indifferent? Can people be good to one another, or does self interest make them necessarily antagonistic? Is the future bright or are we doomed? Is life worth living or would it be better had we never been born? Is history a story of progress and enlightenment or an embarrassing chronicle of horrors? Is there a caring God who watches over us, or are we playthings of cruel forces? Optimists know which side of these questions they’re rooting for, and they may be tempted to tilt the scales a bit as they weigh the evidence.

The philosophy of optimism

Optimism has at times been taken to unlikely extremes. There was an 18th-Century vogue for philosophical proofs that the universe we inhabit is not merely good, but [“the best of all](#)

possible worlds." The more typical optimist is satisfied with the belief that life is on the whole significantly more good than not.

Optimism at first glance seems to presuppose a scale of good and bad on which you measure your life, or history, or the universe, or whatever, in order to then find it good. Some philosophers have tried another approach to optimism, which is to *unconditionally* love their life without first judging it in this way. Nietzsche put it this way: "My formula for greatness in a human being is *amor fati*: that one wants nothing to be different, not forward, not backward, not in all eternity. Not merely bear what is necessary, still less conceal it—all idealism is mendacity in the face of what is necessary—but *love it*."^[10]

Most philosophies and religions are optimistic, even the ones that are very critical of people or very suspicious of the material world. They typically either explain why things are good, or explain how to get out of a bad predicament. LessWrong-style rationalism is full of this sort of optimism: we are plagued by cognitive biases, yet with effort we have the power to more closely approach the truth; artificial intelligence threatens to extinguish human values, but we can learn how to restrain it; death is everyone's destiny, but perhaps we can be the generation to defeat destiny!

Some philosophies make an effort to show how certain paths to optimism are dead ends but others remain open. If we condition our optimism on being able to *avoid* things like sickness, aging, and death, the Buddha says, we will inevitably fail—yet we can still *transcend* suffering. If we lay up treasures upon earth, where moth and rust doth corrupt, where thieves break through and steal, we will be disappointed, says Jesus—but we can lay up for ourselves treasures in heaven and then we've got it made in the shade. If we condition our happiness on things that are not in our control, we will be forever at their mercy, says Epictetus—but we can stop doing that, and when we do so, the world is our oyster.

It is a rare philosopher who says "nope; it's hopeless"—even Schopenhauer, who is usually trotted out as the poster child for pessimism, seems to have left an escape route open to certain exceptional people. The optimist in me thinks that this is a good sign: whatever else they disagree on, the philosophers agree that things are more-or-less okay and maybe even marvelous. The pessimist in me wonders whether philosophers are protesting too much—shouting optimistic words into an uncaring void in a whistling-past-the-graveyard way—or whether there's a bias at work in which pessimistic philosophers fail to find followers (or publishers) in favor of those who hold out false hopes.

Pessimism considered harmful

Pessimism has become medicalized. Humoral theory held that pessimism was associated with an excess of black bile (from which we get the word "melancholy"); optimism with blood (from which we get the word "sanguine"). Nowadays, excessive pessimism and hopelessness are considered to be pathological symptoms of a medical diagnosis of depression (though not sufficient in themselves for such a diagnosis). If you come to believe that the universe is a cold, uncaring place, that all human effort is pointless vanity as all we come to love in our futile attempts to muffle our howling loneliness decays and is ripped from our fingers by ever-looming death, which is usually preceded by decrepitude, confusion, and painful suffering... you may be asked to consider one of an array of medications that shows promise for treating people who find such concerns to be of urgent import. A certain amount of pessimism is tolerated in the eccentric cranks among us, but above a certain threshold it becomes suspiciously pathological.

Optimism on the other hand is associated with better life outcomes. There's an obvious correlation/causation problem to be cautious about here, but there has also been a lot of research done that attempts to tease out a possible causal role for optimism. Optimists tend to use engagement-based (as opposed to disengagement-based) coping mechanisms to deal with adversity (addressing vs. avoiding the problem). For example, they take proactive steps to protect their health and wind up healthier as a result. They seem to do better in terms of

educational persistence, relationships, and income. A good overview of the research on benefits and drawbacks of optimism can be found in [Charles S. Carver, Michael F. Scheier, Suzanne C. Segerstrom “Optimism” *Clinical Psychology Review* 30 \(2010\).](#)

Trust

“What loneliness is more lonely than distrust?” —George Eliot[\[11\]](#)

“Suspicion often creates what it suspects.” —C.S. Lewis[\[12\]](#)

“Is not he a man of real worth who does not anticipate deceit nor imagine that people will doubt his word; and yet who has immediate perception thereof when present?”
—Confucius[\[13\]](#)

You can extend trust to someone particularly or generally. That is to say, you can trust them to do something (water your plants while you're on vacation) or trust them more generally (you consider them a trustworthy person). You can trust someone to do what they say they'll do, to follow-through, to make good on their promises. You can also trust someone *with* something, that is, you trust that they will be a good caretaker of it. You can also trust someone in the sense that you believe that they are being truthful in what they say. You can also trust someone's judgement: you trust them to make good decisions; you feel comfortable leaving things in their hands.

You might trust someone by virtue of the role they play in your life. For instance you might trust someone as-a-friend to follow through on your expectations of what friendship entails (e.g. keep secrets). Or you might trust someone as-a-fellow-X (Mason, proletarian, Dodgers fan) to show certain signs of solidarity (e.g. never testify against a fellow-cop). If trust of these sorts is betrayed, this can undermine the pillars of the preexisting relationship (e.g. “I thought you were my friend!”).

Trust might be extended in reciprocal expectation: I behaved in a trustworthy way towards you, so now I can expect that you will honor my trust in return. Two people may work together to establish a relationship of trust by alternately extending trust and honoring trust. There's a sort of prisoners' dilemma at work here, where each party is most apt to flourish by living a life in which they have relationships of trust, but each relationship of trust requires that they be willing to extend trust in an act that may make them vulnerable.

Trust towards people in general, or by default

You can also have a more or less trusting outlook towards people in general: You might come to assume that most people have more-or-less good intentions, or on the contrary you might come to assume that everyone is ultimately selfish and has ulterior motives for their ostensibly benevolent actions. Excess distrust of that kind can be an overcorrection in the face of bad experiences. The saga of “learning to trust again” when you've been hurt in romance is a common human tale. Distrust also seems very contextual: healthy distrust in one context would be an unhealthy level of suspicion in another.

People who abuse trust can sometimes be very crafty about doing so, and difficult to detect. Learning to detect untrustworthy people, or filtering them out of your life, can help you maintain a more trusting attitude in general. This way, you do not need to distrust people in general in order to avoid harm from those people who would abuse your trust. This is a difficult skill, and more of an art than a science, but is an important facet of trust-as-a-virtue.

Humans are so extraordinarily helpless as infants and children, and as a result are so reliant on others, that trust relationships are crucial parts of our early lives. Someone who grew up with abusive or neglectful caregivers may have extraordinary difficulty with trust because they lacked the opportunity to place trust in someone trustworthy as a child.

Sometimes people use “trust” to describe their relationships with institutions or impersonal things. Sometimes this seems mostly metaphorical: trusting gravity to keep your belongings from floating away; trusting that the sun will come up tomorrow. But in the case of human institutions, it’s more of an open question whether “trust” is a metaphor or really is something we can extend to collective endeavors or human-created algorithms: For instance, someone might say that they trust what they read on Wikipedia, or they trust Google to keep their email private, or they trust “the science” about global warming. They might feel betrayed if they trusted their software to keep reliable backups and they find out those backups were corrupted.

Trust seems to be one of those “golden mean” virtues. If you have too little trust, you’ll miss out on important opportunities for cooperation. If you have too much, you’ll be a gullible mark who is easily taken advantage of.

Trust vs. expectation and reliance

Trust is more than mere expectation. You might expect that someone will act in some way because it would be in character for them, or because it would be in their best interests. But if you *trust* them to act in some way, it seems to imply that you expect them to do so in part because of the trusting relationship you have with them: you trust them to act in such a way because you believe they consider themselves duty-bound to do so — perhaps because of the trust you have extended to them. (And an attitude of [“trust, but verify”](#) is hardly one of trust at all.)

This also suggests that in order to trust some person to do something, that person needs to be aware that they *are* trusted to do that thing, or at least they need to believe that doing that thing is part of what makes them trustworthy. You can’t, in other words, reasonably trust someone to do something they wouldn’t otherwise do, without bothering to communicate your trust to them somehow. If someone betrays your trust, this probably implies that you find them blameworthy for doing so; if so, then it doesn’t make sense to put your trust in someone in such a way that they would be blameless for betraying it (for instance if they didn’t have any way of knowing they were being trusted, or if you trusted them to do something you knew to be impossible, or if they did not accept the trust you offered). You can hope that your doctor cures your cancer; you can trust your doctor to take all reasonable and necessary steps to treat your cancer; but you can’t reasonably *trust* your doctor to cure your cancer, because it might not be curable.

Not all trust is explicitly granted and accepted, however. Some is implicit. You may extend a certain amount and type of trust to a stranger implicitly. For instance if you stop to ask someone directions, you implicitly trust them not to mislead you. You would be reasonable to feel that they betrayed your trust if they instead sent you on a wild goose chase, even if they never explicitly accepted your trust or never promised not to mislead you. The amount and type of trust that you are willing to implicitly extend in this way is one measurement of your expectations for people. Extending more such trust means you have higher expectations for them (and also, by extension, for yourself). This can be a way of promoting other virtues by setting-the-bar (to trust people to be honest, kind, etc. is also a way of saying I expect people to be honest, kind, etc. and I will judge them poorly if they are not).

Different cultures and subcultures have different expectations of trust: to whom it can be safely extended, in what areas, and to what extent. Violating the trustworthiness expectations of a culture can provoke a sort of exile from that culture in the form of shunning or shaming. Not extending trust to someone in a context where there is a cultural expectation of trust can be seen as insulting. Inter-culture conflicts are sometimes rooted in different trust expectations. Being “multilingual” in your sense of trust can be valuable, particularly if you straddle cultures. You are not merely a passive interpreter of your culture’s trust norms, of course: you also help to shape them. By being somewhat-more-trusting than the norm, you can help move the norm in a more trusting direction.

In my discussion of the virtue of honesty I mentioned Thoreau's observation that "[It takes two to speak the truth—one to speak, and another to hear.](#)"^[14] If you cannot trust someone to be honest with you, you become deaf to honesty: it becomes more difficult for you to hear the truth. If someone knows that you do not trust them, this may also demotivate them from attending carefully to the the truth, precision, and clarity of what they say to you ("why bother, after all you're just going to believe what you want to believe").

Trust and vulnerability

Trust is closely related to vulnerability. When you extend trust to someone, you may be making yourself doubly-vulnerable: the person you trust may not do what you trust them to do (which may have negative consequences for you, e.g. your plants go unwatered and die), and they may betray your trust (which is an additional negative consequence, e.g. you feel a fool for having trusted them). Being vulnerable in this way can require courage.

Extending trust, and accepting the cost in vulnerability to do so, can be a sort of generosity or charity, in some contexts. (If you extend different amounts of trust to people for unfair reasons, this may have implications for [justice as a virtue](#) as well.) Because extending trust makes you vulnerable, it is also a way of demonstrating strength. In other words, if you extend trust, you communicate by doing so that you feel you are strong enough to risk a certain amount of vulnerability.

How can we improve in hope, optimism, and trust

Most of the advice I have found about how to improve in these virtues has to do with how to be more hopeful, more optimistic, and more trusting. In other words, it assumes that you have undershot the mark. Advice is harder to come by on how to [stop clinging to exhausted hopes](#), accept that life is vain suffering, and stop being so darned gullible.

Benjamin Franklin noted that some people tend to dwell on the negative things in life, others on the positive, and the former are at a great disadvantage in life. He asserted that the pessimistic outlook is a "disposition... perhaps taken up originally by imitation, and unawares grown into a habit, which though at present strong, may nevertheless be cured."^[15] His cure involves deliberately adopting the habit of redirecting one's attention away from the negative things in life. This suggests that [gratitude and appreciation](#) can be ways of becoming more optimistic.

[Cognitive-behavioral therapy](#) is a modern approach to Franklin's cure. It is used to combat the persistent pessimistic thoughts that are symptoms of (and may exacerbate) depression, and the cognitive biases that reinforce pessimistic assessments.

Positive psychology researcher Martin Seligman suggested that there is a "[learned optimism](#)" that people can cultivate.^[16] You become more optimistic by changing how you describe things that happen in your life. If you describe good things using language that implies that they are "permanent and pervasive," and bad things with language that implies that they are "temporary and narrowly-focused" an optimistic outlook will result. (And a pessimistic one will result if you get this backwards.) So, for example, rather than describing a negative occurrence as an example of "my bad luck" you describe it as just a one-time setback. This is a sort of ["framing"](#) technique used as a mind hack.



Agnes Callard
@AgnesCallard

...
People are at their most optimistic when answering the question, “what practice now will seem morally reprehensible in the future?”
5:49 AM · Dec 18, 2021 · Twitter for iPhone
<https://twitter.com/AgnesCallard/status/1472202296238952454>

1. [^](#)
Richard Rorty, *Contingency, Irony, and Solidarity* (1989)
2. [^](#)
C. Peterson & R.S. Vaidya “Optimism as Virtue and Vice,” in *Virtue, Vice, and Personality: The Complexity of Behavior* (2003) pp. 23-27
3. [^](#)
Christopher Peterson & Martin E.P. Seligman, *Character Strengths and Virtues* (2004), chapter 25
4. [^](#)
Joseph Delany, [“Hope” Catholic Encyclopedia](#) (1910), Vol. 7
5. [^](#)
Nancy Snow, [“Hope as an Intellectual Virtue”](#) *Virtues in Action: New Essays in Applied Virtue Ethics* (2013)
6. [^](#)
William James, “The Will to Believe” *New World* (1896)
7. [^](#)
Mark Neyer, “The Hypnotoad Pill” *Apxhard* (18 January 2021)
8. [^](#)
Fitz James Stephen, *Liberty, Equality, Fraternity* (1874)
9. [^](#)
“Meliorism is the belief that the specific conditions which exist at one moment, be they comparatively bad or comparatively good, in any event may be bettered.”
—John Dewey, *Reconstruction in Philosophy* (1919)
10. [^](#)

Friedrich Nietzsche, *Ecce Homo* (1908)

11. ^

George Eliot, *Middlemarch* (1872)

12. ^

C.S. Lewis, *The Screwtape Letters* (1942)

13. ^

Analects of Confucius, XIV.xxxiii

14. ^

H.D. Thoreau, *A Week on the Concord and Merrimack Rivers* (1849)

15. ^

Benjamin Franklin, ["The Deformed and Handsome Leg"](#) *The Columbian Magazine* (October, 1786)

16. ^

Martin Seligman, *Learned Optimism: How to Change Your Mind and Your Life* (1990)

Notes on Good Temper

This post examines the virtue of **good temper**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is good temper?

“To bear trials with a calm mind
robs misfortune of its strength and burden.”
—Seneca^[1]

Good temper is the proper regulation of anger in particular. It may be best described by contrasting it with its opposite, which goes under names like “volatility,” “fury,” “hot-headedness,” “seething,” “flying off the handle,” and such. If you suffer from any of that, then even if you are generally level-headed, when you are provoked to anger you no longer make wise and sensible [decisions](#) but instead act rashly in ways that you would otherwise avoid.

If you have good temper, you keep your wits about you even when you are angry or have cause to be angry, and the quality of your decisions does not suffer as a result.

(Something about on-line interactions seems to easily provoke “flame wars” and other counterproductive displays of anger. Internet trolls play social media like a video game in which they score points by provoking other people into pointless displays of rage. So we would be wise to be especially on our guard about letting our anger interfere with the wisdom of how we behave on-line.)

Related virtues

Good temper is related to *poise*, in the sense of being unflappable. Someone with good temper is well-composed, stays cool, rolls with the punches, and doesn't get thrown off their game.

Sometimes good temper is subsumed under [temperance](#) and [self control](#) — in such a scheme the temperate person is not prone to extremes of anger, and the person with self control does not let anger run away with them even if it does rise to extremes.

The pros and cons of anger

“Rightly to be great
is not to stir without great argument
but greatly to find quarrel in a straw
when honour's at the stake.”
—Hamlet^[2]

There are disagreements over whether anger itself is a useful thing or whether it should be suppressed. Some philosophers recommend cultivating a stance of equanimity and serenity, and treating others' faults with charity and forbearance, and so they see anger as a pathology. On the other hand, [righteous anger or indignation](#) can help prompt your concern for [justice](#), and so a lack of anger might indicate that you are unhealthily indifferent to injustice, or perhaps a push-over who is vulnerable to being taken advantage of.

Edwin Abbott (of *Flatland* fame) wrote: "There is an anger that is always right, such as one feels at the sight of cruelty, injustice, and oppression, a moral recoil of sentiment from evil... Resentment then is a Virtue, and a man who feels no resentment at the sight of injustice is destitute of a true sense of sin. There is almost as great a deficiency of resentment in the world as there is an excess of vindictiveness."^[3]

However, Seneca, who wrote a book on anger,^[4] thought anger didn't have a place in justice: "It is not for the dignity of a judge, when he comes to pronounce the fatal sentence, to express any emotions of anger... for he condemns the vice, not the man... nor is there any need of an angry magistrate for the punishment of foolish and wicked men." He thought that philosophers had tried anger and found it useless: "Democritus laughed, and Heraclitus wept, at the folly and wickedness of the world, but we never read of an angry philosopher."

Aristotle went for the middle ground (of course), finding a golden mean of good temper in between the opposite extremes of wrathfulness and indifference, though he believed that people more often err on the wrathful side and that this is the more harmful of the two extremes.^[5]

These days I can imagine a good argument for the other side: that a "deficiency of resentment" (as Abbott put it) is more common and more corrosive. There is a lot of stupid belligerence around, to be sure, but on the other hand, I'm amazed at the variety and severity of insults to [dignity](#) that people routinely put up with without complaint. The commonplace mendacity of politicians or advertisers is such a variety of insult — accepting it without complaint has become so typical that it looks eccentric to behave as though one took offense at being lied to or treated like an idiot.

The Stoic approach: do not get angry

"How much more harmful are the consequences of anger and grief than the circumstances that aroused them in us!" —Marcus Aurelius^[6]

The Stoics believed that anger was unpleasant, unhealthy, undignified, uncivil, unhelpful, and unnecessary, and that with training in philosophy you could rid yourself of it. Anger, according to the Stoics, is caused when you condition your peace of mind on something that is not in your control turning out a certain way, and it doesn't cooperate in turning out that way.^[7] For instance, you get angry because someone else divulged a secret or parked in your parking place, or because the store was out of your favorite brand of coffee or the weather spoiled your picnic.

The solution is to stop conditioning your peace of mind on things that are out of your control. Those things are just the background against which you live your life; let them do what they will, and concentrate instead on the things that *are* in your control (for instance your attitudes and choices).

If you go along with this prescription, the Stoic philosophers have some sensible advice to go along with it. For example, this from Epictetus: “Start with small things. For example, you have spilled something on the carpet or something small is stolen from you. [Instead of getting angry] Say to yourself, ‘This is such a small price to pay for tranquility and peace of mind.’”

Or this, from Marcus Aurelius: “Say to yourself in the early morning: today I shall meet meddling, ungrateful, violent, treacherous, envious, and uncharitable men.”^[8] (That way you will be prepared to meet them gracefully, and when you do meet them, you won’t get bent out of shape, but you’ll just say “aha, there you are; I was expecting you.”)

“The greatest remedy for anger is delay,” says Seneca. “Beg anger to grant you this at the first, not in order that it may pardon the offense, but that it may form a right judgment about it: — if it delays, it will come to an end.”^[4]

Some free Stoic ebooks:

- Cicero, [*Tusculan Disputations*](#)
- Epictetus, [*The Enchiridion*](#) and [*Discourses*](#)
- Marcus Aurelius, [*Meditations*](#)
- Seneca, [*Dialogues*](#) (includes *De Ira*)

Seneca suggests that anger be approached in two ways: by trying to avoid becoming angry in the first place, and by avoiding doing wrong if one nonetheless becomes angry. The second of these ways is the modern discipline of “anger management.”

The anger management approach: be angry skillfully

I am encouraged to see that for this virtue at least there is a scientific discipline devoted to it. Anger management is more of a psychological than philosophical discipline. Because anger is at the root of a lot of violent crime, professional anger management counselors are now deployed in the service of criminal rehabilitation and crime prevention. Because of this, the main focus of anger management is on helping people who have excessive anger and/or who behave unwisely under its influence (not, in other words, people who have insufficient anger or who suffer quietly while nursing a grudge).

The anger management discipline typically disagrees with the Stoics about whether anger is something that can be or should be entirely avoided. It sees anger as a natural and unavoidable emotion, and tries to teach ways to react to the arising of that emotion that are relatively harmless.

Anger management is a complex subject: There are a variety of possible causes for why people react to anger poorly, a variety of ways in which such people misfire when angry, and a variety of techniques deployed by anger management counselors to try to improve matters. I don’t have any particular expertise here, so I’ll just [link to Wikipedia and be done with it](#).

A web search for the term “anger management” will lead you to a multitude of sites with tips for the layman. Seneca’s advice of *delay* seems to still be very popular:

count to ten, take a deep breath, etc. Dispelling the adrenaline of anger by means of innocuous physical activity like exercise is another popular gambit.

Anger is apparently physiologically similar to anxiety and fear.^[9] These share at their root a feeling of being out of control and an urge to regain control, and the physiological promptings are those that make you more alert and eager to take physical action. Displays of rage can be ways of trying to manage an out-of-control situation: You make everyone else freeze in place as your histrionics take center stage, and in this way you exert some temporary control. This suggests that the more chaotic a situation appears to you, the more likely you will go into rage mode to try to confine it. Noise, crowds, and the presence of other stressors can exacerbate this. So there may be some environmental changes that you can make that will make it easier for you to avoid extremes of anger.

Holding grudges

Most of what I've found about anger and about responding to anger skillfully has to do with acute anger: the immediate insult, the sudden rush of blood to the face, the struggle to resist the urge to flail out without thinking. But there is also the phenomenon of the simmering grudge, the gnawing complaint, the all-consuming vindictive crusade. If you have been treated unjustly and you never feel like this was addressed properly — you never got justice, or revenge, or vindication — you may have a hard time letting go. It may feel like unfinished business and may continue to bother you long after there is practically anything to be done about it.

If you've ever dragged around the rotting carcass of a grievance like this, you'll know how unpleasant it is and how little good it does. Over time the harm caused by the original insult can be dwarfed by the harm caused by how rotten you feel every time you return to dwell on it. But in spite of this it can be frustratingly difficult to just drop it and move on.

The Stoics would argue that you are making the classic mistake: conditioning your peace of mind on things outside of your control. You cannot force the person who wronged you to become contrite and offer restitution, or for the insult to be undone, or for the world to admit that you were wronged. You feel terrible because you are out of control; you have no influence over these things that you are allowing to tug your emotions this way and that. What you *can* control are your own attitudes and choices, and if you want to feel better, you need to exercise that control.

Well, thank you Stoics, but all that is more easily said than done. Once you have developed the habit of chewing on the bone of an old complaint, it can take a lot of persistent effort to break that habit.

For an approach that is less philosophical and more methodical, rumination-focused cognitive behavioral therapy (CBT) looks promising. That approach involves carefully investigating the triggers that reevoke the grudge, and deliberately working to replace your grudge-reinforcing responses to those triggers with more consciously-chosen and rational ones until this becomes habitual. There are therapists who specialize in CBT who can help walk you through the process.

The classic cure for holding a grudge is [forgiveness](#), though this can also be easier said than done, and may not always be appropriate.

1. [^](#)
Lucius Annaeus Seneca (the Younger), *Hercules Oetaeus*
2. [^](#)
William Shakespeare, *The Tragedy of Hamlet, Prince of Denmark* (~1600), act IV, scene 4
3. [^](#)
Edwin Abbott, *Bible Lessons* (1870)
4. [^](#)
Lucius Annaeus Seneca (the Younger), *De Ira*
5. [^](#)
Aristotle, *Nicomachean Ethics* IV.5
6. [^](#)
Marcus Aurelius, *Meditations* I.18
7. [^](#)
See also: Duncan A Sabien, ["Anger as evidence"](#) Jan 18, 2020
8. [^](#)
Marcus Aurelius, *Meditations* II.1
9. [^](#)
David Hanscom, ["The Connection Between Anxiety, Anger, and Adrenaline"](#)
Psychology Today (15 December 2019)

Notes on Fairness

This post examines the virtue of **fairness**. It is meant mostly as an exploration of what others have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

Fairness may seem like a things-I-learned-in-kindergarten virtue, but experimental evidence shows that it can be difficult for mature adults to wrap their heads around. To summarize: you can set up an experiment in such a way that a subject declares that being fair is very important, expresses that A is more fair than B, does B anyway, and then congratulates themselves for their fairness, all seemingly with eyes wide open. More about that later...

What is fairness?

I struggled to define fairness. It seems to be different things, or at least to emphasize different things, in different contexts. (See also Eliezer Yudkowsky's [The Bedrock of Fairness](#) and [Is Fairness Arbitrary](#), and the comments to those.) I'll stick to being descriptive rather than prescriptive:

One aspect of fairness is impartiality. A process tainted by favoritism, self-dealing, invidious discrimination, or nepotism is unfair. (Why am I still wearing these rags while my wicked stepsisters get all gussied up for the ball?)

A [decision](#) can be unfair if it was based on things that are not justly relevant, for instance if prejudice was involved. A decision can also be unfair if it is not based on anything at all: if it is arbitrary.

Fairness can be connected to merit. It is fair to give the trophy to the winner. It would be unfair to give it to the second-place finisher instead. (Such a travesty would probably be a case of prejudice, partiality, or arbitrariness, so maybe it's subsumed by the previous cases.)

That said, the trophy might fairly be withheld from the winner if it was not won by means of fair play. Fair in that sense means by-the-rules (or sometimes by the unwritten rules of sportsmanship). Being unjust in general is sometimes also called unfair: it's unfair to go back on your word, or to counterfeit money, or to defraud your customers, for example.

You are supposed to be best able to make fair decisions if you can adopt an objective and impartial perspective, unaffected by irrelevancies. Political philosopher John Rawls created a thought experiment to this end with his "[veil of ignorance](#)" behind which you could make your decisions about what would be fair.^[1] As a simplified example, imagine that there is going to be a robbery but you don't know whether you are going to be the robber or the person being robbed (because you are behind the veil of ignorance). You can decide: would you rather have this take place under a system in which robbers are caught and forced to relinquish their ill-gotten gains to their victims, or one in which the robbers get away with it and can brazenly keep what they have

stolen, or some other alternative? Because you do not know whether you would personally gain or lose by your decision, you can decide based on fairness rather than on partiality.

How is fairness to be distinguished from [justice](#)? Is fairness just a component of justice? or maybe a primitive form of justice? Can you be unfair without also being unjust? Maybe so. For example: If I'm handing out bonuses at the end of the year, and I give higher bonuses to men than to women, or to relatives than to non-relatives, I'm arguably being unfair even if justice did not obligate me to give bonuses of any sort to anyone.

Fairness can be in tension with the virtue of [loyalty](#), which can come packaged with an expectation of partiality.

Cake-cutting algorithms

One way that fairness has been studied rigorously has been through cake-cutting algorithms. These involve dividing a cake among multiple people. The trick is to prove whether there is some algorithm by which a cake can be divided by the people involved such that some criterion of fairness is respected. There are various criteria of fairness that you might choose, such as:

1. *equality*: everybody gets a slice of cake with the same value (size, in a simple model)
2. *proportionality*: everybody gets a slice of cake that they value at least as much as the slice they would have gotten if the cake had been equally divided among everybody by some omniscient authority
3. *contentment*: nobody has reason to wish they had someone else's slice of cake instead of their own

A simple case is dividing a cake between two people. If person A slices the cake in two, and person B makes the selection of which slice goes to which person, this incentivizes A to slice the cake such that they would not prefer either slice over the other, as presumably if one slice is better, B will choose it. This algorithm guarantees proportionality and contentment. Note that this works even if one person would most prefer the bigger slice of cake, and the other person the slice with more cherries on top. They don't have to use the same criterion for valuing the slices. Proportionality- and contentment-guaranteeing algorithms of this sort have also been proven for dividing the cake among *any* number of people. [\[2\]](#)

For something more practical than cake-dividing, imagine this scenario: A set of roommates is moving into a rental house. They need to decide how to divide up the rent, and who gets which room. The rooms are different: some are larger than others, have better window views, are nearer or farther from the noisy neighbor, etc. Is there a fair way to distribute the rooms and divide up the rent? Yes: it's an application of Sperner's Lemma and it comes out of the same branch of mathematics as the cake-dividing stuff. [\[3\]](#)

It's nice to know that at least in some cases, you don't have to eyeball it, but instead there is a proven method for arriving at a fair result — at least among people who can agree on which criterion of fairness to use.

Ultimatum and dictator games

An ultimatum game is a variety of game theory scenario. In its basic form (there are many variations), player A is given money to split with the other player, B. A decides how much of the money each player gets. B can decide either to accept the portion A has granted them, or to reject it, in which case both player A and player B get nothing at all. (Dictator games are severe variants of the ultimatum game in which B is reduced to helplessness. Player A divides the pot, and player B gets what proportion player A decides to give them, without any opportunity to reject this.)

Naively, it is always in B's utility-maximizing interest to accept any non-zero portion, as this is better than the zero portion B would get by rejecting it. And correspondingly it is always in A's utility-maximizing interest to offer B a tiny portion. (If the game is repeated, or under certain other assumptions, this calculus can change.) However, experimentally, people do not utility-maximize in this way: It is most common for A to offer an even 50/50 split, and B will often reject an offer if it is low but non-zero. This may suggest that people are biased towards fairness at the expense of naive utility-maximization. (In other games, people will go out of their way to *punish* unfairness in ways that also would not be predicted from naive utility-maximization; see [this example](#).)

However, a number of modifications to the experiment have been tried that change the results in revealing ways. [I get nearly ten thousand results for “ultimatum games” on Google Scholar](#), and [ten thousand more for “dictator games”](#). Researchers have delved into all sorts of interpretations of these games, and have designed a multitude of variants meant to clarify various nuances.

For the purposes of this discussion of fairness-as-a-virtue I'm most interested in the results from the varieties of the dictator game that were developed by [Daniel Batson](#) and extended by those who followed his lead. There's [a good summary](#) of these at the [Stanford Encyclopedia of Philosophy](#) entry on [“Distributive Justice and Empirical Moral Psychology.”](#)

Participants in these experiments had the dictator-game-like role of assigning themselves and another subject to two tasks, one of which was described as being clearly more favorable than the other. “After making the assignment privately and anonymously, participants were asked about what was the morally right way to assign the task consequences, and to rate on a 9-point scale whether they thought the way they had actually made the task assignment was morally right.” Almost none of the participants said that simply assigning the better of the tasks to themselves was the morally right thing to do, but a large majority of them did assign the tasks such that they got the better of the two.

In another variation, the participants were pointedly given a coin that they could flip to make the task assignment randomly should they so choose. 70% of participants agreed that assigning the tasks by using the results of a coin flip was the correct thing to do. Half of the participants followed through on this and actually flipped the coin. Of those who did not flip the coin, 90% gave the favorable task to themselves. Of those who *did* flip the coin, 90% gave the favorable task to themselves as well. When they were later asked to rate the fairness of their decision-making process, those who had flipped the coin rated themselves as having been significantly more fair than those who hadn't.

This (and other permutations of Batson-style games) suggests to me that people are very vulnerable to self-perceptions of fairness that do not match our actual behavior, and that inclinations to fairness are not very strong or widely-held. If I want to have the habitual characteristic — the virtue — of fairness, therefore, I should expect that I will have to be extraordinarily vigilant and skeptical about my behavior.

Political fairness

Political systems, whether or not they like to admit it, depend on the consent of the governed. If enough of the governed come to feel that the system is unfair, particularly if they are on the raw end of the deal, this is a threat to political stability. Those who have a stake in the political status quo will therefore try to see to it that the state is portrayed as a fair one, and may go so far as to support reforms that make the system more fair.

How a political system can demonstrate that it is a fair one, however, is a matter of much debate and little consensus.

“Equality” is a perennially popular variety of political fairness, though equality-how-exactly can be a wiggly issue. Equality before the law (“The law, in its majestic equality, forbids rich and poor alike to sleep under bridges, to beg in the streets, and to steal loaves of bread.” —Anatole France^[4])? Abolishing officially-enforced castes? Distributing equal ownership shares of everything?

“From each according to their ability, to each according to their needs.” is an aphoristic formula that caught on in some circles.

John Rawls made a name for himself with an ingenious defense of a sort of maximin version of fairness.^[1] An outcome is the fairest one, he thought, if in no other outcome would the worst-off people in it be better off than the worst-off people in this one.

Libertarians tend to prefer Robert Nozick’s rejoinder.^[5] In his reckoning, if you start with a fair system, *any* other system that can be reached from that starting point through intermediate steps that are themselves fair (no force or fraud was involved, for instance) is also fair. In other words: Fairness is not so much about where you’re at, but how you got there.

I race through these shameless oversimplifications of political philosophy just to give some idea of the breadth of notions about what’s “fair” that are out there. We do not agree about what is fair. We don’t even agree about what fair is. It is enough to make one suspect that highfalutin theories of distributive justice have mostly to do with coming up with impressive reasons why the outcome you would find preferable is also the fair one.

“So convenient a thing to be a reasonable creature, since it enables one to find or make a reason for every thing one has a mind to do.” —Ben Franklin^[6]

How to develop fairness

I didn’t find much about how to become more fair. There is an intervention called “transactive discussion” in which two people with slightly different ideas of fairness discuss a moral dilemma together (e.g. the Heinz dilemma). In the results of some

studies, this appears to help the person with a “lower” level of fairness sophistication raise the quality of their fairness evaluation in a measurable and lasting way.^[7] So it might be helpful to discuss moral dilemmas (or just ordinary moral quandaries) with others.

1. ^

John Rawls, [A Theory of Justice](#) (1971)

2. ^

Ariel D. Procaccia, [“Cake Cutting Algorithms” Handbook of Computational Social Choice](#) (2016)

3. ^

Albert Sun, [“To Divide the Rent, Start With a Triangle” New York Times](#) (28 April 2014)

4. ^

Anatole France, *The Red Lily* (1894)

5. ^

Robert Nozick, [Anarchy, State, and Utopia](#) (1974)

6. ^

Ben Franklin, *Autobiography* (1791)

7. ^

Marvin W. Berkowitz, “The Role of Transactional Discussion in Moral Development” *Sum* (1980)

Marvin W. Berkowitz & John Gibbs, “Measuring the Developmental Features of Moral Discussion” *Merrill-Palmer Quarterly* (1983)

Notes on Endurance

This post examines the virtue of **endurance**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is endurance?

Endurance is the virtue exhibited by someone who characteristically puts up with adversity, pain, discomfort, hardship, or suffering in a way that allows them to maintain their poise. To be endurance-as-a-virtue, you must be enduring because you have the confidence that you *can* endure; in other words, you aren't just hopelessly enduring because you have no choice but to endure.

Related virtues

There are several other virtues that are closely related to endurance or that are dependent on it:

- A near-synonym for endurance is **fortitude**.
- When you both endure and do not let whatever it is you are enduring distract you from your worthy projects, you exhibit **persistence, perseverance, steadfastness, or tenacity**. It's also trendy these days to use the words **grit or sisu** to describe this.
- Being resistant to the negative effects of adversity (to the extent where maybe you do not even feel you have to endure anything) is **hardiness** or **toughness**.
- The ability to recover well from adversity is **resilience**.
- When you both endure and don't get bent about whatever it is you're enduring, you're also exhibiting **forgiveness or patience**.
- In the context of physical exertion, endurance is sometimes called **stamina** (sometimes this term is also used analogously in terms like "mental stamina").
- To endure something it helps to have **hope/optimism** that it *can* be endured.

Aristotle pointed out parallels between endurance and **self control**: Self-control is resisting the temptation of things that seem immediately appealing ("I can do without that"); endurance is resisting the discouragement of things that seem immediately uncomfortable ("I can put up with that"). It can be difficult to distinguish them in some cases: is the regretful alcoholic reaching for the bottle because they cannot resist the temptation of a pleasing drink, or because they cannot endure the discomfort of withdrawal?

What good is it?

"The Romans 'made Fortune sirname to Fortitude,' for fortitude is that alchemy that turns all things to good fortune." —Thoreau^[1]

I've noticed that authors who praise the virtue of endurance often quickly pivot to describing the benefits of a related virtue like persistence or resilience. So it may be that what is best about endurance is that it is an essential ingredient of these other virtues. That doesn't mean it is not worth considering on its own, as there may be useful ways of analyzing or of strengthening endurance in isolation.

I'll try to stick with discussing endurance itself in this post, and save discussion of related virtues for later.

Endurance is often contrasted to things like "giving up," "defeat," "surrender." This characterizes life's various struggles as *battles*, adversity as the *opponent*, and makes endurance a crucial ingredient in *victory*. Simone de Beauvoir wrote that "the characteristic feature of all ethics is to consider human life as a game that can be won or lost and to teach man the means of winning."^[2] If you consider life in this way, endurance is one of the keys to winning the game.

Endurance and gender

In the culture I grew up in and am most familiar with, endurance seems to be often represented as a particularly masculine virtue, or at least a virtue whose absence is particularly unattractive in men. Men who lack endurance are caricatured as small and scrawny, with high-pitched whine, a hand-wringing lack of confidence, and a stooping lack of dignity: in short, as a sort of un-matured male, still dragging his blankie behind him into adulthood. Epithets like nebbish, pansy, or wimp usually imply both a lack of endurance and that the person so characterized is failing at being a man.

In women, endurance or the lack thereof does not seem to register so strongly as part of their gender identity, and there isn't such a rich set of epithets to describe women without endurance. There are caricatures of women who, for example, faint away at the first sign of adversity, or who simply cannot go on because of some stereotypically feminine need (my nails! my hair! my heels!), but these usually do not imply a *failed* femininity but merely a delicate *variety* of femininity; and conversely a woman who *does* demonstrate great endurance is not violating gender norms by doing so but is usually portrayed as being admirable and not necessarily any less feminine (though sometimes men around her will have their masculinity called into question by comparison).

How to improve endurance

In the case of physical stamina, the path to endurance is pretty clear: repeated exercise in which you approach the limits of your endurance causes gradual physical changes that have the effect of expanding the limits of your endurance.

For other forms of endurance, something similar seems to happen, though the mechanism is probably different. If you are not confident you can endure something, you may give up early. If you persist (or if you don't have the choice not to) and so you do endure something you were not confident you could endure, you learn that your limits were greater than you expected, and so in the future when you encounter similar situations you may have more confidence, and therefore more endurance when tempted to give up.

So one way to build endurance may be to “fake it ‘til you make it” — behave as though you have endurance even if you don’t, in order to prove that you *can* endure and build the confidence that endurance requires.

It is possible that what you think of as your lack of endurance is really masking a failure in another virtue, and so you can improve your endurance by working on that virtue instead. For example, if your difficulty enduring typically shows itself in the face of *frightening* circumstances, maybe it’s really *courage* and not endurance that you’re lacking. If you cannot endure waiting, maybe it’s really patience you need. If you cannot endure loneliness, more self-reliance might come in handy. Some introspection at the point where your endurance seems to be giving out may be key here (“what exactly is the terrible consequence I expect will happen if I persist beyond this point?”).

It is easier to endure adversity if you have more resources in general. So for instance, if you have a strong social network, good money management understanding, confident interview skills, and so forth, it can be easier to endure a bout of unemployment simply because it is less daunting than it would be for someone without those resources. Maybe endurance is best thought of, then, not as a single, isolatable virtue, but more as a wisely-developed constellation of resources that can be drawn on in times of need, along with the skill to deploy them well.

It may also be helpful to familiarize yourself with the stories of other people who have gone through hardships and come out the other side thanks to the endurance they expressed along the way. There are many evocative and influential fictional stories along these lines (e.g. *The Old Man and the Sea*, the *Odyssey*), but also plenty of based-on-real-life stories. As you become more aware of just how much people *have* endured, you may have more confidence about what you personally *can* endure.

1. ^

H.D. Thoreau, [*The Service*](#) (1902)

2. ^

Simone de Beauvoir, *The Ethics of Ambiguity* (1947)

Notes on Benevolence

This post examines the virtue of **benevolence**. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is benevolence?

Benevolence (sometimes “goodness” or “goodwill”) is the belief that some things are good, others bad, accompanied by a determination to promote the good and discourage the bad. In short: if you are benevolent you commit yourself on the side of good. Most typically, benevolence concerns *moral* good and bad — being good to others, in particular — rather than less morally-tinted goods like cleanliness or virtuousness.

A brief definition like this hides a lot of complexity and difficulty. How can it be the case that some things are morally good and other things morally bad, for example, and how can we know the difference? Why is there debate about what is morally good or bad, and why does there seem to be no consensus about how such debates can be resolved through appeals to facts and logic? Why ought one choose to side with good over bad instead of the alternative or instead of remaining neutral and making your decisions based on some other metric entirely?

The benevolent person doesn’t necessarily think that they have all the answers to these questions, but is satisfied that moral right and wrong is real and worthy of respect, believes they can discern the rightness or wrongness of a thing or a course of action well enough to act on that discrimination, and is persuaded that consistently siding with the good is the correct course of action.

Ethical philosophy is in part a gigantic junkyard strewn with the wreckage of ingenious attempts to prove or disprove any of that. If you want a philosophical vehicle with which you can rationally drive yourself either into or out from a stance of benevolence, you can take your pick, but good luck getting it off the lot.

Benevolence and the virtues

From a virtue ethics perspective, the question is whether benevolence is an essential ingredient in human flourishing, or whether one can flourish just as well as an evil or morally indifferent person.

Most virtue traditions at least imply that benevolence is correct. Some virtues are morally neutral: both good people and evil people can thrive better with a good helping of fortitude, fitness, boldness, discipline, confidence, and so forth. But many virtues imply taking moral sides. What makes sympathy better than *schadenfreude*? concern better than disdain? gentleness better than cruelty? [honesty](#) better than deceit? sportsmanship better than ruthlessness? If you decide not to be benevolent, you almost need to find a mirror-world set of [Machiavellian](#) virtues in order to learn how to thrive.

“Good” certainly won the public relations war when it chose that name. People more often choose the side of good, or claim to, anyway. The Marist Institute for Public Opinion does a poll about New Years Resolutions every year, and “be a better person” is usually near the top of the list, next to [resolutions](#) about exercising more and losing weight.

Sympathy for the devil

But we also tend to love a good villain. Shakespeare’s [Iago](#) and [Richard III](#) are among his most captivating characters, in large part because they are so thoroughly and energetically evil. They show us what an inverted set of malevolent virtues might look like, by showing us a character who has decided to try to flourish as a human being aligned with evil.

Adam Kotsko wrote a fascinating book several years back called [Why We Love Sociopaths \(excerpt\)](#) that examines why sociopathic characters had become so enormously popular on television — folks like [House](#), [Dexter](#), [Tony Soprano](#), [Jack Bauer](#), and [Don Draper](#) in TV dramas; cartoon characters like [Eric Cartman](#), who often steals the show on *South Park*; or the various ruthless and amoral reality TV competitors like those in [Survivor](#) or the U.S. Presidential Election — and what this collective pursuit of a fantasy of being sociopathic might mean about us and about the phase our society is going through. In short:

My hypothesis is that the sociopaths we watch on TV allow us to indulge in a kind of thought experiment, based on the question: “What if I really and truly did not give a fuck about anyone?” And the answer they provide? “Then I would be powerful and free.”^[1]

Can you flourish as a human being on the side of evil, or is that just a conceit of fiction? It certainly seems like evil people in real life are making their evil decisions under the theory that they will benefit by doing so. How would you argue against them?

Well, for one thing, evil deeds can have bad consequences for the person who does them. But this isn’t very persuasive. Evil people, if they’re not also stupid, have already figured that into their calculations. Avoiding detection and deflecting consequences are among the evil virtues. And it is an ancient complaint that, at least superficially, evil people often prosper while good people get trod underfoot.

We are all familiar with the eternal-soul gambit that sometimes gets trotted out in response to this. Certainly it *seems* like evil people prosper and good people suffer, but that’s only because you’re focused on the short term. If you look at things from the perspective of eternity, boy howdy do the good people make out like bandits while the bad people wail and gnash their teeth in fire & brimstone. Cool story bro.

Another tack is for good people to say that ungood people are being mistakenly short-sighted. They’re aiming at the sorts of petty rewards that even the evil can obtain, while missing out on the grander things in life that are only available to good people. For example, the good person may wax poetic about the joys of [love](#) being in large part the fulfilling feeling of *loving*, not the silly selfish pride of being adored. An evil person may be able to manipulate someone into giving them the latter, but will never be able to obtain the former.

But ungood people are likely to think that good people are just fooling themselves about that sort of thing: “It is better to give than receive? [More terrible to commit injustice than to suffer it?](#) Yeah, sure. Tell me another one. I’ve been down that road, and brother, let me tell you: nice guys finish last.”

More cynically, they assume that good people are just other evil people trying to manipulate each other. Like the child who told the other child: “You should learn to share. Do you know what sharing is? It’s when you have something that I want and you give it to me.”

When I did a quick Google search for “how to become good if you are evil” to see if there was something obvious I missed, perversely almost all of the results seemed instead to be about how good people turn evil. The upshot of all this is that I despair of coming up with either motivations or methods for *choosing* benevolence if you are not already so inclined.

But that isn’t even the end of the troubles. Consider the enormous differences of opinion about what good and evil consist of. Some radical thinkers ([Nietzsche!](#) [Rand!](#)) insist that a lot of the common-sense ideas we have inherited about good and evil have been so corrupted that they are almost the opposite of what they should be. So even if you make the leap of deciding to be benevolent, you’ve got your work cut out for you figuring out just what that means.

There are even those who insist that if you side with good over evil, you show that you are subscribing to a primitive, dualist [Manichæism](#), and dooming yourself never to integrate your shadow side and become one with the greater cosmic harmony in which evil plays its useful supporting role.

So where am I going with this?

“I have a belief of my own, and it comforts me... That by desiring what is perfectly good, even when we don’t quite know what it is and cannot do what we would, we are part of the divine power against evil—widening the skirts of light and making the struggle with darkness narrower.” —George Eliot^[2]

All of that was mostly to say that if there’s a failsafe way to *reason* someone into benevolence, I’m not familiar with it, and so you won’t find it here. And even if you don’t need convincing on that point, and you plan to take up your sword and shield for the good team, you then have the trouble of finding out how to *be* good and what *is* genuinely good, hidden as it is among the many counterfeits.

In the same spirit that the rationalist community adopts when they vow to become “less wrong,” I think it’s probably a better idea to try to be “less bad” than to declare yourself to be firmly, once and for all, on the side of good. This is in part because of the phenomenon of [moral licensing](#) — a cognitive bias in which if you do something good, or think of yourself as basically good, or even *imagine* yourself doing something good, you may give yourself a sort of “get out of jail free card” as a reward, and feel less guilt about indulging in something bad. Somewhat related is [noble cause corruption](#), in which if you’re convinced you’re on the side of the angels, you use that as a great excuse for your atrocities.

If you can avoid strapping on your paladin kit right away, and instead more humbly commit to becoming “less bad” you will be less likely to fall victim to these fallacies,

you'll be more investigative about how to be better, and you'll be less defensive when you learn about your shortcomings or have them pointed out to you.

I wasn't a big fan of [Robin DiAngelo's book *White Fragility: Why It's So Hard for White People to Talk About Racism*](#), but I thought its main argument was convincing: that when white Americans become defensive about their identities as not-racist-people it can cause them to avoid confronting their participation in behaviors and structures that perpetuate racism. DiAngelo wishes white Americans would adopt something of an "I'm trying to perpetuate racism less" point of view instead of the less helpful "I don't have a racist bone in my body" defensiveness.

"Less bad" is something like this. It acknowledges that good is not something you can award yourself and forget about, like a ribbon or a badge, but something you have to always be in pursuit of, and always in doubt about whether or not you're on the right trail.

Good is hard

It is frustratingly difficult, time-consuming, and uncertain to discern the right from the wrong. It is maddeningly easy to come up with convincing justifications for why the wrong thing that would be easiest to do is really the right thing if you look at it properly.

The upshot is that if you aim for benevolence, it's going to take work: clear thinking, painstaking discernment, strong [willpower](#). What lures most people away from benevolence and into evil is probably not the usual temptations of the devil but simple laziness — unwillingness to do this hard work:

"The sad truth is that most evil is done by people who never make up their minds to be good or evil." —Hannah Arendt^[3]

Arendt made *thinking* (of a particular type) the key to becoming good. Evil — "banal" evil as she famously put it^[4] — is committed, in her theory, by people who do not think. This isn't to say that these people are not intelligent, or cultured, or knowledgeable. "Thinking" has a particular meaning in Arendt's framework:^[5] it is a process of internal dialog, one that is necessarily done in withdrawal from society and real-world concerns (that is, you can't think at the same time you are working or conversing). This withdrawal she calls "solitude" but it is a solitude that you share with yourself in a peculiar duality that enables the dialog to take place: you split in two and converse with yourself (and, crucially, you realize that you have to live with yourself).

Thinking of this sort is not a method for *determining* hard-and-fast eternal truths about good & evil, but is a process of doubting and testing. "We have a tendency to think of people who are in the habit of examining basic propositions and standards as destructive. We have every reason to change our minds on this subject. Doubters and skeptics are more reliable, not because doubting is wholesome or skepticism good but because such people are used to make up their own minds — to live together with themselves."^[6]

If you do not think, you are "rootless" — at the mercy of the winds of fashion that might blow you into some new, pathological moral convention. It's not necessarily the case that having roots means that you're wisely-rooted, but it does mean that you

have a stake in your own personality and self-imposed limits on what you are capable of doing. Without these roots, you have no limits, you are capable of anything, and your own character is a matter of indifference to you. In short: you are dangerous.

Thinking, which is to say being in dialog with yourself, is what gives you this stake in your own character — it “results in conscience as its by-product.” You don’t want to be spending your time in dialog with a monster: “If I do wrong I am condemned to live together with a wrongdoer in an unbearable intimacy.”^[3]

Well, that’s an interesting perspective, but I’m not sure I buy it. It seems to me that if you could not stand to live with yourself if you were a murderer or a crook or what have you, then you’re already most of the way to benevolence before you start “thinking.” But how do you *acquire* this idea that being locked in internal intimacy with a wrongdoer is so awful — especially when all you have to do to avoid it is to sink into the luxurious bliss of not-thinking?

Arne Johan Vetlesen is one of those who thinks Arendt got a little carried away by her theory and by her own intellectualism.^[7] Vetlesen thinks that the feeling of right and wrong and the *shame* of being a wrongdoer have more to do with things like *empathy* and sensitivity than thinking. *Adolf Eichmann* may have been thoughtless, Vetlesen thinks, but he did the evil things he did not because he didn’t think but because he was astonishingly devoid of *compassion*.

Conclusion

I have no conclusion.

1. ^

Adam Kotsko, [“Why We Love Sociopaths”](#) *The New Inquiry* (4 April 2012)

2. ^

George Eliot, *Middlemarch* (1872), spoken by the character Dorothea Brooke

3. ^

Hannah Arendt, *Responsibility and Judgment* (2003)

4. ^

Hannah Arendt, *Eichmann in Jerusalem* (1963)

5. ^

Hannah Arendt, *The Life of the Mind* (1978)

6. ^

Hannah Arendt, “Moral Responsibility under Totalitarian Dictatorships” (undated)

7. ^

Arne Johan Vetlesen, *Evil and Human Agency* (2005)

Notes on Ambition

This post examines the virtue of **ambition**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

Is ambition a virtue or a vice?

The noble Brutus
Hath told you Caesar was ambitious:
If it were so, it was a grievous fault.^[1]

“Ambition” has undergone a shift of meaning over time. Today, ambition is often associated with positive things like having lofty goals, drive, initiative, aspiration, being a hard worker, not settling for mediocrity, and that sort of thing.

But it wasn't that long ago that “ambition” was almost always a bad word that described a vice. It was more associated with a blinkered and ruthless pursuit of power, influence, and position. [Lady MacBeth](#) might be thought of as the poster child for this sort of ambition.

So, as was the case with the virtue of [prudence](#), which also had a shift of meaning over time, we need to be especially cautious when we read about the virtue (or vice) of ambition that we understand what the author had in mind.

Aristotle ran into a similar problem when he tried to identify the [“golden mean”](#) concerning ambition (*φιλότιμος*) or lack of ambition (*ἀφιλότιμος*).^[2] In Greek, both of those words had either good or bad connotations depending on context. If someone was being ambitious to an unseemly extent, you might compare them unfavorably to a properly unambitious person. But if a person failed to set their sights high, you could also chastise them for not being ambitious. Aristotle complained that “as there is no recognized term for the observance of the mean, the extremes fight, so to speak, for what seems an empty place.”^[2]

Several years ago, [Swimmer963](#) shared a couple of insightful posts about how she wrestled with ambition. She noted that the [ambition-is-evil](#) sense of ambition can discourage people from developing the good kind of ambition:

I can't trace the roots of this idea completely, but for whatever reason, I spent a long time thinking that being ambitious was in some way immoral. That really good people lived simple, selfless lives and never tried to seek anything more. ... [!]t's a way to feel superior to people who've accomplished cooler things than me, of whom part of me is actually jealous, and that's not the person I want to be.

What is ambition?

Ambition, in the good sense, seems to have these components:

1. You have a goal and a strong desire to reach that goal.
2. That goal is not merely a wish or [hope](#) for a future outcome, but is something that necessarily involves your own effort. So your strong desire comes packaged with an intent to follow through and do the work [earnestly](#). (This suggests a role for the virtue of [determination](#).)
3. The goal, to be an ambitious one, should be challenging to meet and high-impact in its results.

Sometimes people define ambition such that the goal is necessarily *about* yourself: for instance a goal to be an Olympic swimmer, or a successful entrepreneur, or to win an Emmy. But I also see ambition used to describe other sorts of goals: someone who decides to end world hunger, cure cancer, etc. could be considered to be ambitious, even if they did not have as part of their goal that they be personally honored or acknowledged for having accomplished those things.

A question I have is whether ambition is properly to be thought of as a virtue of its own, or whether improperly-tuned ambition is more a *symptom* of failures in other virtues. For example, if you do not show enough ambition, this may be because you have a fear of failure or of responsibility, you don't have faith in yourself, you give up at the first sign of trouble, or you are lazy. Those are things that implicate virtues like [courage](#), boldness, confidence, [endurance](#), or [industriousness](#). If you are ambitious in the bad way, this typically demonstrates itself through ruthlessness, betrayal, dishonesty, and things like that (doing whatever it takes to get ahead). These things also implicate virtues like [honor](#), [loyalty](#), [honesty](#), and so forth. It may be the case that if your other virtues are well-tuned, proper ambition will just naturally arise as part of the package.

Ambition and aspiration

Philosopher [@AgnesCallard](#) contrasts ambition with aspiration.^[3] To oversimplify (and I haven't read her book yet, only some interviews about it): Ambition has to do with acquiring things of already-ascertained value: money, power, fame, and the like. Aspiration is more transformative: it anticipates that one might radically change one's own values and viewpoints; it is more unsure about where it is going and what it will find there.

In ambition, you know what the answer is (e.g money, power, fame) and set out to get it. In aspiration, you know you want the answer but you're not sure what that answer is, so you set out to find out.

For example, one might have the *ambition* to become a professor because you like the idea of people listening to you lecture, being able to dole out grades on your whim, having a respectable job, and that sort of thing. But you might *aspire* to become a professor because you anticipate that you will be transformed by that role in unexpected but beneficial ways. The ambitious person will be consumed with questions like "how do I get tenure?" or "how do I impress the hiring committee?" while the aspiring person will be consumed with questions like "how does a professor think?", or "how does one 'profess' well?"

You may, for example, *aspire* to appreciate jazz because you expect that you will find it valuable, even though you don't really get it yet. Over the course of learning about jazz and listening to jazz, you discover the things that are valuable about it and so you acquire an appreciation of those values, but those were not things you originally had

any ambition to acquire, simply because you had no idea what they were. Or maybe over the course of learning about jazz, you learn that you really prefer blues, or you really like dimly-lit cocktail bars whatever happens to be playing, or you come to value improvisation and spontaneity. Your quest is more tentative: you are still actively driven, but more flexible about your destination.

Ambition and honor

Ambition is sometimes defined as the love of honor.

The *megalopsycche* or great-souled man that Aristotle describes (and that I recapped in [my post on the virtue of honor](#)) is someone who values honor above all else, and is single-minded in pursuit of it.^[4] Aristotle suggested that *megalopsychia* was something like ambition on a grand scale, in the same way that an extravagant display of philanthropy might be considered generosity on a grand scale.

When genuine honor is not what is being sought, but only the fame and admiration that go along with being honored, then the ambitious person is vulnerable to being taken in by flattery and [other sorts of counterfeits](#).

This seems to be one example of a broader ambition failure mode. Other examples would be having a goal of *being* a rock star instead of *making* great music; of *being* a best-selling author instead of *writing* a great novel; of *being* a hero instead of *doing* something heroic. Such mistakes mean that you are more likely either to leave your ambitions stillborn at the daydreaming stage or to seek for shortcuts that leave you short of a really ambition-worthy goal.

This could also be described in terms of Callard's ambition/aspiration distinction: The ambitious person wants to be more or less the same person they are now but with the added prestige of being a rock star; the aspiring person wants to change who they are such that they become the kind of person who makes excellent rock music.

Ambition and certain other virtues

Ambition can be more or less wise. Knowing which goals are realistically attainable (if only with difficulty), and knowing to avoid adhering to ambitions that come with unacceptable risks, are skills for which [prudence](#) (both in the sense of practical wisdom and in the sense of [caution](#)) is helpful.

Ambition can be thwarted by a nihilistic sense that nothing really matters much anyway. Why put in extraordinary effort to achieve some difficult goal when free will is an illusion, everybody dies, and the heat death of the universe is only a few aeons away? For this reason, virtues like [hope](#), [reverence](#), enthusiasm, a sense of purpose, and [optimism](#) may come to the assistance of ambition.

Ambition can also fall victim to poor self esteem. Who do you think you are, anyway, to have such ambitions? What makes you think you're special, to think you can do something extraordinary that other people aren't doing? So self-worth, self-respect, self-esteem, and pride can also come to the aid of ambition.

Ambition is also assisted by the virtue of confidence. One way to become more confident is to do more things successfully. But if you only do things at which you

confidently succeed, you aren't really stretching yourself in an ambitious way. If you do stretch yourself ambitiously and fail, this may mean you take a hit to your confidence which may tend to make you less ambitious. This dynamic may seem like something of a dilemma, but can probably be better characterized as a difficult balancing act. Some suggestions:

- Try not to fail *traumatically*. Choose ambitions such that if you fail, the failure will not be so catastrophic that it will scar you and leave you overly-timid.
- Try to recharacterize failure as a less-awful thing. The "[fail fast](#)" buzzword is an example of this: it characterizes failing skillfully as a variety of successful outcome.
- Learn from your failure. If you do not simply turn away from failure in anguish, but try to wring as much experience as you can from it, your failures can make you *more* confident rather than less.
- Beware of the fallacy of overgeneralizing from a single failure. Sometimes people fail at something and then immediately jump to the conclusion that they aren't the sort of person who succeeds at things. If you catch yourself doing this, notice the [fundamental attribution error](#) you are making and correct for it.

How to know if you need to adjust your ambition?

Swimmer963, in her series of posts on developing ambition, noted:

[I don't fail at things very often. Far from being a success, this is likely a sign that the things I'm trying aren't nearly challenging enough.](#)

In other words, a good heuristic for whether or not you're sufficiently ambitious is whether or not you're failing occasionally. If you never failing, you're probably not challenging yourself as much as you should.

I don't think that we can expect this heuristic to work in reverse. If you fail very often, that *might* mean that you're too ambitious, but it could also mean that you lack follow-through, are easily-distracted, lack the patience to develop skill in the things you attempt, are unwilling to endure setbacks, unwisely choose goals that are inherently unattainable, or fail for a number of other reasons.

1. ^

Mark Antony, in Shakespeare's *Julius Caesar*

2. ^

Aristotle, *Nicomachean Ethics*, book IV, section 4

3. ^

Agnes Callard, [Aspiration: The Agency of Becoming](#) (2018)

4. ^

Aristotle, *Nicomachean Ethics*, book IV, section 3

Notes on Perseverance

This post examines the virtues of **perseverance**, **persistence**, **resilience**, **grit**, **fortitude**, **tenacity**, **sisu**, and others in that bailiwick. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more and to become better at it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What are these virtues?

People who exhibit these virtues rebound from setbacks and are not easily discouraged. They are more apt to overcome challenges than to be overcome by them.

Some of these virtue-words emphasize coping-with-adversity (resilience, fortitude), others emphasize continued effort (tenacity, perseverance, persistence), and others are a combination of those and other things (grit, *sisu*).

"Grit" in particular has been investigated under the positive psychology banner in recent years, and so deserves some extra attention here. As the researchers define it, grit is a combination of perseverance and passion. Perseverance in turn is defined as consistency of effort, while passion, less intuitively, is defined as consistency of interests. (Passion, in other words, has less to do with the level of present enthusiasm than with sustained interest over the long term. Researchers measure passion in people by asking them how little they identify with statements like "new ideas and projects sometimes distract me from previous ones," or "my interests change from year to year."^[1]) Grit is correlated with (and, some critics say, difficult to distinguish from) "conscientiousness" in the "[big five](#)" model of personality.

["Sisu"](#) seems to be a sort of *in extremis* perseverance: persistence to the bitter end in spite of terrible odds, when you feel you have already given your all.

Related virtues

Sometimes these virtues are accompanied by a mental attitude of determination, [resolve](#), steadfastness, firmness, or [duty](#).

A related virtue that is more passive but has a similar sense of putting up gracefully with adversity is [endurance](#). Another along the same lines is [patience/forbearance](#).

Fortitude and resilience can be assisted by [toughness, hardiness, and strength](#).

The virtues of resourcefulness, inventiveness, and creativity can help you to persevere by coming up with new ways to overcome challenges when old ways fail.

One way of looking at the virtue of [courage](#) is to consider it as a sort of specialized perseverance that operates when the main challenge to overcome is fear.

The virtue of [self-control or discipline](#) resembles perseverance. Some researchers contrast perseverance and self-control in this way: self-control is more about staying on task in the short term in the face of temptations, while perseverance is more about staying on task in the long term in the face of obstacles. Eliezer Yudkowsky, in his essay [“On Doing the Impossible”](#) suggests that we might meaningfully analyze varieties of perseverance on even more fine-grained timescales: Just persevering past evaluating a problem as “impossible” to evaluate it more accurately to be merely “very very difficult” may be an important and useful skill.

Related vices

If you lack these virtues, you may be a quitter: you fold too quickly, you give up too early, you’re easily burned-out or overwhelmed.

You can also fail, however, by over-doing it. You might be called bullheaded, obstinate, pertinacious, or stubborn if [you don’t know when to quit](#). If you stick with a lost cause because you are indulging the [sunk-cost fallacy](#) or [status quo bias](#)... or if the only reason you refuse to stop the car and turn around is because you want to put off the moment when you have to admit you went the wrong way... then maybe it’s time to dial back or fine tune your perseverance. At the extreme, perseverance can be an obsessive monomania: flogging the bones of a long-dead horse, [confusing beating your head against the wall with making productive effort](#).

So perseverance is another virtue that maps well to [Aristotle’s golden mean theory](#).

Conflicting virtues

Perseverance comes highly recommended. But there is another vision of the good life that doesn’t put much stock in perseverance. Perhaps, its advocates argue, it is better to be *flexible*: to flow around obstacles rather than trying to wear them down. When fate interrupts your plans, pivot gracefully and choose new plans rather than trying to stubbornly fit today’s realities to yesterday’s agenda. [Fail fast](#) and move on. Life is a dance, not a skirmish, and you are better off trying to match the tempo and mood of the music you hear than trying to stomp your feet against the rhythm in the hopes that you’ll defeat the drummer.

Maybe both perspectives are right in their own way. If you lack the virtue of perseverance, you will not be able to persist when persistence is called for. But whether or not persistence is called for in any particular situation is something that requires wisdom and discernment to know. Perseverance is a capability that can be helpful, not an ideological stance for all occasions. But if so, is it still a *virtue*? A virtue is a characteristic habit. If you are *characteristically* persistent, maybe that’s a bit overmuch. The *virtue* concerning perseverance might then be better described as the capability to persist combined with the wisdom to “know when to hold ‘em, know when to fold ‘em.”

What good is it?

Some opportunities in life are only available to those who can persevere. If you want to become a top-notch pianist who delights others with your excellent playing, for

example, you have to endure thousands of hours of being a mediocre pianist, practicing alone. If you have the virtue of perseverance, you can expand the range of opportunities available to you to include things like these. If you cannot persevere, you limit yourself to living off the low-hanging fruit of life.

To persevere sometimes means to reach the end of what you think your capabilities are and what you think can be done, and then to keep going anyway. To the extent that this succeeds, you not only come closer to meeting whatever immediate goal you had, but you also improve your knowledge of your own capabilities and of what is possible. You become more confident and capable.

“The existence of reservoirs of energy that habitually are not tapped is most familiar to us in the phenomenon of ‘second wind.’ Ordinarily we stop when we meet the first effective layer, so to call it, of fatigue. We have then walked, played, or worked ‘enough,’ and desist. That amount of fatigue is an efficacious obstruction, on this side of which our usual life is cast. But if an unusual necessity forces us to press onward, a surprising thing occurs. The fatigue gets worse up to a certain critical point, when gradually or suddenly it passes away, and we are fresher than before. We have evidently tapped a level of new energy, masked until then by the fatigue-obstacle usually obeyed. There may be layer after layer of this experience. A third and a fourth ‘wind’ may supervene. Mental activity shows the phenomenon as well as physical, and in exceptional cases we may find, beyond the very extremity of fatigue distress, amounts of ease and power that we never dreamed ourselves to own, sources of strength habitually not taxed at all, because habitually we never push through the obstruction, never pass those early critical points.” —William James^[2]

Grit in particular

A variety of perseverance has been studied under the name “grit”—I mentioned it and its idiosyncratic definition earlier. It is said (by researchers like [A.L. Duckworth](#), who has made it her particular focus of study) to predict success in a variety of fields —from spelling bees to West Point graduation— better than other metrics.

Duckworth believes that grit, which does not correlate with intelligence, can explain why higher IQ does not reliably result in greater success, achievement, and happiness.

People with more grit are happier, more optimistic, have better self-control, and report more life satisfaction. They are more likely to engage in activities that require (or benefit from) sustained practice.

People with more grit “tend to make fewer career changes,” and “progress farther in their formal education.”^[3] They are in general less able to imagine better paths than whatever path they are currently on. This could be interpreted in at least a couple of very different ways (e.g. they are good at finding and taking optimum paths and can usually be found on them, or, they are poor at finding and taking optimum paths because they are less able to envision alternatives they might choose from). One paper suggests this interpretation: “that grittier individuals generate lower estimates of the opportunity cost of their current pursuits.”^[3]

It’s important to keep in mind, when reading the enthusiastic literature about the power of grit, how grit is measured. Typically in these studies a person’s grit is measured with a number that is a function of how much the person says they identify

with a small set of self-assessment statements. These statements include things like “I have overcome setbacks to conquer an important challenge,” and “I finish whatever I begin.”^[1] So when you hear the researchers say that someone has more grit, what they’re really saying is that someone is more likely to think statements like that apply to them. And when the researchers then say that someone with more grit is more likely to be able to get through “beast barracks” at West Point or make it to the final round of a spelling bee, this may mean that someone who reports that they are the sort of person who *has* risen to the occasion and met challenges has, true to form, yet again risen to the occasion and met challenges. There’s a limit to what that can tell us about what factor *causes* people to be able to repeatedly rise to the occasion and meet challenges, and calling that hypothetical X-factor “grit” may make it seem more explanatory than it is, given how grit is defined.

Here are a couple of papers that make for a good introduction to the grit lit:

- Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). [Grit: Perseverance and passion for long-term goals.](#) *Journal of Personality and Social Psychology*, 92(6), 1087–1101
- Eskreis-Winkler, L., Gross, J.J., & Duckworth, A.L. (2016). [Grit: Sustained self-regulation in the service of superordinate goals.](#) In K.D. Vohs & R.F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory and applications* (3rd ed.). New York, NY: Guilford

How to improve in perseverance

If there is a well-tested silver bullet for improving perseverance in general, it has not yet come to my attention. Instead in this section I will briefly describe some ideas that seem promising, or some insights that might apply in limited circumstances.

- In some domains, you can learn to persevere for a long period or through great difficulty by incrementally persevering through modestly longer periods and modestly greater difficulties. This may be helpful for the sorts of challenges that lend themselves to repetition and where you have control over the duration and difficulty level. For example, I now have a 30-minute-per-day meditation habit, and while I find meditating *well* a challenge, I’m no longer challenged by just sitting still and making the attempt for half an hour. But when I was trying to establish the habit, just staying on the cushion without fidgeting for ten minutes seemed at times like more than I could stand. Steady, incremental practice has transformed the intolerable into the easy.
- Optimistic people tend to describe adversities as temporary, out-of-the-ordinary, and changeable; pessimistic people tend to describe them as typical, run-of-the-mill, and indelibly woven into the fabric of life. This may partially explain the correlation between grit and optimism: if you are optimistic you are more likely to see the obstacles that confront you as surmountable. So one possible sideways avenue for improving your perseverance might be to work on your [optimism](#).
- The Eskreis-Winkler paper^[3] discusses two angles for improving perseverance, which they call logic-of-consequence and logic-of-appropriateness. When you use the logic-of-consequence angle you try to become more aware of the expected positive outcomes of persevering: you do a more explicit cost-benefit analysis, and hope that if you learn thereby that you would [rationally](#) benefit from persevering, your willpower will go along with it. The logic-of-

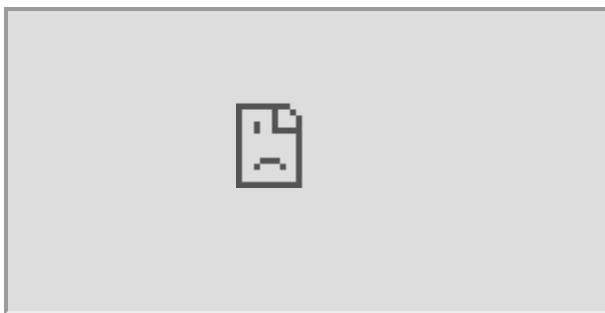
appropriateness angle has more to do with your self-image. What sort of person do you want to be? Are you a milquetoast who wilts at the first sign of adversity, or a hero of the sort that laughs in the face of the storm? [What would Brian Boitano do?](#) There is some indication that just by spending a little time daydreaming — envisioning yourself as the sort of determined, hard-working person who meets difficult challenges — you can give your perseverance a temporary boost. The motivating power of [the St. Crispin's Day speech](#) from Henry V, for example, comes in part from how it asks the listeners to imagine themselves in the future, *having overcome* the challenge before them.

- One of the benefits of persevering is whatever reward is on the other side of the obstacle before you. But there is also the potential reward of being the sort of person who overcomes obstacles, who has the virtue of perseverance. If you value perseverance itself, beyond its value in overcoming any particular obstacle, and think of perseverance as something that would burnish your character, that may help motivate you to practice it. You may come to see challenges not just as obstacles to getting what you want or need but also as opportunities to demonstrate and practice a valued character trait.
- Giving up is one effective coping strategy for dealing with stress. You turn your back on what is causing you stress, and your stress goes away. It may be that you over-rely on that coping strategy because your alternative strategies are underdeveloped. You might be able to become more persevering if, when you encounter a stressful situation, you have other strategies you can try that don't rely on escape from the stressor. Learning and practicing such strategies, and having a plan to deploy them when they will be helpful, may help you persevere.
- Obstacles sometimes take the form of negative emotions like fear, anger, sadness, or embarrassment. If you see such emotions as awful things to be avoided at all costs, you may be more reluctant to persevere. If you are more forgiving to yourself and see these emotions as part of the game, you will be better able to shrug them off and continue.
- Perseverance is not just a matter of character, it is also a matter of resources. [\[4\]](#) If you are living on the edge, you may have to cut your losses quickly in some endeavor to avoid losing even more elsewhere; whereas if you have a surplus in your life, you may be able to use it to subsidize going further out on a limb. Maintaining a surplus buffer of time and resources in your life can be a way of improving perseverance.
- Among these resources are social resources and social skills like being able to ask for help or advice. [\[5\]](#) For example, if you have a spouse or trusted friend, you may be able to say "this project is taking a lot longer than I thought it would; could you pick up the kids from school so I can keep at it?" If you help your friends when you have excess resources and slack, you can better expect help from your friends when you are tapped and stressed. This can make you more resilient by smoothing-out acute challenges. Other people can also help you discover and explore new avenues to meet your challenges, can give encouragement, and can suggest ways to bolster your resolve.
- Similarly, you can reduce how much you have to rely on extraordinary reserves of fortitude by planning carefully and by anticipating and preparing for possible difficulties. If you plan for difficulties, then when difficulties arise your first impulse will be to put your plans into effect rather than to contemplate surrender.
- Sometimes people overgeneralize from examples of personal failure. They fail at some task, and then generalize to say that they are a failure at tasks like that. I flubbed that speech... because I'm terrible at public speaking. That interview went horribly... because I don't really interview well. This can make

perseverance seem hopeless. If you catch yourself generalizing in this way, consider applying some extra skepticism to your reasoning.

- Some people report being able to strengthen their perseverance by making a pledge, taking a vow, or something of that sort. There's usually magic of some sort involved: making the declaration publicly, or in writing, or using formal language, or invoking something sacred — something that sets the vow apart as unusually potent and compelling.
- The presence of witnesses can help or harm your persistence, depending. If what you are trying to persist at involves a series of minor successes leading (hopefully) to a final triumph, the presence of witnesses to your intermediate successes can help you persist longer. However, if what you are trying to persist at involves a series of failures that eventually (hopefully) culminate in an ultimate success, the presence of witnesses to your failures can discourage you and cause you to give up.^[6]
- Sometimes we persevere because perseverance is forced on us. These are typically situations that we would not choose, but, finding ourselves in them, we can learn something about ourselves from them. There is no shortage of thrilling stories of people who endured horrible experiences through tenacity and unexpected reserves of fortitude, and they often include moments of surprised and admiring self-reflection, and also an attitude of looking with amused contempt at the once-daunting smaller day-to-day challenges of life. There's reason to beware of [survivorship bias](#) here, but maybe there's something to be said for trying to wring some extra wisdom out of our most difficult moments.
- The theory of [learned industriousness](#) posits that people can practice exercising persistence in some arena in which they are rewarded for that persistence (perhaps some artificial task invented for this purpose), and then this will help them apply persistence in other unrelated areas. So you may be able to improve your persistence by engaging in hobbies or other pastimes that have a persistence component (video games, jigsaw puzzles, marathon running).

A.L. Duckworth, "Can Perseverance be Taught?":



1. ^

"[12-Item Grit Scale](#)" from A.L. Duckworth, C. Peterson, M.D. Matthews, & D.R. Kelly, "Grit: Perseverance and passion for long-term goals" *Journal of Personality and Social Psychology* (2007)

2. ^

William James, "[The Energies of Men](#)" *Proceedings of the Meeting of the American Philosophical Association* (1906)

3. ^

L. Eskreis-Winkler, J.J. Gross, & A.L. Duckworth, ["Grit: Sustained self-regulation in the service of superordinate goals."](#) In K.D. Vohs & R.F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory and applications* (2016).

4. ^

Kathryn Britton, ["Resilience in the Face of Adversity"](#) *Positive Psychology News* (7 December 2008)

5. ^

Rob Cross, Karen Dillon, and Danna Greenberg, ["The Secret to Building Resilience"](#) *Harvard Business Review* (29 January 2021)

6. ^

R.G. Geen, ["Effects of being observed on persistence at an insoluble task"](#) *British Journal of Social Psychology* (1981)

Notes on Kindness

This post examines the virtue of **kindness**. I wrote this not as an expert, but as someone who wants to learn more about kindness and how to become better at it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is kindness?

"It's a little embarrassing that after 45 years of research & study, the best advice I can give people is to be a little kinder to each other." —Aldous Huxley^[1]

I had more difficulty than I expected coming up with a good definition of kindness. I painted myself into a corner by doing previous write-ups on [compassion](#), [respect for others](#), [care](#), [benevolence](#), and [courtesy](#). Those cover a lot of the same territory, and seemed to leave kindness with little ground of its own. But I think there is some distinct kernel to kindness that makes it different from these things, and this may also suggest different ways of improving it in ourselves.

For example, there's the concept of "random acts of kindness." It's harder to imagine "random acts of care" or "random acts of compassion": those virtues seem to require more intimacy with the recipient and their needs than "random" acts would be appropriate for.

Courtesy and respect for others may just discourage you from being unkind, without also prompting you to kindness.

And there's the trope of the *benevolent* person who loves humanity, but can't seem to get along with anybody in particular and treats individual people condescendingly or brusquely. Kindness seems an especially important skill for an aspiring benevolent to cultivate. As one wise person put it: "Try to always be kind because you never know when you're incompetent."

[B]ecause being wrong feels exactly like being right, you're almost always better off being nice. Kindness covers a multitude of incompetences, including incompetence you didn't even realise you had.^[2]

My attempt to define kindness

I think **you are being kind if you are doing something with the motive of making some other creature's life better for them and you do so skillfully enough that you are likely to succeed in this**. To unpack this:

1. Kindness involves *doing something* (a kind word might be enough, but thoughts-and-prayers probably aren't). It is an activity, not a sentiment.
2. It is prompted by a kind *motive*—a desire to be helpful: to make some creature's life better (or prevent it from getting worse).
3. This motive is to make that creature's life better *for them*—on their terms: not your idea of what would be better, but theirs... typically, anyway. In exceptional circumstances, this might be your wise understanding of their enlightened self-interest even when at cross-purposes to their present desires: e.g. taking your dog to the vet, preventing a suicide.
4. What you do, prompted by this motive, is done *skillfully enough* that it works, or anyway would likely have worked barring unforeseeable accidents or the intransigence of someone who is just dead set on having a bad day. Good intentions without skillful follow-through are not sufficient to be kind.

Quasi-kindness

Does my definition leave out acts of kindness that are done for mercenary motives: to prompt reciprocal acts of kindness, to boost one's reputation or build one's network, to bask in gratitude, and so forth? Such acts might not seem as pure but aren't they nonetheless *kind*? It might be argued that in such cases you are doing *kindnesses* but are not doing them *kindly*. You are not quite practicing the virtue of kindness, but are using skills of kindness for other purposes.

But maybe these can be ways of getting practice... like the way Phil in *Groundhog Day* cynically does good deeds in order to impress Rita, but in doing so eventually comes to appreciate the value of doing good deeds.

If you are trying to do something kind, to make some creature's life better for them, *in order that* you might gain something else besides, I don't think this means you haven't been kind. It just means you've been kind with ulterior motives. Kindness doesn't necessarily require that you are selflessly motivated to make some creature's life better for them as an end in itself. Such a disinterested kindness might arguably be a nobler or better form of kindness, but I don't think it's the *only* form. You may genuinely want to make some creature's life better for them, but as a means to some other end, and by doing so you are still being kind.

I think people sometimes confuse cases like that with cases in which someone merely pretends to be kind—does something that is only superficially or conventionally kind—because they are never motivated to make some creature's life better for them but only to appear kind or to cheaply reap the benefits of having been seen to have attempted a kindness. Such things are not self-interested kindness or kindness with ulterior motives, because they are not really kindness at all.

Evolutionary psychology suspects that subconscious mercenary motives are hidden behind many of our acts of kindness. For example: we are kind to our children and other family members because [kin selection](#) has led to the evolution of such altruism in the service of the selfish gene. [Reciprocal altruism](#) perhaps adequately explains some acts of kindness as ultimately self-interested. Perhaps [sexual selection](#) leads us to be more broadly and flamboyantly kind as a way of broadcasting how comparatively fit we are.

Some of this stuff is very plausible; some is cynical [just-so storying](#) without much evidence behind it. It does suggest some possible avenues for improving one's kindness. For example: if you are already hard-wired to be kind to your kin, can you become more kind in general by expanding your concept of kinship? "All men are brothers" and other similar sentiments seem designed to hack this mechanism in this way.

Kindness as a virtue

"I think probably kindness is my number one attribute in a human being. I'll put it before any of the things like courage, or bravery, or generosity, or anything else... Kindness—that simple word. To be kind—it covers everything, to my mind. If you're kind, that's it."

—Roald Dahl^[3]

How is kindness a *virtue*? What connects being kind with being a flourishing human being? There are a number of ways to consider this.

Kindness can be evidence of flourishing. This is for at least two reasons: For one thing, if you are focusing on the needs of those around you, this probably implies that your own life is in pretty good shape: You've got slack, surplus, ease. For another, if you are kind it may be because you learned kindness as a child, which suggests that you had a caring upbringing from people who successfully modeled kindness for you. This is indirect evidence that you

started out with at least some of the helpful foundation of a flourishing human life. By contrast, if you are unkind this can suggest that you feel yourself to be struggling desperately in an environment where you have to lash out against hostile competitors, or that you were raised by wolves or something.

Kindness is also widely esteemed. People often revere exemplars of kindness, and approvingly share anecdotes of people acting kindly. This suggests that we see kindness as part of being human well. (Though in some contexts and in some subcultures, kindness is looked down on, and cruelty and ruthlessness are admired, so this is not entirely reliable.)

There is also evidence that behaving kindly improves a person's measurable well-being. A wealth of studies have tried to test the hypothesis that doing something kind makes you better off for it. A recent meta-analysis of 27 such studies tried to be extra-careful to rule out publication bias, mere findings of correlation, and other such potential pitfalls.^[4] It found that the interventions studied (usually measuring short-term effects after brief acts of kindness, in [WEIRD](#) research subjects) did seem to support the hypothesis that acting more kind improves your well-being.

For many people, this will probably seem a ridiculous thing to test. Does being kind to people make you happy? Does water slake your thirst? Is the sun warm? People tend to behave in certain ways when others are mean to them and life is going sour, and in other ways when others are being kind to them and things are getting better. If you prefer the way people act in the latter case, it takes no great leap of understanding to note that you can facilitate that by being kind to them.

But this preference is not self-evident. Some people seem to prefer it when those around them are miserable, frightened, angry, and discouraged. If anyone has carefully tested the alternative hypothesis that you can improve your life by being mean to someone, I haven't heard about it, but maybe that works too, or works for some people in some contexts.

Kindness and intimacy

Kindness can be a good opening gambit in an exploration of potential intimacy and friendship. A friend of mine put it this way:

I desire deep connection. Kindness is like testing the waters. If I am kind to someone (based in sympathy and fueled by generosity and general care), the feedback I might receive can inform me if they would like a more brief engagement of shared minds, or if they would like to explore deeper possibilities, wherein we would communicate the vulnerability necessary to be empathetic.... Kindness is sort of a tipping point. Am I going to simply be respectful or considerate, to make sure I don't make things worse? Or am I going to tip into my deeper sense of humanity and strive to be altruistic and find an empathic connection with another person?

Kindness to animals

I enjoy petting cats. Part of this I suppose is the tactile pleasure of soft fur, but I never find myself going out of the way to pet strips of disembodied fur, at least not since my MDMA phase. Without a cat attached, it's not nearly so fun. What I like about petting a cat is the obvious delight the cat experiences from being pet. I give the cat delight by scratching it behind the ears just so, and that cat's delight becomes my delight somehow by some sympathetic magic.

This simple interaction between a person and a cat may help to demonstrate kindness stripped down to its essentials, and also to demonstrate how kindness can contribute to the

flourishing of the person who exhibits it.



"A scout is kind—He is a friend to animals. He will not kill nor hurt any living creature needlessly, but will strive to save and protect all harmless life."

"A scout is kind" is part of the Scout Law. The original wording was "A scout is a friend to animals," and this part of the Law for a long time was defined with kindness-to-animals as the paradigmatic example.

(Contrariwise, cruelty to animals by children is a recognized warning sign that something is amiss.)^[5]

Kindness to animals and to children can be some of the biggest bang-for-your-buck ways of benefiting from showing kindness to others. (Some) animals and children express their delight or relief in uninhibited and legible ways (when compared to the often more restrained and ambiguous ways in which adult humans do so). If you take pleasure from contributing to the joy of others, being able to do so by petting a cat or laughing at a child's joke is a pretty cheap and easy win: it does not require much creativity ahead of time or much second-guessing afterwards. This is in part what makes children delightful (sometimes): they can find joy in simple things that adults are too jaded to notice, and then transmit that joy to us.

The ability to take pleasure in someone else's joy is a nice trick. Rather than having to figure out how to get off your own hedonic treadmill once again, you can just look around you for someone else you can easily delight or relieve, and then bask in their joy.

Random acts of kindness

The term "[random acts of kindness](#)" was coined as a sort of cutely ironic counterpart to news media clichés like "senseless acts of violence," "pointless acts of vandalism," and the like. In a random act of kindness, for seemingly no reason aside from eccentric caprice, you make a special effort to do something kind for someone, usually someone you don't know: sometimes, by design, someone you will never know because by the time the good deed has

hit its target you are already far from the scene. Random acts of kindness also invite you to wonder what it might be like if you lived in a world where nobody was safe from the threat that any random stranger at any moment might do something unexpectedly nice for them without warning. (See also [“pronoia.”](#) [“prospiracy.”](#))

If “random acts of kindness” are acts of kindness, and this isn’t just a cute name for something else, then this highlights how kindness can be anonymous and at least somewhat distant. You can be kind to people you don’t have preexisting relationships with; you can even perhaps be kind to them without them knowing about it. It might be considered a kindness, for example, to remove a tree branch that has just fallen on a bike path, even if nobody but the person doing the kindness ever sees the obstruction or its removal.

People who perform clandestine kindnesses like these cannot as easily rely on the positive feedback they would get from the reactions of those who directly benefit. If such reactions are important motivators for them, they must use their imaginations (or perhaps hide in the shadows or set up cameras). Alternatively, other motivators may be in play: the actor wants to bolster their self-image as a kind person, for example, or they have philosophical reasons for wanting to increase kindness.

Kind intentions

To be kind, are kind intentions necessary, sufficient, both, or neither? I think kind intentions must be at least *necessary*. If you do something that is unintentionally kind—commit a kindness by accident—I don’t think you’ve really *been* kind. Might kind intentions be *sufficient*? I have my doubts. If you are sad and I bake you a cake, but I’m no good at baking cakes and so you now have a cake that you’d rather not eat but feel guilty not eating, I haven’t really done you much of a kindness, have I? Or if you’re sad because you’ve just gotten the news that you’re pre-diabetic and will need to radically adjust your diet and lifestyle to avoid serious health problems, me showing up at your doorstep with a cake to try to cheer you up is not a kind thing to do and I ought to know better. My good intentions were poorly-informed, short-sighted, incurious, counter-productive: insufficiently developed to be the kernel of kindness. Arguably such things do not even demonstrate good intentions: to actually *intend* good you have to be willing to try harder to get it right.

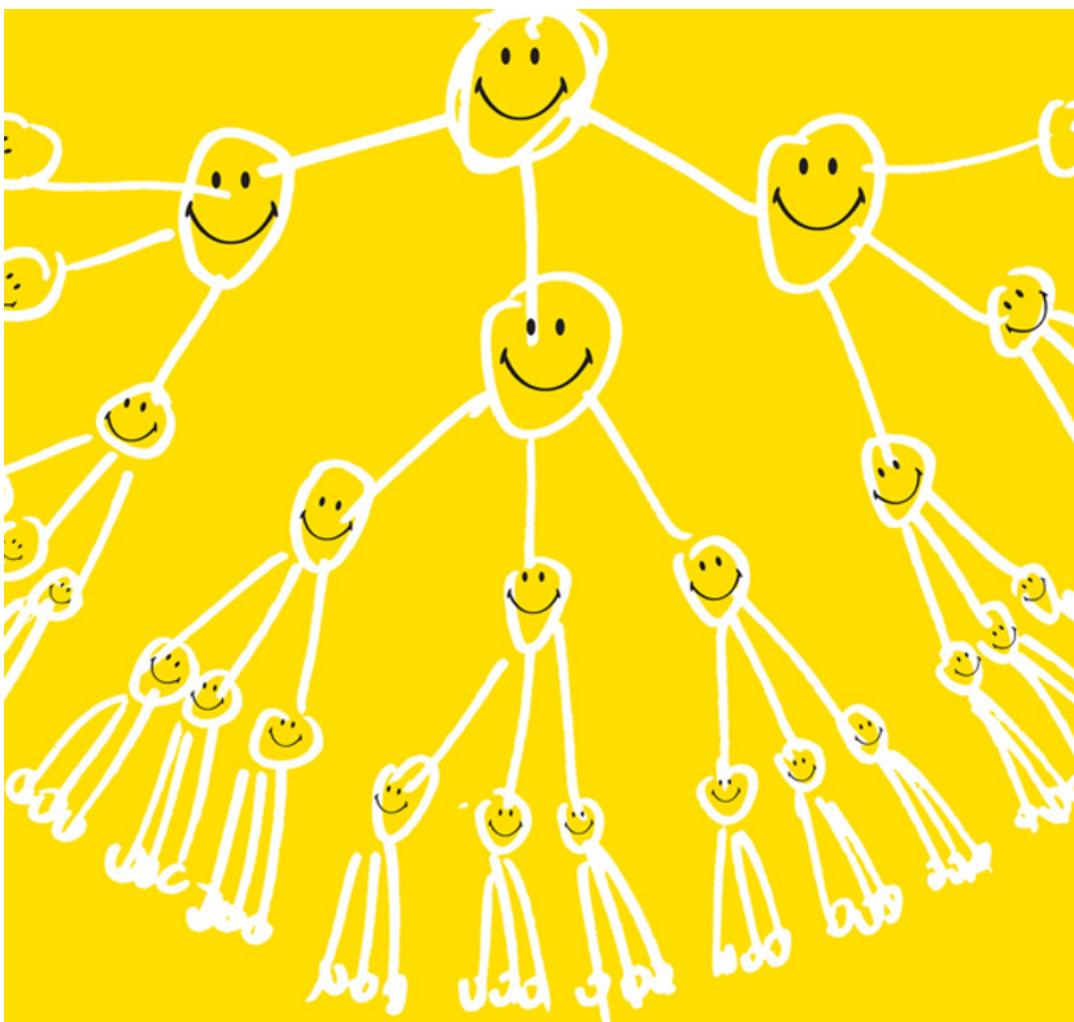
On the other hand, *sometimes* the wrong thing to do can become the right thing by virtue of being done with good intentions. A child comes up to you and says “I heard you were sad, so I made you this drawing.” A child’s drawing was not what you needed to feel less sad, but a child handing you a drawing with the innocent expectation that it will help somehow works as a package deal. The fact that someone cares enough to try to make your life a little better might itself be helpful and kind, even if the attempt they make isn’t very on-point. Attempted kindness communicates a message like “I want you to be happy. If you are not, I wish you were. It would give me joy to find that you are. And so I am genuinely interested in helping that come to pass.” If that message comes as news, it is usually good news.

Non-utilitarian gestures of kindness—bringing flowers to someone in the hospital, for example—are a sort of *benign harmlessness plus good intentions* that may functionally add up to kindness when there isn’t much room for practical intervention. (However such things can also backfire when more practical intervention is called for.)

I’m tempted to say that you can be kind even if your kindness fails to make anyone else’s life any better, if this is for reasons that you cannot foresee or that are out of your control. If you try carefully to console someone who is just plain inconsolable, I think you are still being kind to try. If you do a skillful kindness for someone that they don’t notice because they’re distracted by an unexpected meteor strike or something, I think you have still *been* kind even if the kindness wasn’t fertile. That was kind of you; too bad about the meteor.

Compounding kindness

People commonly observe that they are more likely to act kindly when they are relaxed and unthreatened and have been shown kindness themselves, and conversely that they are more likely to be mean when they are stressed or when people have been mean to them. This, along with the observation that people frequently copy the behaviors they see modeled for them by others, has led to the speculation that acts of kindness may have compounding effects: You are kind to person A, which makes person A more likely to be kind to others, and may also influence bystander B to do something similar; A's and B's kindnesses then similarly affect C and D and so forth. In this way your kind act has both direct and indirect effects.



KINDNESS *is* CONTAGIOUS

GOOD VIRUS PRESENTS A DAVID GAZ FILM "KINDNESS IS CONTAGIOUS"
FEATURING JAMES FOWLER, CATHERINE RYAN HYDE, DACHER KELTNER, DAVID RAND, MUSIC BY BRANDON SCHOTT
CINEMATOGRAPHY BY BENJAMIN KANTOR, JUSTIN KANE, OWEN COOK, ALEX POLLINI, BRAD WILHITE, DAVID GAZ, PRODUCED BY CHANET JOHNSON,
SARA GLASER, ANNELIE WILDER, ROY BODNER, ROBERT HENSLEY, DAVID GAZ, EXECUTIVE PRODUCER AL CATTABIANI, DIRECTED BY DAVID GAZ

CINEDIGM GATHR KINDNESS-IS-CONTAGIOUS.COM /GOODVIRUSKIC @GOODVIRUSKIC #GOODVIRUSKIC

The documentary *Kindness is Contagious* promoted this hypothesis.

This makes intuitive sense, and is pleasant to think about, but it may prove difficult to verify. [6] There must at least be some dampening factor, or by now an out-of-control kindness feedback loop would have made happy saints of us all.

One possible dampener is that it may be only extraordinary acts of kindness that trigger this effect. Ordinary, par-for-the-course acts of kindness only reinforce the ordinary, par-for-the-course status quo. So as the level of kindness in a culture rises and the status quo adjusts, it takes increasingly unusual or difficult acts of kindness to raise the bar further. Furthermore, it becomes easier to do things that now appear *unkind* that in an earlier regime might have just been considered unremarkable. (This speculation was inspired by KatjaGrace's "[Limited kindness is unappreciated.](#)")

The advice to "[pay it forward](#)" encourages people to amplify the kindness positive-feedback loop by acting kindly toward others as a way of showing [appreciation](#) for kindnesses received.

Killing with kindness

"If thine enemy be hungry, give him bread; and if he be thirsty, give him water to drink; for by doing so thou shalt heap coals of fire upon his head." —*Proverbs xxv 21-22*

"You should respond with kindness toward evil done to you, and you will destroy in an evil person that pleasure which he derives from evil." —Tolstoy^[7]

Kindness can be a way of blunting the force of an adversary's attack. You show your adversary kindness instead of the counter-attack they were expecting, and this surprises, baffles, and disarms them. This can short-circuit conflicts that have become disconnected from any grievance beyond the conflict itself—where the point of insults given is just to counter previous insults received. It is also particularly helpful in internet flame-wars or other such situations in which trolls and shit-stirrers seem eager to gin up animosity for its entertainment value. If you out-ridicule someone who is trying to belittle you, you may come across as witty; if you out-ridicule someone who is trying to be kind to you, you probably come across as an asshole. Oops.

Problems with kindness

There are some complications with kindness. For one thing, kindness doesn't happen in a vacuum, but has a context and side effects. It is possible to do kindness to A in a way that is unkind to B, for instance. "I thought you might like a new bicycle so I stole you this one" is a certain sort of kindness, but a problematical one.

Acts of kindness can communicate solidarity. If you are kind to A you may thereby appear to be saying "I am on A's team." This in turn may be interpreted as a slight or an unkindness to people who are on the outs with A. You may have to take pains to make known that you are not [breaking bread with tax collectors and sinners](#) because you've joined team tax collector and sinner.

Outrage mobs and things of that nature use public displays of unkindness to signal disapproval of what a person symbolizes or stands for or is seen as sympathetic with. To be kind instead in such a circumstance can be seen as breaking ranks, being disloyal.

To the extent that you are selective about whom you show kindness to, and in what situations, your criteria for making these discriminations may be interpreted (or misinterpreted) to reveal things about your character or allegiances.

Intended acts of kindness can misfire in various ways. You may intend to be kind by doing something for someone else, but they perceive it as condescending or as a way of putting a spotlight on their neediness or as an affront to their pride. You may try to make supportive statements but you don't realize that they are not perceived as intended (a friend gave me this example: "when someone tells me they can't tell my son has autism I know they mean it in a nice way, but it really doesn't improve anything and their perception wasn't the point in the first place.") Ostensibly encouraging sentiments like "look on the bright side," "it's not so bad," "you'll look back on this and laugh" often fail to be perceived as kindnesses.

Sometimes kindness can exacerbate dependency and ultimately be unhelpful if doing something kind for someone prevents them from learning how to do it for themselves. Occasionally you may do a kindness for someone who then interprets that kindness as something they are entitled to and you as a resource that they can tap in a sort of parasitic way, which can be awkward and inconvenient.

Kindness can sometimes ironically make people suspicious and defensive.^[8] What's your angle? Who are you trying to impress with that Mother Theresa act?

Becoming kinder

Kindness can be a surprisingly complicated mechanism, and so it is not a simple matter to describe how to build or repair it. In this section I'll briefly describe some possible angles for tinkering with this mechanism.

To become a kind person it helps to have been raised that way: to have been under the care of adults who were kind to you, who modeled kindness for you, and who valued kindness in you. Nobody is surprised to learn that some person who has been caught being cruel and awful turns out to have had a childhood inflicted on them in which kindness was rare. There are also some explicit things you can do to help a child learn to be kind. For example, you can [notice, praise, and reward](#) kind behavior in a child when you see it (I expect something of this sort probably works in grown-ups too).

If you're reading this essay it's probably too late for you to correct your own childhood: you either got a good upbringing or you didn't, and there's not much you can do about it now. So what can you do as an adult to improve in the virtue of kindness, to be characteristically kinder? To become *habitually* kind means either ① to deliberately practice kindness, as one practices playing an instrument, from an aspiration to become a kind person; or ② to engage in kindness in a way that is itself rewarding and so self-reinforcing.

In case ② maybe you are kind ③ because you desire the other person's well-being and so contributing to that well-being directly via kindness helps you to satisfy your desires. If not, maybe you are kind ④ because their well-being rewards you in a more indirect way (they are more pleasant to be around, they are inclined to do nice things for you in return, the world around you runs more smoothly when it's occupied by people having good days). I'm not sure if ③ can be forced. If you don't give a fig for the well-being of another person, can you decide to change your mind about that somehow? Perhaps. Some forms of [mettā meditation](#) seem like they are trying to brute force a sort of empathetic imagination in this way.

Meanwhile ④ is indirect enough that it may require close attention for you to trace the path from your act of kindness to the eventual desired result. On the other hand, maybe an "aha!" moment is enough: Someone points out to you that life seems to go a lot more smoothly when you're nice to those around you, and you think "you know, that makes a lot of sense. Why didn't I think of that before?" However, ④ is vulnerable to an additional layer of frustration: you might succeed in being kind, but then that kindness fails to bring about the hoped-for fringe benefit. If you're not careful, such a thing could disincentivize your kindness.

Back to case ①, in which kindness is not (sufficiently) inherently rewarding to make becoming kind nearly effortless. You have to work at it. There are several reasons you might want to. Perhaps you have a hard time caring about particular people, but philosophically, in the abstract, you are philanthropic, and you see kindness as a way of advancing this. Or maybe you believe that kindness is admirable (or recognize that it is at least admired) and so you like the idea of *being kind* even though practicing kindness mostly leaves you cold. Perhaps you do care about the well-being of individual people, but something is broken in the feedback circuit such that even when you are kind to people successfully you do not get the sort of reward that reinforces your kindness and so you struggle to form the habit.

In case ① you may first need to strengthen your motivation for being kind. You might want to motivate yourself by artificially tying some reward to your acts of kindness: give yourself a gold star, take yourself out for ice cream, pat yourself on the back, write it down in your book of kindnesses. In case ② the motivation is there but maybe there is something else that interferes with you being as kind as you'd like to be: another motive that overrides the motive to be kind, perhaps, or a lack of skill in translating your motives into kind action. It may take some sustained inquisitiveness to discover where the snag is.

If you have kind intentions but fail to put them into action, this might be from a sort of timidity or fear of failure, in which case you may want to put some effort into improving virtues like boldness, initiative, or [courage](#). If you want to be kind to people, but you fail to notice opportunities to do so, maybe working on [attention](#) would help. A lot of missed opportunities for kindness (and a lot of small, thoughtless unkindnesses) come from inattention and insensitivity.

It is easier to attend to (or just to notice) the needs of others when your own needs are not desperate for your attention. If you can satisfy [or reduce](#) your own needs, you will better be able to be kind to others. If you can accumulate a surplus of resources, you can more easily deploy those resources in kind ways. Being in a good place in your life also helps to make kindness more motivating: contributing to another person's good fortune might have a bitter taste if you have to compare it to your own bad fortune. In short: you may be better able to help others by first shoring up your own life. (And being kind to yourself may be good practice for being kind to others.)

If you find that your attempts at being kind often misfire, you may need to work on imagination and [empathy](#). I know someone who has kind intentions, but his imagination usually begins by coming up with something he would think was kind if it was directed toward him, and then jumps immediately to imagining how grateful the recipient will be if he directs that kindness toward them, without carefully examining the middle stage: will they find it as kindly as he would? It might take more careful attention for you to imagine a way to intervene in someone else's life that will be seen as a kindness by them. It's a real art to be able to imagine and empathize with another person's point of view. You might consider asking for advice from someone you know to be kind: tell them about someone you would like to be kind to and ask them if they can walk you through their thought process as they consider ways to do that well. Another option is to choose simple ways of being kind that don't require as much subtlety and second-guessing.

More about mettā

Mettā or “loving-kindness” meditation is an attempt to deliberately evoke kind motives and thereby provoke kind behavior in oneself. At first glance it seems tailor-made for someone seeking to improve their own kindness. However, mettā meditation often explicitly aims at much grander goals than mere kindness—loving all living things, aiming at the end of suffering for all beings living and yet to live, and grandiose things of that nature. I don't have a good feel for whether it would be appropriate or useful in strengthening more ordinary, secular kindness, but it seems worth investigating.

A typical mettā meditation takes you through a sort of [expanding circle](#) visualization. You begin by looking at your own life and its fundamental needs and your desire for fulfillment and relief from suffering & anxiety. You [hope](#) that these desires will be met and this relief will come. You then consider certain people that you care about, their needs and so forth, and extend this hope to them. You then consider mere acquaintances, strangers, even enemies, members of other species, and so forth, pushing the boundaries of your kind wishes and intentions along the way.

Conclusion

I hope this has suggested some ways in which the virtue of kindness can be a valuable thing to develop, and some possible angles you can explore if you want to strengthen or widen your kindness. In researching kindness, a lot of what I came across amounted to [applause lights](#) that didn't strike me as very actionable or insightful. I would welcome your recommendations for resources I have overlooked.

1. ^

quoted in Laura Huxley, *This Timeless Moment: A Personal View of Aldous Huxley* (1975)

2. ^

["Try to always be kind because you never know when you're incompetent"](#) DePonySum (22 September 2020)

3. ^

interviewed by Brian Sibley, *BBC World Service* (November 1988)

4. ^

Oliver Scott Curry, Lee A. Rowland, Caspar J. Van Lissa, Sally Zlotowitz, John McAlaney, & Harvey Whitehouse, "Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor" *Journal of Experimental Social Psychology* (2018)

5. ^

Joni E Johnston, "Children Who Are Cruel to Animals: When to Worry" *Psychology Today* (27 April 2011)

6. ^

Jamil Zaki, ["Kindness Contagion: Witnessing kindness inspires kindness, causing it to spread like a virus"](#) *Scientific American* (26 July 2016)

7. ^

Leo Tolstoy, *A Calendar of Wisdom* (~1910) January 30

8. ^

See, for example, u/thecontradictin, ["Is it bad that I'm always suspicious of kind acts that are recorded?"](#) *r/Stoicism*

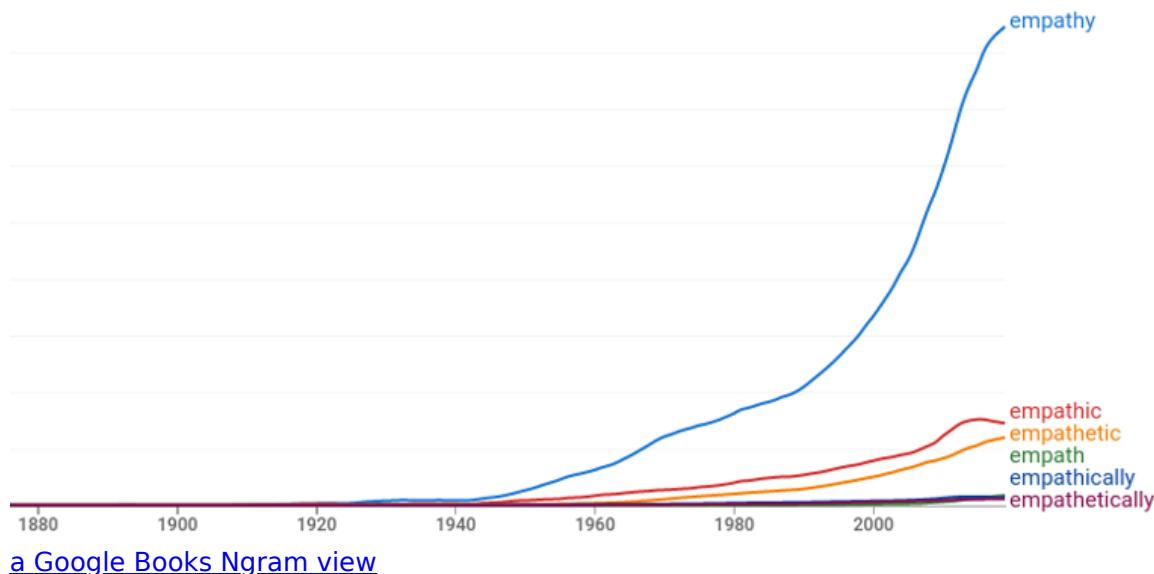
Notes on Empathy

This post examines **empathy**, as part of [a sequence of posts about virtues](#). It is mostly an exploration of what other people have learned about empathy, rather than my own research or opinions about it, though I've been selective about what I found interesting or credible. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will help people who want to know more about empathy and how to practice it skillfully.

This was a particularly interesting (and difficult) post to research and write. There is a wealth of literature about empathy, and I found a lot of it surprising. I feel kind of like the kid who opened up the back of dad's watch to see how it works and now is sitting amidst a cubic yard of springs and gears wondering how to put it all back together.

What is empathy?

The word "empathy" is pretty new. It apparently was coined in 1908, and only really caught on in the last half of the 20th century.^[1] (The concept itself is not so new; earlier authors sometimes deployed the word "sympathy" to cover similar experiences.)



[a Google Books Ngram view](#)

The definition of empathy is contested.^[2] What sort of thing empathy is also defies agreement: is it a sense or emotion (an immediate and visceral mirroring of another person's state), or is it more like an intellectual feat (accurately discerning another person's viewpoint), or maybe a social skill (the ability to respond appropriately to another's condition)?

Some authors narrowly define empathy as a sense or feeling, like "the coexperience of another's situation"^[3] or "feeling what you think others are feeling."^[4] Others apply the term to a constellation of feelings, cognitive interpretations, and responses, for example "a social and emotional skill that helps us feel and understand the emotions, circumstances, intentions, thoughts, and needs of others, such that we can offer sensitive, perceptive, and appropriate communication and support"^[5] or "the art of stepping imaginatively into the

shoes of another person, understanding their feelings and perspectives, and using that understanding to guide your actions.”^[6]

In some descriptions of empathy it is restricted to feeling distress upon noticing someone else exhibiting distress in circumstances such as pain, fear, or loss. But in others, you might just as easily empathize happily with someone experiencing more enjoyable emotions like joy or triumph (“I’m so proud of you!”), or empathize curiously with someone who is not experiencing any strong emotion at all but whose perspective you want to explore for other reasons (“I wonder what’s on his mind.”).

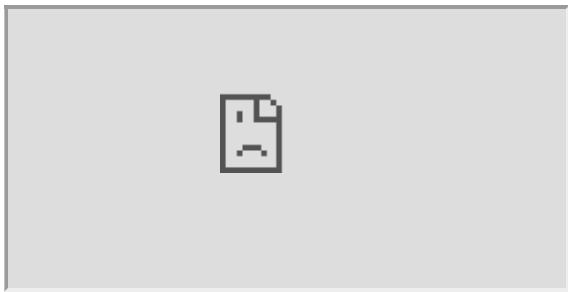
If all of this weren’t confusing enough, authors will sometimes slide between different definitions without seeming to realize that they’re doing it. (I’ll probably end up doing this too.)

There is plenty of interesting speculation and research about the evolutionary pathways that led to human empathy, the neurological correlates to empathic behavior, the expression of empathy (or its building blocks) by other animal species, and the developmental stages of empathy in children. I’m mostly going to leave such stuff out of my summary of the topic here, which is going to focus on how adult humans can characteristically empathize well. One important take-away from such research, however, is that different components of empathy have different foundations in development and in the brain. So people will probably differ in which components they are better or worse at. If I find myself to be unskilled at empathy I might want to look closer at which specific component seems to be lagging. And if I try to strengthen my empathy, I may have more luck if I look for ways to strengthen individual components.

People sometimes speak of empathy as a perception that operates automatically and subconsciously—or even as a variety of emotion. Other times empathy is described as a deliberate and cognitive skill of creating and working with mental models about other people. Some models incorporate both of these possibilities, working in tandem. I will also be considering empathy as a variety of *virtue*, which puts it more in the deliberate-skill category (a virtue is also a *characteristic* skill). So I want to examine (a) whether characteristic empathy is good for you (is it part of how to flourish as a human being); (b) if so, how is empathy best practiced; and (c) what practical steps can you take to get better at it.

An [applause light](#) often accompanies the word “empathy.” Authors who write about empathy sometimes gush about other things that fall into the halo surrounding empathy—like being kind, civil, tolerant, cosmopolitan, peaceful, or gentle—seemingly without realizing they’ve changed the subject. A lot of short-form popular writing on empathy doesn’t look at it critically, but assumes that it is good, full-stop, and that everybody just needs more of it. But on closer examination empathy seems not to be a more-is-better sort of virtue, but one that follows the Aristotelian “golden mean” theory: virtuous empathy is empathy characteristically practiced at the right times, in the right ways, and to the right extent. Even Karla McLaren, the most enthusiastic empathy booster of those I read, agrees that “there’s a sweet spot with empathic skills” between too little and too much.^[5]

Just about everybody empathizes. One of the more empathy-skeptical writers I read nonetheless concluded that “If there were people lacking empathy completely, we would not recognize them as people.”^[3] Indeed, people seem to be “hyperempathic.” It is a little weird how eager we are to empathize, for instance when we invent gods and spirits whose moods and whims we intuit to explain natural phenomena, or when we attribute emotions and motivations to things we have no reason to believe possess any, such as animated geometric shapes:



What are the components of empathy?

There seems to be some agreement that empathy is not a simple thing, but is a composite of things. There is less agreement about what those things are, or which things are part of empathy and which ought to be carved off and assigned to something else.

One common way to divide empathy is between its affective ("experience sharing," "emotional empathy") and cognitive ("understanding," "mental state attribution") aspects.^[7]

Some authors restrict the term "empathy" to the affective aspects.^[8]

Others add a third category that I'll call "behavioral empathy" (a.k.a. "empathic concern"^[9] or "perceptive engagement"^[10]). This concerns how, once you have felt and understood another person's point of view, you make use of that to take some appropriate action.

Affective empathy

The affective aspects are usually thought to be temporally, developmentally, and evolutionarily primary. They begin with simple mirroring or mimicry of the current emotional state of another person. You see them startle, you get frightened too (before you know what you're supposed to be afraid of). You see them crying and tears come to your eyes (before you know what's wrong). You laugh along at [the punchline](#) (before you get the joke). You watch [Two Girls, One Cup](#) and find yourself gagging as if you were one of the stars.

(Inevitably, "[mirror neurons](#)" are invoked to explain some of this, so I feel obligated to mention them here. These are neurons, first discovered in macaque monkeys, that fire both when a creature performs some particular action and when they view someone else performing that same action.)

Sometimes this mirroring is called "emotional contagion," but other authors use that phrase to refer only to a pathological empathic overreaction in which you become so overwhelmed by an empathic emotion that you lose track of where it came from and feel it as though it originated with you. (Still other authors recognize that same pathology, but give it different labels, like "empathic distress" or "personal distress.") In any case, mirroring primes the pump for the rest of the empathic process. You begin by mimicking another person's affect, and then, by perceiving what feelings this mirroring evokes in you, you get some insight into what feelings might be going on in the other person.

People can get tripped up at this early stage of empathy. If they find empathically-evoked emotions overwhelming or unwelcome, they may try to escape from them and to suppress the affective empathy response. For example, people who have learned aversion to certain emotions, or people who learned from bad childhood experiences that emotional extremes in others can be precursors to abuse, may react in that way. People with heightened sensitivity to emotional stimulus may also find affective empathy overwhelming. Karla McLaren

theorizes that the empathy deficits measured in people with autism spectrum conditions may be caused this way.^[5]

A subset of “empathic accuracy” is sometimes teased out from the rest of affective empathy. ^[5] When you mirrored the other person’s affect, did you hit the target? Your empathy can get off to a bad start without this sort of accuracy: They laugh nervously, you laugh uproariously; they are pleasantly surprised, you become bewildered; they express amused chagrin, you become indignant on their behalf. If certain emotions trigger things in you idiosyncratically (e.g. fear⇒panic or embarrassment⇒shame), this can also interfere with empathic accuracy.

Elizabeth Segal adds “affective mentalizing” as a subcomponent of affective empathy. ^[10] This additional step provides more raw material for cognitive empathy to operate on. It involves appraising another’s emotional state by pulling together additional clues like body language, facial expressions, knowledge about the other’s beliefs & situation, and context (what just happened). This process can cause you to engage in an additional cascade of mirroring behavior. You can also summon affective mentalizing deliberately, without the trigger of an initial acute emotional outburst from someone else, just by deciding to pay closer attention to someone. When you develop an emotional response to a character in a story or a daydream, for example, you are doing so by means of affective mentalizing.

Cognitive empathy

“When I condole with you for the loss of your only son, in order to enter your grief, I do not consider what I, a person of such a character and profession, should suffer, if I had a son, and if that son was unfortunately to die; but I consider what I should suffer if I was really you; and I not only change circumstances with you, but I change persons and characters. My grief, therefore, is entirely upon your account, and not in the least upon my own.” —Adam Smith^[11]

Cognitive empathy also has a few subcomponents.

“Emotion regulation” is a damper on the emotional contagion process.^[5] It allows even an alarming empathic emotion to appear as a cool blue info-light on your dashboard rather than a blinking red warning. Without emotional regulation, you might let such an emotion run away with you in a way that doesn’t do you or anyone else any good.

You will sometimes hear people described as “extremely empathic” because they overreact to the emotions of others—going into hysterics at another person’s distress, for instance. But this is better thought of as poorly-calibrated empathy. Some people respond to others’ pain/distress/etc. with avoidance, anti-social behavior, aggression, or personal distress, any of which demonstrates that something is amiss here.^[12]

“Self-other awareness” enables you to empathize with another person while keeping your sense of self intact. It helps you realize that you are empathizing—that you are experiencing a borrowed emotion—and that you can therefore more dispassionately decide how much you want to get wrapped up in it. Without it, you can become mistaken about the nature of your feelings and your empathy can become narcissistic (“look at all the suffering your distress is causing me”).

“Perspective taking” is when you try to imagine the other person’s experience from their point of view. The first part of this is to shift from considering how you feel *about* what they feel to considering *how they feel*. This is the difference, for example, between being happy that your child is happy because it’s more pleasant to be around a happy child and being happy that your child is happy because you share in their happiness.

(This is where [“theory of mind”](#) tends to enter the discussion, but I’ll try to avoid the temptation to go down that rabbit hole.)

There are two ways to go about this shift in perspective. The first and easiest is to imagine how you would feel *if you yourself* were in the other's situation. But this can be misleading. [13] Another approach is to imagine how you would feel *if you were them* in their situation. That requires you to engage in some difficult and speculative modeling of what it is like to be the other person, but as a result may give you a more accurate model of them to work with.

If you know the other person and their agenda well, taking their perspective is easier. If they're a stranger to you, you have little to go on but your model of a typical person, combined with whatever stereotypes you have about people-like-them, and whatever clues you have picked up during "affective mentalizing." This suggests that you will be better able to empathize accurately if you have better models of people in all their rich variety: which things are peculiar to you and which are more general to humanity, in what ways people are diverse, what sorts of traits and attitudes tend to cluster in people, and so forth.

Knowledge of context and social nuance comes in handy at this stage too. What signals are people putting out? Are subtle things—like plausibly-deniable slights, back-handed compliments, or damning with faint praise—coming into play? If you do not have a sophisticated understanding of social dynamics, it can be easy in some circumstances to draw the wrong conclusions about another person's state and what triggered it.

It's important to distinguish all of this from a colder sort of assessment: "so-and-so is in such-and-such a situation, and so I intend to feel for them in some way." In the empathic process, you go through at least an initial phase in which you feel, not *for* but *with* them. In the cognitive phase of this process, you may begin to move from this feeling-with to deciding on an appropriate feeling-for (e.g. sympathy, compassion, pity, concern), which leads us to behavioral empathy.

Behavioral empathy

So you are experiencing empathy, you feel what you believe to be a reasonable facsimile of what someone else is feeling... now what? Part of the promise of empathy is that this insight into the other person's headspace will help you to come up with a way to interact with them that optimally fits the situation. Perhaps they have a need you could help them to meet, or perhaps they're about to blow their top and you should beat a hasty retreat, or perhaps they're feeling generous and now would be a great time to ask for that favor.

Typical examples of empathy-provoked responses include [compassion](#), [kindness](#), [care](#), and consolation. These are often altruistically helping behaviors prompted by another's distress.

One theory about this is that empathy prompts us to help a person who is experiencing a distressing emotion because we find empathically sharing that emotion to also be distressing and we'd like that to stop. This theory leaves it something of a mystery why we should *help* rather than *stop empathizing*, which would be as effective for that purpose and probably simpler. Empathy seems sticky: you can turn away but you can't get it out of your head; you *could have* helped and that bugs you. Experimentally, even when people are given the option of an easy, clean exit from an empathically distressing situation, they still often opt to help instead. [14]

Paul Bloom believes a better explanation is that empathy helps to trigger preexisting humane motives (or, perhaps, "virtues"). Empathy does not goad us to be compassionate in order that our compassion will indirectly relieve our empathic distress, but it makes us aware that an occasion for compassion has arisen so that we can rise to the occasion if that is our wont. "It's not that empathy itself automatically leads to kindness. Rather, empathy has to connect to kindness that already exists. Empathy makes good people better, then, because kind people don't like suffering, and empathy makes this suffering salient. [In contrast, if you made a sadist more empathetic, it would just lead to a happier sadist...]" [15]

Another example of an empathy-provoked response might be [forgiveness](#), which is easier for us to give when we empathically verify the other person's remorse. And, from the other direction, apology is more effective if it appears to come from a place of vivid awareness of the harm caused (not just from an abstract acknowledgment of wrongdoing), and you can better understand that harm if you feel it empathically.^[16]

Social empathy

Elizabeth Segal would like to add "social empathy" to the list. Social empathy involves a more sophisticated understanding of large-scale social dynamics: how people's perspectives are affected by things like class, race, culture, history, systemic barriers, and so forth. It gives you the "ability to understand people by perceiving or experiencing their life situations and as a result gain insight into structural inequalities and disparities."^[17] It seems to be in part an attempt to shoehorn sometimes-parochial modern social justice concerns into empathy, and I was skeptical about how Segal et al. went about it.

A proposed Social Empathy Index test includes questions like "I think the government needs to be a part of leveling the playing field for people from different racial groups," "I believe adults who are poor deserve social assistance," and "I believe that each of us should participate in political activities," that seem to uncritically import certain contested ideas into the definition of social empathy.^[18] (For example: What if I think I should give a hand to adults who are poor even if they don't "deserve" it? Does that make me more or less socially empathic? What if I think they could probably use financial assistance more than "social assistance" whatever that is?) Another such question—"I believe government should protect the rights of minorities"—seems to me like it only potentially measures *empathy* for others if the answerer doesn't consider themselves one of the minorities who would be protected (that Latinos in the U.S. have higher SEI scores than Whites shouldn't be surprising even if you suspect both groups have a similar amount of "social empathy," given questions like that).

In part, "social empathy" is meant to address well-documented common biases in human empathy, such as those that favor ingroups over outgroups. "How do we cultivate the type of empathic reactions that people demonstrate toward friends and family members in their responses to groups who differ from them, particularly groups that have historically been the focus of prejudice and oppression?"^[19]

Other models

Fritz Breithaupt has an idiosyncratic model for empathy.^[20] If I understand him right, he says that a typical person has most of the various "abilities and mechanisms related to empathy" described in the Affective, Cognitive, and Behavioral sections above, "and a [hyperempathic] tendency to use them;" but also has a vigorous set of empathy suppression mechanisms. These suppression mechanisms operate at multiple levels, from the neural, to the conscious (e.g. evaluating when someone is emotionally manipulating us; deciding when someone had-it-coming), to the societal (e.g. learning which animals we ought to feel empathy toward and which ones we can mercilessly chop up for bait). They prevent empathy from disturbing us unless certain conditions are met. In Breithaupt's model, these suppression mechanisms are as important to well-regulated empathy as any of the positive abilities and mechanisms, but they also have biases and idiosyncrasies.

How is empathy measured?

There has been a lot of research into empathy, of a more-or-less scientific nature. Much of the challenge of such research comes from how to measure empathy. A typical approach is to give people self-assessment questionnaires in which they are asked how much they feel

certain descriptions apply to them (e.g. “When someone else is feeling excited, I tend to get excited too” or “I become irritated when someone cries”).

But this means researchers are not measuring empathy itself, but how much people report that statements about empathy apply to them. This sort of indirect measurement is flawed. It is not unheard of for people to have inaccurate self-images, for example. And people may respond to such questions aspirationally (what they would like to be, or what they think they should be) rather than accurately.

One puzzled researcher found that students scored as *less* empathetic on such an empathy questionnaire after taking her empathy class. Why? After learning more about empathy, the students became more exacting in how they judged their own feelings and behavior, expected more from themselves, and so rated themselves more severely.^[19] So someone with a low “empathy quotient” might either be someone who has little empathy or someone who has plenty but has even higher standards.

Similarly, when gender differences in empathy scores are found, it turns out to be very difficult to tease out whether this is because one gender is really more empathic than another, or whether people are just conforming to gender expectations in their self-image or in how they think it is appropriate to present themselves.

Another problem is that as the science of empathy has matured, and additional facets of empathy have been teased out or theorized by researchers, new questionnaires have been developed to try to capture these nuances. There are now multiple competing empathy questionnaires, which can make it difficult to compare results from study to study.

Even when the same questionnaire is used, it may not really measure the same thing over time because language changes. One study, which used a single questionnaire to measure 13,737 college students between 1979 and 2009, found that empathy measures had been falling substantially over that time.^[21] The press of course went with the “kids these days” angle.^[22] But one critic noted the results might be explained by the fact that the wording of the questionnaire had become anachronistically quaint (it used idioms no longer in common use, like “tender feelings”, “ill at ease”, “quite touched”, or “go to pieces”), and today’s students might not identify with such statements for that reason.^[3]

Other ways to measure aspects of empathy include fMRIs and things of that sort, in which subjects are asked to perform empathy-related tasks while their brain activity is measured. Researchers can monitor different sorts of people (e.g. sociopaths, autistic people, children, meditators) or people in different circumstances or undergoing different interventions to see if their brains light up in different ways, for what that’s worth.

Some researchers have tried to measure empathic accuracy by filming one person, later asking that person what was going through their mind at various points during the filming, asking a second subject to watch the film and to try to empathically discern those subterranean thoughts and feelings, and then comparing that person’s answers to what the first person reported. It’s a complex experiment design that leans heavily on human memory, introspection, language-use, and subjective judgment. But it may be as close as we can expect to get to measuring and comparing the subjective and ephemeral.

Is empathy altruistic? is it moral?

Paradigmatically, empathy is a tender feeling toward someone else in crisis, followed by consolation or kind assistance. You feel someone’s pain and then you work to relieve it, for their sake. Is empathy essentially altruistic? [This is also contested](#). It turns out to be pretty easy to tell a story in which seemingly altruistic acts turn out to be motivated primarily by self interest after all. But it is difficult to test whether this says as much about the seemingly altruistic acts as it does about our story-telling abilities.

"You say: How sad to think that the noblest altruism is, after all, merely a refined kind of selfishness.

"I say: How good to think that selfishness, when it is purified and stops being stupid, is exactly the same thing as the noblest kind of altruism."

—Pierre Cérésole[\[23\]](#)

Empathy has been proposed as the primitive foundation on which humans established moral ideas (Martin Hoffman calls empathy the “bedrock of morality” for example).[\[24\]](#) Once empathy allowed us to see things from other people’s points of view, we then became able to entertain ideas like The Golden Rule, mutual tolerance, and so forth.

This further extends into political rights: If other people have aspirations, needs, etc. just like I do, they perhaps ought to have rights that I should respect, and vice-versa. Jeremy Rifkin for this reason called empathy the “soul of democracy.”[\[25\]](#)

Empathy skeptics, however, argue that because of the strong, demonstrable, and not particularly defensible biases of human empathy, it makes a poor foundation for morality or for political rights (see below for more about this). But Roman Krznaric points out that in historical accounts of societal moral progress, such as the abolition of the slave trade, successful appeals to empathy can seem to have played a crucial role. “Empathy and reason are not polar opposites, as critics like [Paul] Bloom would have us believe, but rather mutually reinforcing ideals on which we can build a more humane civilization. Indeed, it is ‘the gut wrench of empathy’ [quoting Bloom] that forces open the door of our common concern—and only then does reason have a chance to wedge it open with laws and rights.”[\[26\]](#) While it may be important that we use reason rather than biased empathy to make decisions about the worth of human lives, “the explanation for *why* we believe all humans should be treated and valued equally... is because empathy has made us care about the plight of strangers outside our local community.”[\[26\]](#)

Filtered empathy

Fritz Breithaupt is one of those who disagrees with the assessment that empathy is altruistic. “Like most other human abilities, empathy probably serves the empathizer first and foremost and not the target of empathy.”[\[3\]](#) He believes that empathy is better understood as a self-serving feeling: a mostly aesthetic indulgence and a way of satisfying our curiosity about others.

Breithaupt points out that in humanitarian empathy in particular it’s questionable whether we really empathize with those in need or rather with a (perhaps imaginary) rescuer whose hero role we empathically inhabit. Such empathy allows us to feel not the suffering of someone in crisis, but the praise due to the rescuer, and so makes empathy more alluring—we don’t feel bad for them but good for us. He calls this “filtered empathy”—empathy that is absorbed by a third party before reaching its ostensible target. (This may help to explain the common Hollywood trope of telling stories of struggling people in a roundabout way through outsiders who intervene to help them, e.g. *Schindler’s List*, *Amistad*, *Avatar*.)

This also may explain why we sometimes empathize with people we cannot possibly help (victims we hear about on the news, fictional characters, historical figures). Such empathy can’t culminate in helpful action, but if Breithaupt is correct, that is not its purpose: instead it culminates in a pleasant [Walter Mitty](#)-like daydream along the lines of “if only I had been there, I would have helped heroically.”

Empathic accuracy

Native Mandarin Chinese speakers overestimated how well native English-speaking Americans understood what they said in Chinese, even when they were informed that the listeners knew no Chinese. These listeners also believed they understood the intentions of the Chinese speakers much more than they actually did. This extreme illusion impacts theories of speech monitoring and may be consequential in real-life, where miscommunication is costly.^[27]

When I empathize with someone, I have the sense that I really feel what they feel. Part of this comes from how visceral the emotion is: it doesn't seem like I'm making up a story but like something has affected me. But do I have good cause to believe my empathetic feelings are any more accurate than any other guesswork I might engage in about the contents of other people's heads?

Is there even theoretically enough information available to a maximally astute observer to accurately surmise the inner state of the person observed? How transparent are we, really? People often seem to throw up clouds of ink in the waters around them, and can be mysteries even to themselves.

Even given the information that is available, just how clever can we expect imperfect mortals to be in putting it all together? Can we trust our intuition about how well we do at this complex task?

[Correspondence bias a.k.a. the fundamental attribution error](#) is one example of how we predictably mislead ourselves by observing others and then jumping to conclusions about their motives and outlooks.

Even assuming you have good, representative, unambiguous data about another person, and you use that data wisely to come up with a good model of how they work, is that anything like understanding how it feels to be them? [What is it like to be a bat?](#) I necessarily use myself as the only available model for what a subjectivity feels like, and so [I'm bound to make errors if I try to generalize from my n=1](#). Every once in a while I learn something new, for instance that [some people don't have an inner monologue](#), and I have to throw out big hunks of my model of what makes other people tick. [There's lots of stuff like this, apparently.](#)

Fritz Breithaupt is among those who think empathy is deceptive in this way. An empathizer simplifies the situation of the person they are empathizing with to what seem to be the most relevant features; to that other person, the situation is messier. “[T]he feelings of the other person become a fact to [the empathizer], appearing transparent and perceptible.”^[3] This can give the empathizer “a clarity not available to the other”—a clarity that may sometimes even be helpful, if the person being empathized with is lost in a thicket of details and uncertainty.

But this oversimplification may just be true of rudimentary, naive empathy. Another author stressed that more proficient empathy doesn't have to be so reductive. “Empathy means acknowledging a horizon of context that extends perpetually beyond what you can see... Empathy means realizing no trauma has discrete edges.”^[28]

Empathic accuracy is hard, but I don't want to just give up. Awful things can sneak in through that door. For instance, you can justify doing cruel things to others if you can convince yourself that there's just no way of knowing whether it's really cruel from their point of view: e.g. “the Arab mind is different from ours,” “homeless people probably just like to live that way,” or “animals probably don't really suffer.”

Karla McLaren recommends that to improve the accuracy of your empathic sizing up of someone else, you just frickin' ask them. “The way to gauge your Empathic Accuracy is both very simple and infinitely hard: you ask people if what you're sensing from them is true.”^[29] You're never going to become a supernatural mind-reader. People exhibit themselves in diverse and complex ways, and to some extent each person has to be

interpreted anew as if they were written in a foreign language. Your guesses will be prone to errors, but you can always go to the expert: the person whose feelings you are trying to discern. McLaren recommends that you ask like this:^[30]

“When you [*objectively described behavior*], it seems to me that you [want/feel x]. Did I get that right?”

For example:

“When you cross your arms like that, it seems to me that you feel impatient. Did I get that right?”

(She stresses that it’s important that you not import your guesses about their feelings or inner state into your description of their behavior. So, for instance “when you gaze out the window...” is okay, but “when you start ignoring me...” is less helpful.)

When we learn that our empathetic assessments are incorrect, we can recalibrate and bring the other person into better focus. Unfortunately, according to Fritz Breithaupt, “false empathy is a powerful drug” and it can be difficult for people to abandon bad guesses once they’re empathically established.^[31] Sometimes people respond to being told their empathic accuracy is off by accusing the person they are empathizing with of not knowing their own mind (this isn’t to say all such accusations are false, but it does take some chutzpah to assert that I know your feelings better than you do), or by getting angry at them for not feeling the way they’re supposed to or for not being sufficiently legible.

Relation to other virtues

Empathy can contribute to other-attending virtues like [compassion](#), [kindness](#), [care](#), sympathy, pity, consideration, [courtesy](#), *nying je*, consolation, altruism, recognition, [respect for others](#), persuasion/education/tutoring, [loyalty](#), [amiability](#), and connection.

It can be helped by [mindfulness/attention](#), emotional intelligence, curiosity, and imagination. Roman Krznaric thinks “sheer [courage](#)” can come in handy too, for instance in asking difficult questions that can improve your empathic accuracy.^[31]

There is some tension between empathy and objectivity, [impartiality](#), [fairness](#), and [justice](#). On the one hand, it can exacerbate biases that hurt such things; on the other hand it may help us more vividly see “both sides” of a case and may assist mercy and *epeikeia*.

The good and bad of empathy

There is extreme disagreement about the value of empathy: ranging from Karla McLaren, a self-described hyperempath, whose book calls empathy “Life’s Most Essential Skill,”^[5] to Paul Bloom, whose *Against Empathy* says we’d be “better off without it.”^[4] I’ll start with the case in favor of empathy.

What good is empathy?

“If there is one emotional intelligence skill that we would recommend developing, it’s definitely empathy. Empathetic people are happier, more self-aware, self-motivated, and optimistic. They cope better with stress, assert themselves when it is required, and are comfortable expressing their feelings.” —Dr. Ilona Jerabek, president, PsychTests^[32]

Empathy is (nearly) universal in humans. It develops in children in a predictable way at an early age. Forms or components of it are found in other animals. All of this gives the impression that it's something that's been selected for, and so is probably good for us at least in the for-our-reproductive-success sense of "good." Why might this be? Here are a few common theories for how empathy might have been selected for:

- Mirroring allows us to react more quickly to threats. When we see a fight-or-flight reaction in someone else, we can get our guard up before we know why.
- Empathy helps us intuit and meet the survival needs of our ridiculously inept children.
- Empathy also helps us as children to understand and manipulate the adults whom we rely on to get our needs met.
- Empathy helps us navigate our social environment. It tells us whether someone means us help or harm, helps us better detect deception, helps us discern who is allied with whom, helps us distinguish an accidental jostle from a hostile poke, and so forth.

If you mimic the posture, facial expressions, and vocal style of someone you are with, you can thereby encourage them to help you and to form a favorable opinion of you.^[33] Savvy business people consciously take advantage of this in job interviews, negotiations, and sales pitches. Police interrogators use this to build rapport with suspects. This may seem pretty far-removed from what we usually think of as empathy, but it may be more evidence for how the early mimicry/mirroring stages of empathy can help strengthen one's social position.

Empathy seems to improve the quality of social relationships. People who score more highly on empathy questionnaires also report having more positive relationships with other people.^[34] People also tend to value empathy in their friends and romantic partners. A friend goes up a notch in my book if it seems they "get where I'm coming from," and I think demonstrating empathy is often an effective way for me to say that I care about a relationship and want to deepen it. Empathy might also improve your gift-giving proficiency.^[35]

Being empathic seems to correlate with feeling better about your life. In one longitudinal study of adults, researchers found that "[p]eople with higher empathy scores reported greater life satisfaction, more positive affect, less negative affect, and less depressive symptoms than people who had lower empathy scores."^[36]

Children who exhibit more empathy also have more resilience.^[37] Empathically sampling other people's situations and emotional states may help you to prepare for such things in your own life. This may also help you to learn culturally-legible ways to request an empathic response from others. When I roll my eyes and throw my hands up to heaven, I reenact a theatrical way of broadcasting "empathize with me in my exasperation!" that I had to learn somehow.

Thinking about other people's problems can be a welcome distraction from your own (or can put your own into perspective), thus making you feel better. In one study, for example, peers who helped other peers with their depression found their own depression symptoms improving.^[38] (Another explanation, but one that also favors empathy, is that when you give advice to others you may realize it's good advice and take it yourself, but you wouldn't have come up with it on your own for some reason.)

Empathy helps you to [read the room](#), which can make you more courteous, persuasive, and so forth.

Empathically feeling unusual emotions in others (or in fictional characters) helps us to recognize them in ourselves and broadens our emotional intelligence. The more emotions we encounter in others, the more we are able to associate certain behavior and visible bodily changes with those emotions, and the better emotional vocabulary we develop. That helps us to potentially better understand ourselves, and also feeds back to improve our empathic accuracy.

When we look at the world through someone else's eyes, we can use this as a mirror to look back at ourselves and get some idea of how we appear to others.

Empathy can be an aesthetic pleasure, "by widening the scope of that which we experience... by providing us with more than one perspective of a situation, thereby multiplying our experience... and... by intensifying that experience."^[39] It can sometimes be harmlessly pleasurable, for instance when we take joy at the joy of children discovering things or playing make-believe, or when we use empathy to satisfy our curiosity about other people's lives.^[40]

Paul Bloom suggests that empathy might work best in a backwards way: rather than a empathic person becoming distressed by seeing another person in distress, a person in distress may look at a bystander who is not so freaked out and become calmer as a result.^[16] "Calm is contagious," as the saying goes. For example, Leslie Jamison spent some time as a "medical actor" helping to teach med students to demonstrate empathy toward their patients. When the time came for her to be a genuine patient, she realized what she wanted from her doctor was different from that sort of empathy: "I wanted to look at him and see the opposite of my fear, not its echo."^[41]

Religious applications

Empathy is occasionally used to heighten religious feelings. Group worship and ritual (and chanting and song) can have the effect of synchronizing the worshipers' emotions and outlook.

But there are also some practices that seem to use empathy to strengthen communion with the divine. For example, Christians may [imagine themselves on the cross](#), in all of that gory awfulness, so as to better appreciate Christ's sacrifice. Some Christians are said to have developed sympathetic wounds ([stigmata](#)) as the result of intense contemplation of the crucifixion. In some varieties of Buddhist tantric meditation, the meditator tries to [merge their identity with a deity](#) so completely as to lose their own identity.

Business applications

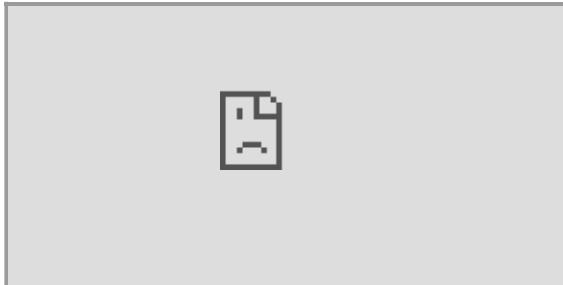
Empathy has become a business buzzword. It is supposed to have applications in management and marketing for example. *Harvard Business Review* (US) and The Empathy Business (UK) have produced "[most empathetic" rankings](#) of businesses. The "leadership consulting firm" DDI found in 2016 that 20% of U.S. employers offered empathy training to managers.^[42]

If you can empathize with your (potential) customers, identify their problems, and come to understand how you can help them, you presumably can develop new products & services for the niches you thereby identify. This can be done through field observation & study of (potential) customers, and by empathically trying to get into their heads: why do they act the way they do; what are they trying to accomplish; how might we help them meet their goals more effectively? "It's a process informed by deep qualitative data rather than statistical market data."^[43]

[Patricia Moore](#) was a pioneer of this technique. In one example, she used makeup and prosthetics to simulate the experience of elderly people, and used the insights she gained from this to inspire friendlier products for that customer segment.^[44] Design engineers at Ford Motor Company wore prosthetics to simulate effects of pregnancy and of old age, in order to help them design cars that would work better for a broader set of customers.

[\[45\]](#) Fidelity uses a virtual reality training application to put its phone bank "associates" in a (dramatized) customer's home so they can see what it is like to be on the other side of their conversations.^[46]

Customer empathy is exploited in advertising, which often invites us to empathize with characters in brief vignettes (who eventually buy product or service x and feel glad they have done so). See for example the old television ad below. It tells the story of a protagonist who meets and overcomes a challenge. We are meant to empathize with her in her struggle (and to imagine successfully seducing our husbands back with Yuban).



Charity marketing is particularly sensitive to the empathy⇒compassion⇒generosity pipeline. It is notorious for the sometimes ham-fisted ways it plays to the biases that accentuate human empathy, for example our propensity to empathize more with individuals than groups, or with the adorable over the homely:



We can expect marketing to become increasingly manipulative in this way, especially now that AI is being trained to intuit and respond to human emotions.^[47] After all, even [ELIZA](#) was oddly engaging.

What about psychopaths?

A lack of empathy is a defining feature of the [psychopath](#). Because of this, sometimes the psychopath is trotted out as an example of what empathy is good for (the implication being that lack of empathy is what made the psychopath the way they are).

But this is less clear-cut than it may seem. For one thing, the logical structure of that implication is faulty. (A lack of hair on the scalp is a defining feature of alopecia, but avoiding alopecia is not a good reason to let your hair grow long.)

For another, it isn't entirely clear what the nature of the empathy deficit is in psychopaths. There is some evidence that psychopaths are perfectly capable of empathizing, but just don't typically care to.^[48] Psychopaths can be very manipulative and deceptive, which means they likely have good models of other people's emotions, points-of-view, and so forth; they typically do fine on theory-of-mind tests.^[49] They just don't seem to give a damn about anyone else. On the other hand, as I was putting the finishing touches on this post,

researchers announced that they found people with psychopathic tendencies are less likely to yawn contagiously, implying that the deficit may prefigure affective empathy.^[50]

The abhorrent behavior we associate with psychopaths may have less to do with their supposed lack of empathy than with other things that diagnosed psychopaths typically have in common (such as poor impulse control, criminal history, and low emotional engagement).^[4]

What **bad** is empathy?

“The problems we face as a society and as individuals are rarely due to lack of empathy. Actually, they are often due to too much of it.”

“If you are struggling with a moral [decision](#) and find yourself trying to feel someone else’s pain or pleasure, you should stop.” —Paul Bloom^[4]

Empathy boosters invited a backlash with some of their over-hyped claims. From reading Roman Krznaric’s *Empathy: Why It Matters, and How to Get It*, for example, you would think that without empathy, mothers would let their children starve, charities would dry up, and nobody would help anyone else or commiserate with suffering friends. Without empathy, we would be damned to a “heartless world of indifference”,^[51] and yet even now we are plagued by an “empathy deficit” and “epidemic of narcissism” that explains everything from Syrian civil war to child-molesting priests to insufficient action on global warming.^[52] “Empathy is like a universal solvent,” wrote researcher Simon Baron-Cohen. “Any problem immersed in empathy becomes soluble.”^[53]

So now there are many strong and persuasive criticisms of empathy. From the point of view of empathy-as-a-virtue it is important to discern whether these describe ways empathy *is* wrong, or just ways empathy *can go* wrong. If the former, we should consider dropping empathy from the list of virtues we aspire to; if the latter, we need to carefully attend to how to be empathic more wisely and well.

Note also that some of these criticisms hinge on a particular definition of empathy. Paul Bloom, for example, in *Against Empathy*, mostly draws the line against *affective* empathy, and has kinder things to say about being understanding (part of *cognitive* empathy) and compassionate (part of *behavioral* empathy), which he thinks are valuable things that should be insulated from (affective) empathy in order that they should operate more effectively.

Some of the criticisms also push back against the idea that empathy is altruistic, or is the wellspring of altruism. They insist instead that empathy is really mostly for the benefit of the empath. From my point of view, though—examining empathy as a virtue—this is not damning in the least. I should hope that empathy helps the empath to thrive: then I know I’m not barking up the wrong tree.

All that said, here are many of the criticisms of empathy:

Empathy may make you more vulnerable to emotional predators and parasites. Dishonest panhandlers, for example, rather than asking for money outright will sometimes concoct an intricate “sob story” to encourage you to cough it up. If you are used to taking guidance from your empathy, you may be more likely to fall victim to such cons.

If you are not sufficiently critical in your empathic response, you may get caught up in someone else’s enthusiasm or rage when you would be wiser not to. This too can make you vulnerable to manipulation, both by individuals and by institutions that want you to help further their agendas. For example, it’s common for sympathetic victims of particular (real or imagined) atrocities to be repeatedly and vividly referenced in the rhetoric of those drumming up wars, pogroms, moral panics, and the like. When rational appeals to justice

and interest fail to persuade, the unavenged suffering of that innocent child orphaned by the Hun might do the trick.

If you know you can evoke empathy in people, it can be difficult to avoid the temptation to play to the camera. Instead of feeling and behaving in an authentic manner, you feel and behave dramatically, with an eye to the effect this has on others. You see this even in young children, who may, for example, skin their knees and then look around to see if anyone else is watching before they decide whether or not to cry. (Is this a bad thing, or is it just a human thing? Is there such a thing as an authentic, genuine emotion that precedes its expression, or are emotions essentially signaling devices from the get-go?)

Empathy can discourage you from making necessary decisions that hurt or disappoint other people (for example, the decision to subject your child to a vaccination, or to say “no” when asked for an imposing favor).

Empathy may enforce conformity with other people’s standards, and discourage independent thought. If you would empathically feel someone else’s disappointment, disgust, judgment, or upset, in an unpleasant way, this may encourage you to conform to their desires and expectations instead, even when non-conformity would have been a better option.

The aesthetic appeal of empathy, or the emotional fix you can get from it, can be intoxicating and can sometimes cause what looks like pathological empathy-seeking behavior. This may be a factor behind things like [outrage porn](#), [poverty porn](#), and the popularity of videos capturing other people’s embarrassments, failures, and *Jackass*-style injury-inflections. On the other hand, this same appeal probably explains a lot of the attraction of literature and drama. For example: when you read a mystery story you want to share the curious befuddlement and eventual a-ha of the intrepid detective. That’s part of the fun.

Empathy can be a mirage. People seem prone to believe their empathic intuitions are more accurate than they really are (see “Empathic accuracy” above). If you act on the strength of such belief without adjusting for this bias, you may act poorly. If you have an exaggerated belief in your empathic powers, you may lose the curiosity you need to really know what’s going on with someone else.

Clumsily-expressed or overconfident empathy can seem presumptuous and condescending. A politician turns on the puppy-dog eyes and tells me “I feel your pain.” Do you really? Leslie Jamison, who acted the part of a patient to help physicians in training with their bedside manner, rankled at the awkward tropes of empathy they were trying out:

I grow accustomed to comments that feel aggressive in their formulaic insistence: *that must really be hard* [to have a dying baby], *that must really be hard* [to be afraid you’ll have another seizure in the middle of the grocery store], *that must really be hard* [to carry in your uterus the bacterial evidence of cheating on your husband]. ... “I am sorry to hear that you are experiencing an excruciating pain in your abdomen,” one says. “It must be uncomfortable.” [\[54\]](#)

Empathy can be emotionally wearying, especially in the case of distressing emotions or situations. Medical caregivers for example may need to suppress empathy in order to avoid burnout. [\[55\]](#) It seems people only have so much empathy to give, and so they need to ration it. If you use up your empathy at work ([or in fretting over the benighted people of Borrioboola-Ghâ](#)), you might not have any left for your family. [\[56\]](#)

If you cannot control your empathy, you may instead turn away from things that could really use your attention, if those things trigger an unpleasant empathic response you want to avoid. And that can encourage you to come up with quasi-rational excuses for not caring about those things, which can make you irrationally callous or neglectful. A less empathic,

more level-headed problem-solving approach might have more staying power and therefore be more helpful.

To the Stoic philosophers, to condition your emotional disposition on the emotions or fortunes of someone else would seem to be foolishness of the first order. Cicero, summarizing (though not fully endorsing) the Stoic point of view, said that someone who feels distress at another's misfortune is committing as much of an error as an envious person who feels distress at another's good fortune.[\[57\]](#)

Is empathy unnecessary?

Empathy (feeling-with) is often defended as important in the development of compassion (feeling-for). Paul Bloom wonders why we cannot simply be compassionate without jumping through the biased and disorienting hoops of empathy first? When we are happy about our children's triumphs, why feel the need to first be triumphant when we can just be wholeheartedly glad for them? When someone close to us is sad or nervous, what's wrong with going straight to being consolatory or calming rather than mirroring their sadness or nervousness back at them first?[\[16\]](#)

Bloom shares the results of an fMRI study of a Buddhist meditator (Matthieu Ricard) doing "various types of compassion meditation directed toward people who are suffering" and reports that this "did not activate those parts of the brain associated with empathetic distress —those that are normally activated by non-meditators when they think about others' pain." The same meditator "put himself in an empathetic state" and got the more typical results, but also found it comparatively exhausting. Experiments on meditation-naive people in which they're trained in either compassionate or empathetic mediation apparently show something similar: the former is more pleasant and also leads to kinder behavior.[\[4\]](#)

Some researchers have had difficulty finding the expected evidence that empathy makes you a better person, or that lack of empathy makes you worse.[\[58\]](#)

Empathy does seem to be effective at goading people to altruism, but its biases and flaws mean that this altruism is often poorly-targeted and -executed. The cases of [Baby Jessica](#) who was rescued after falling down a well ("sympathetic strangers showered the family with teddy bears, homemade gifts, cards, and cash"[\[59\]](#)) and [the #Kony2012 phenomenon](#) are examples of empathy-induced altruism being both impressively strong and questionably-targeted.

Empathy is biased

Empathy is demonstrably biased in terms of whom we are likely to empathize with and in what situations and in what manner. These biases are hard to defend as bases for our compassion, kindness, respect, and so forth. If we rely on empathy to guide our decisions and priorities, we may unthinkingly import those biases. These are some of the biases that have been documented:

Empathy causes unconscious favoritism

If you feel empathy towards someone, this may cause you to practice unjust favoritism towards that person without recognizing that you are doing so.[\[60\]](#)

Empathy seems to encourage us to take sides with whomever we empathize with first. Fritz Breithaupt goes so far as to claim that "we do not act morally because we feel empathy; rather, we moralize to justify our quick and empathetic side-taking."[\[3\]](#)

This bias is subject to exploitation. When you hear a party to a conflict say something along the lines of "you can't stand aside at a time like this; you have to take a side!" you're also

being implicitly told “and try on my side for size first” which then stacks the deck. Once someone decides to empathize (with e.g. [Brett Kavanaugh or Christine Blasey Ford](#)), that person is likely to view emerging evidence from an empathically-biased perspective and to be less able to evaluate it on its merits. In this way, “empathy not only fails to end conflicts, but deepens them.”^[20]

Another way this plays out is that people who score higher on empathy scales are more likely to advocate harsher punishments of those who transgressed against whomever they are empathizing with.^[61]

Your empathy is likely biased against people who are different or who are in the outgroup

People empathize more with those who are like them in certain ways (such as the usual suspects of language, culture, race, and nationality).

This effect is even measurable at early, pre-cognitive stages of empathy. For example, it's easier to recognize emotions in ingroup members than in outgroup members (even when the ingroup/outgroup distinction is artificially imposed, not based on previous experience).^[62]

The “othering” that applies to empathy (and other things) is somewhat flexible. “For example, in research where temporary group identifications were arbitrarily manufactured, dominant group identity such as race became secondary.”^[63]

A good case can be made that this bias contributes to loyalty and group cohesion / coordination (and perhaps this is why we are biased in this way). In a conflict, you empathize most readily with those who are like you or in your ingroup, that triggers empathy's side-taking bias, and that causes you to line up with your squad in the conflict.

We are also less likely to empathize with people we envy or otherwise dislike.^[64]

Since empathy has biases that make you more likely to empathize with someone who is like you or for whom you have fonder feelings, when you empathize with someone you may thereby inadvertently suggest to others that you feel yourself to be like them or are fond of them. This can make it costly to (and can disincentivize you to) empathize with unpopular people. Why are you empathizing with that person convicted of possessing kiddie-porn? [You some kind of pervert-lover?](#)

The internet allows us to discover people who are uncannily like us in very specific ways. I wonder if this raised the bar for whom we see as similar-enough to empathize with? Since most people we meet aren't like the select people we've become companions with on-line, are we now more apt to find we “can't relate” to them?

Your empathy is subject to change with your social/political power

People tend to empathize less when they have more social or political power. Indeed lack of empathy may be a kind of status symbol.



Melania Trump visited a detention center for migrant children while wearing a jacket painted with the words “I REALLY DON’T CARE, DO U?”

This effect is most noticeable in extreme cases, such as when an abusive parent or a hostage-taker has arbitrary power over another person. [The Stockholm Syndrome](#) is one way this can play out. The vulnerability of the hostages drives them to extremes of empathic awareness so they can try to anticipate their captors’ actions. This has such a strong effect that the captives begin to sympathize with and defend their captors.

But this is also measurable in less-extreme circumstances. For example, people from lower-strata economic backgrounds exhibit better empathic accuracy (judging others’ emotions) than those from higher-strata economic backgrounds.^[65]

We should be cautious, given how some “priming”-style social science studies failed to replicate, but with that disclaimer attached: In a variety of experiments, people who were asked to recall a situation in which they had power over someone else then demonstrated reduced ability to mirror others, to comprehend their viewpoints, or to learn from others’ perspectives.^[66]

However for some people who attain power, “if they feel responsible for those who have less power or they already value being empathetic” they can be more than typically empathic, counter to the usual trend.^[67]

Given the various flaws and biases associated with empathy, maybe it’s good that powerful people typically don’t empathize as much or as well. Maybe this helps them make tough, rip-the-bandaid-off decisions that cause short-term pain for long-term gain. Or maybe their lack of empathy helped them make more rational decisions that also helped to empower them.

There are many suggestions for why people exhibit this bias, among which are these:

- If you're already high on the ladder, you don't need to put as much effort into understanding others. You can relax; you've made it.
- If you have power, you have authority and coercion on your side and you don't need persuasion as much, so you don't need to get inside others' heads to get them to do what you want.
- It's a valuable survival skill to understand the inner workings of those in the dominant culture if you're not in it; it's not so valuable to understand the dominated culture if you're in the dominant one.
- This could be an effect of a pyramid-shaped hierarchy with more powerless people and fewer powerful ones. It's easier for each employee to be concerned, say, about whether their manager woke up on the wrong side of the bed, than it is for the manager to keep track of all of the various headspaces of their many employees. (However, supervisors seem to score higher on accurately recognizing the emotions of others than their reports do.)^[68]
- Some leaders have a management style that involves giving direction in terms of vaguely-stated aspirational goals, and then humiliating and browbeating subordinates who disappoint them. The subordinates are thereby incentivized to "get out ahead" of what the leader explicitly asks for and to try to anticipate what will make the leader happy instead. This can require those subordinates to devote a lot of extra attention to empathic mind-reading.

Other empathy biases

People prefer to empathize in "clear, relevant, and decisive" situations.^[3] And we do so in a way that makes them clearer, more relevant, and more decisive—that is, in a way that papers-over ambiguities and nuances that don't fit a simple story.

We tend to empathize with the here-and-now, with what is immediately available to our senses. Empathy is good at prompting prosocial behaviors that are informal, unplanned, and directed at someone right here right now, but not very good at prompting things that require more abstract, long-term concern (like giving blood, donating to charity, or volunteering).^[69]

We also have a harder time empathizing with what is not immediately apparent (for example, with the struggles of someone with an outwardly invisible brain injury). Empathy can operate on the "out of sight, out of mind" principle.

Empathy operates best at the level of the individual, and so we may empathize with one person facing a plight, but if we hear about multiple people undergoing a similar plight, empathy has a harder time gaining a foothold. (Mother Theresa: "If I look at the mass, I will never act. If I look at the one, I will.")

Empathy favors known/nameable/picturable individuals over unknown/anonymous/invisible ones. An example of how this can pervert decision-making is vaccine reluctance. If some child has a bad reaction to a vaccine, that child is known, has a face and a name, and can therefore be more-easily empathized with. But if a child avoids catching a deadly or crippling disease thanks to vaccines, well, can you even point to them? They're a statistical projection, barely even a rumor so far as empathy is concerned.

People with different levels of experience empathize differently: A child may not empathize with a situation because they do not yet know what its relevant features are and have not had similar experiences of their own to compare it to, while a world-weary person may find it hard to empathize with the anxiety and surprise of a tyro.

We empathize more with people in distress whom we feel are "innocent"—that is, they were not responsible for bringing the distress upon themselves, or in any case that distress is disproportionate to any cause they participated in.^[70]

We are less likely to empathize in “hopeless cases” of distress, in which we don’t think anything can be done to make it better.

It may not be realistic to say “be empathetic, but try to be on guard about your biases.” Paul Bloom thinks advice like that is hogwash, especially when people propose deploying empathy to *mitigate* the harmful effects of ingroup/outgroup biases. “Asking people to feel as much empathy for an enemy as for their own child is like asking them to feel as much hunger for a dog turd as for an apple—it’s logically possible, but it doesn’t reflect the normal functioning of the human mind.”^[71]

Vampiristic and sadistic empathy

Fritz Breithaupt highlights two harmful forms of empathy.

One, he calls “vampiristic empathy.”^[72] This sort of empathy is more aggressive in how it attempts to import the experience of another person. The empathic vampire is not content with sampling that experience in order to understand it better, but they want to go further and appropriate that experience for their own. Examples of this include “helicopter parents, stage mothers, ...fans...[, and] stalkers” who engage in “obsessive observation” of their subjects “while supplanting [the other’s] objectives, goals, or desires with [the empathizer]’s own.” The goal of this sort of empathy is to enrich the self with something that is envied about the other. For example, obsessive celebrity fans empathize with the objects of their obsessions in the hopes of feeling what it’s like to be special, worthy of admiration, and worthy in particular of the obsessiveness of the fan. But this technique is self-frustrating. By obsessively concentrating on the other, the self just gets emptier, and the quest to assimilate the other person’s life becomes increasingly desperate.

The other he calls “empathetic sadism.”^[73] Some people get a charge out of empathically feeling the suffering of someone else. And many of us can be tempted by, for example, wanting to see someone who has hurt us suffer as we suffered. Breithaupt wants us to recognize that in such cases empathy is what allows us to imagine what will most hurt our victim, and also allows us to revel in that pain.

In empathetic sadism “[a] person creates, encourages, wishes for, or tolerates a scenario in which someone else is placed in danger or made to suffer, precisely in order to feel empathy with that person, now cast in the role of the victim.” And this is not limited to brutal psychopaths or revenge fantasies. Why do we enjoy daredevils, like for example tightrope walkers or Evel Knievel-style stunt performers? We enjoy safely empathizing with their peril, maybe with their courage or fear—which they have taken on for our entertainment. Is this not a sadistic form of empathy?

Breithaupt himself puts forward “empathy rape” as an example. Contrary to theories that the rapist treats their victim as an object, devoid of feelings, Breithaupt says that at least to some rapists, the opportunity to empathize with the suffering of the victim is part of the motivation for the crime. He points out that “non-consensual” porn fantasy stories attend in detail to the feelings of the victim (a typical trope is for the victim to start off horrified and in pain, and end up begrudgingly delighted and humiliatingly grateful). Though I suspect that non-con porn fantasy stories probably represent genuine rape at least as unrealistically as other porn fantasies represent their real-world analogues.

There are also milder forms of “manipulative empathy”—“behaviors on the part of the empathizer intended to guide the other into a particular situation in which they will be emotionally predictable and it will be possible for the empathizer to coexperience their emotions.” Consider for example jumping out from behind a bush to scare someone, or telling a puzzling riddle. Internet trolls who try to push people’s buttons on-line are another example.

Sometimes people cooperate to converge on a shared and predictable emotional state: for example when you sing in a choir, attend a concert, watch a thriller, or attend a football game. There is something valued about the shared emotional experience (it wouldn't be nearly so fun to go to a rock concert or a football game if you were the only one in the audience, even if that meant you had a front row seat and never had to wait in a long line to get a beer).

What's common to all of these is "*empathy for empathy's sake*. The empathetic response becomes its own goal, independent of any consideration for—or, indeed, detrimental to—the well-being of the other."

A safer outlet for the sadistic empathy impulse is (theatrical) tragedy. Why do we enjoy a tragedy like *Hamlet*, in which everyone we care about over the course of the play dies and everything goes to shit? Our empathy for the characters evidently does not mean we are disappointed at their unhappy endings. This is a clue that it's "*empathy for empathy's sake*" we're after. "[F]iction allows us to enjoy empathy without compassion or obligation to help."

The Nietzschean critique

Fritz Breithaupt also analyzed empathy from a Nietzschean perspective.^[74] This is my summary of his summary of Nietzsche's ideas, so salt to taste:

In Nietzsche's view, people tend to have poorly-developed, weak senses of self, to their detriment. When you empathize with someone, you further suppress your self in order to simulate their point of view (other ways in which you might respond that do not weaken the self would include loving or hating that person, praising or condemning that person, helping or hindering that person, or judging that person from your own point of view).

Empathy simplifies the other person to make them more comprehensible, and this process of simplification exaggerates how unified and coherent the other person is: in other words, it makes their self look stronger than it really is. The empathetic person "project[s] onto others the self that they are lacking" and "feels empathy for that which they must give up in order to be able to feel empathy: a strong self." This makes empathy somewhat vampiric in the sense described above, and also makes it a source of envy: we can dislike those we empathize with because we envy the strong self we believe they have.

Breithaupt calls this the "empathetic endowment effect" and suggests that it explains, among other things, why people like to follow the goings-on of celebrity stars and charismatic politicians. I wonder if this also explains some of the "[why we love sociopaths](#)" phenomenon. Sociopaths, because they typically don't display much empathy, don't dissolve their selves in other people the way most of us do. This may suggest to us that they have the strong selves that we lack and envy and fantasize about having.

Nietzsche sees people's attempts to be "objective" as a variety of this same disorder. To see something objectively is sort of like empathizing with the point of view of God, and it is subject to the same sort of distortions as empathizing with another person (for example, we can be tempted to project a strong self onto this God that we hypothesize). If you try to be "objective" you undermine your self: You see your own priorities, passions, and judgments as just one of many, not as of any particular importance. You distrust and denigrate what is merely subjectively yours. You don't form firm opinions because after all there are diverse opinions to take into account, and objectively no values by which to distinguish good and bad. (It sounds to me like Nietzsche would find the attempt to counter biased empathetic engagement with rational "effective altruism" just makes the problem worse.)

How can you improve at empathy?

Since empathy seems to be a composite skill, we should be on the lookout both for things that may improve empathy writ large and for things that may improve the functioning or interoperation of its component parts. A good first step for improving your empathy might be to examine more closely which components of your empathy need the most improvement.

There is plenty of advice on how to become better at empathy and its components (and, less usefully, on how to become simply *more* empathic). And there is evidence that empathy is a skill that can improve with deliberate training.^[75] I'll try to summarize some of this advice here.

Asking questions, and listening well

As touched on in the “empathic accuracy” section above, one way to improve your empathy is to use the gift of language. You don’t have to painstakingly intuit what another person is thinking and feeling if they’re happy to just tell you when prompted.

This is a two-part skill on the empath’s part: asking the right questions, and attending well to the answers. It also only works well with a cooperative partner who is not being reticent or deceptive (or poorly self-aware or emotionally inarticulate).

[“Active Listening”](#) and [“Nonviolent Communication”](#) techniques are often recommended for this purpose. Some of this stuff can seem off-putting if it’s not used with a gentle touch: “they’re very obvious *techniques*, and I could see through them even as a toddler,” says Karla McLaren “You don’t have to parrot me to empathize with me. You need to *interact*—honestly, authentically, and as yourself.”^[76]

Sometimes people engage in small talk with the goal of avoiding conversation that will be emotionally taxing. When we want to empathize, we need to shift from that mode to asking questions that go more to the heart of things and that invite more revealing answers.

“Empathy isn’t just remembering to say *that must really be hard*,” says Leslie Jamison. “It’s figuring out how to bring difficulty into the light so it can be seen at all. Empathy isn’t just listening, it’s asking questions whose answers need to be listened to.”^[28]

Another thing that can help is a broad emotional vocabulary that, for instance, doesn’t have to fall back on something vague like “excited” when you really mean something more precise like frantic, aroused, expectant, agitated, enthused, or manic. Karla McLaren insists that “empathy is first and foremost an emotional skill.”^[77] She, and other writers on the subject, recommend improving your emotional intelligence skills as a way of also improving your empathy.

If you want to practice asking good questions and listening well, you can borrow a human from the [Human Library](#) and do just that. Or you might go out for a [Conversation Meal](#). If that seems like too much fuss or involves too much vulnerability, there’s always [Chatroulette](#), which you can enjoy over your computer from the comfort of home.

Karla McLaren also recommends the back-channel communication technique of *gossip*. Gossip can give you important context about the people around you that you might not be able to learn directly from them and that can help you improve your empathetic accuracy about them. Gossip has a bad reputation, for instance because it can surface things like envy and jealousy and because people dislike feeling gossiped about, and malicious gossip can be pretty awful, but good gossip can be informative and helpful in the empathetic project.^[78]

Literature, film, and other such media

“The greatest benefits we owe to the artist, whether painter, poet, or novelist, is the extension of our sympathies. Appeals founded on generalizations and statistics require a

sympathy ready-made, a moral sentiment already in activity; but a picture of human life such as a great artist can give, surprises even the trivial and the selfish into that attention to what is apart from themselves, which may be called the raw material of moral sentiment.... Art is the nearest thing to life; it is a mode of amplifying experience and extending our contact with our fellow-men beyond the bounds of our personal lot."

—George Eliot^[79]

"It's not a big leap to suppose that the habit of reading other people's words could put one in the habit of entering other people's minds, including their pleasures and pains."

—Steven Pinker^[80]

When you engage with good literature (or film, or what-have-you), you will often empathize with its characters. Consider a movie in which the protagonist is at a crossroads; there are pros and cons to both decisions, and the stakes are serious. The protagonist pauses, anxious and indecisive. If it's a mundane but okay movie you'll be *curious* about what happens next. If it's a great movie, you'll be *anxious and indecisive* along with the protagonist.

Good literature can also let you inside of other people's heads in a way that is normally not possible in real life. Consider stream-of-consciousness writing like Virginia Woolf's *Mrs. Dalloway* or *To the Lighthouse*. Woolf shares her extraordinary talent for empathy with us—takes us by the hand and lets us empathize vividly too, by illuminating the insides of her characters for us.

Literature can help you learn what it feels like to empathize with someone very different from yourself or in a very different situation than yourself, when you otherwise might not have known where to start. The author can also spell out for you certain social nuances that you might otherwise have difficulty picking up on your own, and thus prime you to be aware of them in similar real-life situations in the future. Some literature operates as a kind of mystery story in which you have to be attentive to clues that help you piece together the motivations and attitudes of various characters; such fiction may prepare you for real-world challenges of empathetic accuracy. In short, maybe literature is good practice for empathy.

A meta-analysis designed to determine whether existing research supports the claim that fiction-reading causally improves social cognition determined that "fiction reading leads to a small, statistically significant improvement in social cognitive performance."^[81]

Lynn Hunt argued in *Inventing Human Rights: A History* that the concept of human rights developed how it did and when it did in part as a result of the influence of mid-eighteenth-century European novelists, particularly those whose use of the briefly-in-vogue "epistolatory novel" form gave readers a more vivid sense that they were gaining access to the candid details of a real life. These novels became something of a craze, and the culture was swept up in this new, shared experience of empathizing with a (fictional) person in unaccustomed intimacy.^[82] "The epistolatory novel did not just reflect important cultural and social changes of the time. Novel reading actually helped create new kinds of feelings including a recognition of shared psychological experiences, and these feelings then translated into new cultural and social movements including human rights."^[83]

This sort of thing may be true of non-fiction as well. There's a genre of non-fiction (sometimes called "role reporting") that describes the empathetic process by having the authors immerse themselves in an unfamiliar lifestyle in order to see it from within with fresh eyes. I'm thinking of things like John Howard Griffin's *Black Like Me*, George Orwell's *Down and Out in Paris and London*, Barbara Ehrenreich's *Nickel and Dimed*, and Günter Wallraff's *Lowest of the Low*. This is a sort of empathy-extreme. It is designed explicitly to help the reader empathize with the people the author has learned to empathize with, but may also model the process of empathizing more generally. (An interesting fictional example of where extreme empathy is central to the plot, rather than incidental to the storytelling, is the film *Being John Malkovich*.) Biography and autobiography also seem designed to satisfy that

empathic urge in us to know what was it like to [grow up as a reincarnated lama, walk on the moon, cross the continent in a covered wagon, etc.].

One problem with this sort of empathy is that it may leave you at the mercy of cultural creators, who may not be very careful in how they form and direct your empathy. Your child's developing empathy may be piqued by reading *The Diary of Anne Frank* or by [*The Turner Diaries*](#), depending on which one gets into their hands first. Only certain sorts of stories and certain sorts of protagonists can get Hollywood budgets.

Sometimes works of literature use shortcuts to help you quickly decide whom to empathize with and whom you can disregard as an NPC or obstacle (the villain has a scar, speaks with an accent, and so forth), and these may influence your real-world empathy in ways you wouldn't want. Actors exaggerate their body language & other emotional expressions, in order to make them easier for us to read. There may be some danger that this teaches you only an artificial stage-dialect of emotions that translates awkwardly to the real world (imagine trying to understand people if all you knew about them was what you learned from *Noh* drama). However I suspect people who have grown up in the post-television era have learned how to express their emotions in part by what actors have modeled for them, so there's some convergence between the artificial and the real; maybe this just helped people adopt a common dialect.

Fiction writers sometimes try to get their readers/viewers to empathize with a morally repulsive character. (And here I'll plug [*Why We Love Sociopaths*](#) again.) That's all fun and games if fiction is just recreation, but if it does indeed help to shape our empathy, do we really want our empathy shaped in such a way that we see-things-from-their-side when they are a brutal serial killer or what have you?

Empathy may be a limited resource. If you do too much of it, you have to recharge before you can do it again. For this reason Karla McLaren is concerned that people turn to potentially empathy-draining activities like fiction, movies, television, video games, or outrage porn to relax. She recommends that you cultivate some non-emotionally-receptive, solitary recreation options—and that you create physical spaces where you can retreat and escape from the intrusion of others' emotions—if you find your empathy waning. [\[84\]](#)

Because fiction allows us to empathize with a character in a way that keeps us actually aloof from them, it could also conceivably atrophy behavioral empathy. When you empathize with a fictional character, that never requires anything further from you. (In this way, it is a form of the narcissistic "empathy for empathy's sake" that Fritz Breithaupt identifies as "dark empathy".)

Walk a mile in their moccasins

The aforementioned "role reporting" authors went to great lengths to experience the lives they were attempting to empathize with. Sometimes people will try to boost their empathy by doing more limited, short-term performances of that sort. For example, they might spend one night sleeping rough on the streets in an attempt to gain more empathy for homeless people, or live on the spending allowed by a minimum-wage income for a week, or something of that sort. The [*Dialogue in the Dark*](#) museum exhibit takes place in total darkness, with a blind museum guide. Such things have a stony feel to them and can be easily mocked, but there's something to the idea that you can understand more about another person's life if you directly experience some representative elements of that life.

However, this may not reliably produce the effects you might expect (empathy that leads to compassion). Counterintuitively, one study "found that people who have endured challenges in the past... were less likely to show compassion for someone facing the same struggle, compared with people with no experience in that particular situation" and that "people who have endured a difficult experience are particularly likely to penalize those who struggle to

cope with a similar ordeal.”^[85] Why? One explanation is that the [rosy retrospection bias](#) makes people forget how hard past difficulties were.

If you have more experience of life, you may “know how that sort of thing always turns out” and find it difficult to empathize with someone going through it for the first time.^[3] Your adolescent child may indeed feel that their latest crush or craze is the most imperative thing in the world, but try as you empathically may, and even having experienced your own momentous adolescent enthusiasms in your time, you may just be unable to take it seriously enough to go there with them.

Acting (for instance, improv) and role-playing games might also help you to practice empathy by slipping into and embodying characters who live lives, and have outlooks different from your own.

Imagine ways in which the other is not so other

One possible way to overcome the ingroup/outgroup or othering biases that make it difficult to empathize with someone different or distant from yourself is to try to imagine them as though they were not so different or distant. Children are sometimes helped to develop empathy through what is called “multiple empathizing”: they are encouraged to imagine that some stranger in distress were really (say) their own mother, and how they would then feel.

In [Matthew 25:31-46](#), Christians are told that Jesus will judge them in the last days in this way:

Then the King will say to those at His right hand, “Come, you blessed of My Father, inherit the kingdom prepared for you since the foundation of the world. For I was hungry and you gave Me food, I was thirsty and you gave Me drink, I was a stranger and you took Me in. I was naked and you clothed Me, I was sick and you visited Me, I was in prison and you came to Me.”

Then the righteous will answer Him, “Lord, when did we see You hungry and feed You, or thirsty and give You drink? When did we see You a stranger and take You in, or naked and clothe You? And when did we see You sick or in prison and come to You?”

The King will answer, “Truly I say to you, as you have done it for one of the least of these brothers of Mine, you have done it for Me.”

So Christians will sometimes strengthen their empathy by “Jesusing” the other. Mother Theresa again: “I see Jesus in every human being. I say to myself, this is hungry Jesus, I must feed him. This is sick Jesus. This one has leprosy or gangrene; I must wash him and tend to him.”

If you find yourself complaining “I just can’t seem to empathize with such-and-such people,” try instead to isolate a single such-and-such person to empathize with. It’s much easier to empathize with a particular person than with a group of people.

Then, consider what you share with that person, even at the most basic level. This is one of the most common and ancient empathy-related interventions. Consider Shylock’s plea in Shakespeare’s *Merchant of Venice*:

“Hath not a Jew eyes? Hath not a Jew hands, organs, dimensions, senses, affections, passions? Fed with the same food, hurt with the same weapons, subject to the same diseases, healed by the same means, warmed and cooled by the same winter and summer, as a Christian is? If you prick us, do we not bleed? If you tickle us, do we not laugh? If you poison us, do we not die?”

"The shift to seeing others not as different but as similar seems to be the strongest way to influence empathic resonance or insight."^[86] There's also a potential feedback effect, in that the more you empathize with someone the more you recognize (or suppose that you recognize) similarities between them and you.

If you somehow come to feel empathy toward a member of an outgroup, this can also increase your propensity to feel empathy toward other members of the outgroup.^[87] And if you learn examples of other members of your ingroup helping members of the outgroup, this can also increase your empathy for members of the outgroup.^[88]

Try to see things from both/all sides

As mentioned before, when you empathize with someone, you tend to take their side, and this can make it difficult to evaluate conflicts dispassionately. One possible remedy is to make an effort to see things from other sides as well. [Ideological Turing Tests](#) are one way of testing your empathy in this regard.

But it's not clear that this sort of intervention is reliable. In one experiment, grade school students in Northern Ireland were taught to understand the historical conflict there from the perspectives of Catholics and of Protestants, in such a way that the students were able to skillfully explain both perspectives. But this did not seem to reduce polarization in the students. Instead it seemed to help the students in "assimilating the experience of the other into their own frame narrative."^[89] It's even possible that asking the students to employ empathy to process additional examples of the conflict (though seen from various points of view) exacerbated their original empathy-induced side-taking bias.

Be cautious around demands for empathy

When someone pleads with you to show empathy, they can make it sound like they are asking nothing more from you than what you owe them: to see things from their point of view, to show some consideration for how they feel.

But empathy can be taxing. It is emotionally costly. It is not just anyone's for the taking. You have the right to ration it out as your wisdom and priorities dictate.

As mentioned above, the "see things from my point of view" gambit can be a way to encourage side-taking. There are times when you do not want to begin by taking sides but instead you want to evaluate a conflict according to impartial standards of [justice](#). In such cases it may be best to say, "no; I'm going to keep looking from my own point of view."

Some people who are having a bad time think they will feel better if they can make other people around them feel miserable too. Thus the adage "misery loves company." Pleas for empathy can be part of that game. Sometimes people will crave the attention and intimacy they receive from others who empathize with them in their distress, and this incentivizes them to undergo ([or masquerade, or amplify](#)) additional distress. If you are [cautious](#) around demands for empathy, you may avoid feeding dumpster fires like these.

Games

Multi-person games sometimes incentivize trying to "get into the head" of your opponent or teammate. Even rock-paper-scissors has something of that. [Bidding strategy](#) in Bridge is a sort of formalized gameplay-empathy in which you use gameplay methods to intuit what your partner knows. There's a frequently told tennis legend in which Andre Agassi was able

to take advantage of a “tell” in Boris Becker’s body language to determine how Becker was going to serve.^[90]

There are also a variety of games that are meant to be icebreakers or intimacy-builders. Sometimes these can provoke empathic responses, and some are even designed with this in mind. [Questions & Empathy](#), for example, is a card game that is designed to get people to share less superficial parts of themselves by answering probing questions from unfamiliar perspectives. “It’s like a highbrow Cards Against Humanity. It escalates you from small talk to big talk ultra fast.”^[91]

Prosthetics, virtual reality, and video games

Prosthetics can allow you to viscerally experience some aspects of lives that are different from your own. There are, for example, [fake bellies that people can wear to empathize with pregnant people](#). A device called [Sympulse](#) transmits tremors from a Parkinson’s patient to another person, “to help foster clinical empathy... to give movement disorder physicians and caregivers a sense of what their patient or loved one is experiencing in real time.”

Invisible disabilities are particularly hard to empathize with. [Detour: Brain Deconstruction Area Ahead](#) is a (1994) film by someone who tried to show from their perspective how their traumatic brain injury causes distortions in how they perceive the world and the difficulties this causes.

Kaitlin Ugolik Phillips’s *The Future of Feeling: Building Empathy in a Tech-Obsessed World* describes a number of virtual reality simulations designed to prompt empathy. (Some of these are virtual reality in the sense of fully-immersive headset-style presentations, others are more like interactive 360° video essays. Some are dramatizations, while others are more journalistic attempts to place you at a real scene.)

- [1000 Cut Journey](#) — “participants embody a Black male, Michael Sterling, experiencing racism as a child through disciplinary action in the classroom, as an adolescent encountering the police, and as a young adult experiencing workplace discrimination”
- [6x9](#) — “places you inside a US solitary confinement prison cell and tells the story of the psychological damage that can ensue from isolation”
- [Across the Line](#) — “put[s] viewers in the shoes of a patient entering a health center for a safe and legal abortion”
- [After Solitary](#) — “lets viewers walk around the cell with Kenny as he recounts his experiences in solitary”
- [The Alfred Lab](#) — “embody Alfred, a 74-year-old African American man with macular degeneration and high-frequency hearing loss as he spends time with family, visits the doctor, and receives a diagnosis”
- [Becoming Homeless: A Human Experience](#) — “spend days in the life of someone who can no longer afford a home”
- [Carne y Arena](#) — “walk in a vast space and thoroughly live a fragment of the refugees’s personal journeys”
- [Clouds Over Sidra](#) — “Meet Sidra. This charming 12-year-old girl will guide you through her temporary home: The Zaatari Refugee Camp in Jordan.”
- [The Displaced](#) — “portrays the impact of war and displacement on children with heartbreakingly, immersive realism”
- [Enter the Room](#) — “experience the trauma of war through augmented reality”
- [Hunger in Los Angeles](#) — “simulate the experience of watching a man go into diabetic shock at a Los Angeles food bank”
- [Project Syria](#) — “witness a bomb go off in the streets of Syria, turning a normal Syrian afternoon to complete chaos and destruction”
- [Use of Force](#) — “the homicide of Anastacio Hernandez Rojas who was beaten and tasered by more than a dozen border patrol agents”

Another source of interesting experiments in empathy-evoking technology is empathymuseum.com.

I find some of this uncomfortably aggressive. I imagine someone saying “you ought to empathize with so-and-so, and we’re going to keep giving you the helmet until you do.” Advocacy documentaries can be manipulative enough without adding VR to the mix.

I also wonder, why on earth would anyone volunteer to “experience the trauma of war”? What kind of monster would design an app whose purpose is to inflict the experience of the trauma of war on people? (I suppose since there are whole industries dedicated to inflicting the real trauma of war on people, I shouldn’t be so shocked.) Either the viewers know that they’re really only going to get a voyeur’s peek at it (which kind of defeats the purpose) or they imagine they’re going to get the real thing but they naively think of the real thing as though it were something safely scary like a roller-coaster (and is VR likely to change their mind about that?).

Something like *Across the Line* does not seem to me as though it is meant to help people learn to empathize with a patient getting an abortion. It strikes me more as a sort of [waving the bloody shirt](#) tactic to further raise the ire of pro-choice activists toward their opponents. As one critic said of *Carne y Arena*, it “puts so many eggs in the basket of creating empathy, since its power so clearly depends on a foundation of preexisting sympathy.”^[92]

Video games are a sort of virtual-reality-lite. In “[first person](#)-style” games in particular it’s difficult *not* to feel-with the struggles and perils of the character whose actions you control and whose eyes you see through. Does this exercise the same skills as empathy? (Does it do so to our benefit?)

Developers have also designed video games that are meant to teach or improve empathy. Several examples I’ve seen seem to be designed for children and adolescents.^[93] Another, *Life is Strange*, gave the game’s protagonist “an ability to read, experience, and manipulate the emotions of her peers” as part of the gameplay dynamics.^[94]

Dealing with hyperempathy

Karla McLaren says she was hyperempathic as a child and that this caused some difficulties for her (for instance she would sometimes shut down or behave strangely in social situations because of this exceptional sensitivity). While she has now turned this to her advantage by becoming particularly skilled in and understanding of empathy, it was a painstaking process to get to that point. Her book *The Art of Empathy: A Complete Guide to Life’s Most Essential Skill* might prove inspiring to you if you have similar struggles.

I am far from hyperempathic, but I think I can empathize somewhat with this. I can remember times during my days of psychedelics enthusiasm when, under the influence of some drug or other, I became hyperaware of social nuances, subtext, and the multiple levels on which social interactions were simultaneously operating. All of the wheels-within-wheels of conversation and body language and insinuation and so forth left me reeling, unable to make a move in this 10-D chess game because I couldn’t possibly trace all of the ramifications of what I was communicating. While there may have been an element of paranoia / delusions of reference here, I’m inclined to think that this complex, multi-layered social interaction is the norm, but that we usually muddle through reasonably well letting most of it happen sub- or unconsciously. Conscious awareness of the complexity and breadth of it all can be enlightening, but also paralyzing.

It may be useful if you have hyperempathic tendencies to learn how to identify unwelcome empathy and put a damper on it.^[95] Helen Riess of the [Empathy and Relational Science Program of Boston’s Massachusetts General Hospital](#) recommends exercises in deep

breathing, detachment, and mindfulness as ways help you observe others without having your own reactivity or empathically-evoked emotions overwhelm you.

Meditation

There are varieties of meditation that seem to exercise empathy. The Mahayana Buddhist “exchanging self and others” meditation in which you assume the perspective of somebody else towards you is one example. Mettā meditation sometimes includes a component in which the meditator tries to imagine the basic desires of an expanding circle of beings. That sort of meditation is apparently effective at improving empathetic accuracy.^[96]

As an aside, there seems to be some convergence between the Buddhist [no-self](#) teaching and the Nietzschean critique of empathy mentioned earlier. In the Nietzschean point of view, when we empathize, we seek for a strong self in the other and project it onto them. Buddhism would agree that this projected self is an illusion, but would go further and say that we constitute our own illusory selves in much the same way. So Buddhist empathy may imply a peculiar psychology that is hard to compare to other forms.

Miscellaneous advice

Maybe you can learn and memorize certain [universal facial expressions](#) that can help you to gain empathic access to another person’s state-of-mind if you have difficulty doing so in a more immediately intuitive way.

Several authors suggested that you can improve your empathy by engaging in more cooperative pursuits with others.

Becoming more curious about the people around you (or being more bold about indulging the curiosity you already have) may help.

Being more aware of the biases that accompany empathy, and bringing those to mind when you find yourself empathizing, may help you to bring your empathic responses more in line with your values.

Conclusion

Empathy is something people do. We can do it in a way that is better or worse at contributing to our human flourishing. If we characteristically empathize well, we exhibit the virtue of empathy. This seems to be, as with other virtues, something we can learn and improve at with practice. I hope this post will help.

1. ^

Susan Lanzoni [“Empathy’s Evolution in the Human Imagination” Zocalo Public Square](#)
17 July 2017

2. ^

C. Daniel Batson *Altruism in humans* (2011)

3. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019)

4. ^

Paul Bloom *Against Empathy: The Case for Rational Compassion* (2016)

5. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013)

6. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014)

7. ^

Karla McLaren thinks the affective/cognitive description is misleading, though she also subdivides empathy in a similar way to people who use that primary division. Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013) pp. 26-42

8. ^

for example Paul Bloom (*Against Empathy: The Case for Rational Compassion*, 2016) says cognitive empathy isn't empathy but "understanding," while the empathy he's *Against* is mostly the affective kind

9. ^

Daniel Goleman "What Is Empathy?" in *Empathy* Harvard Business Review Press (2007)

10. ^

Elizabeth A. Segal, et al. *Assessing Empathy* (2017)

11. ^

Adam Smith, *The Theory of Moral Sentiments* (1759)

12. ^

e.g. P.A. Miller & N. Eisenberg "The relation of empathy to aggressive and externalizing antisocial-behavior" *Psychological Bulletin* (1988) pp. 324-344

13. ^

Elizabeth A. Segal, et al. *Assessing Empathy* (2017) calls it "emotional egocentricity bias".

14. ^

M. Toi, & C.D. Batson, "More evidence that empathy is a source of altruistic motivation" *Journal of Personality and Social Psychology* 43 (1982) pp. 281-292

15. ^

Paul Bloom *Against Empathy: The Case for Rational Compassion* (2016) chapter 2

16. ^

Paul Bloom *Against Empathy: The Case for Rational Compassion* (2016) chapter 4

17. ^

Elizabeth A. Segal, "Social empathy: A model built on empathy, contextual understanding, and social responsibility that promotes social justice" *Journal of Social Service Research* 37 (2011) pp. 266-267

18. ^

Elizabeth A. Segal, et al. ["Developing the Social Empathy Index: An Exploratory Factor Analysis"](#) *Advances in Social Work* 13 (2012)

19. ^

Elizabeth A. Segal, et al. *Assessing Empathy* (2017)

20. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019), chapter 2

21. ^

Sara H. Konrath, Edward H. O'Brien, & Courtney Hsing "Changes in Dispositional Empathy in American College Students Over Time: A Meta-Analysis" *Personality and social psychology review* 15 (2011)

22. ^

e.g. Maia Szalavitz ["Shocker: Empathy Dropped 40% in College Students Since 2000"](#) *Psychology Today* 28 May 2010

23. ^

Pierre Cérèsole *For Peace and Truth* (1954)

24. ^

Martin Hoffman *Empathy and Moral Development* (2000)

25. ^

Jeremy Rifkin *The Empathic Civilization: The race to global consciousness in a world in crisis* (2009)

26. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014) pp. 178-179

27. ^

B.K.Y. Lau, J. Geipel, Y. Wu, & B. Keysar (2022). ["The extreme illusion of understanding."](#) *Journal of Experimental Psychology: General.*

28. ^

Leslie Jamison [The Empathy Exams](#) (2014), p. 5

29. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013), p. 32

30. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013), pp. 188-189

31. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014) pp. 98+

32. ^

["The Multiplicity of Empathy: Study Reveals Both Interpersonal & Intrapersonal Benefits of Being Empathetic"](#) (press release)

33. ^

R.B. van Baaren, R.W. Holland, K. Kawakami, & A. van Knippenberg "Mimicry and prosocial behavior" *Psychological Science* 15 (2004) pp. 71-74

B.C.N. Müller, A.J. Maaskant, R.B. van Baaren, & A. Dijksterhuis "Prosocial consequences of imitation" *Psychological Reports* 110 (2012) pp. 891-898

R.B. van Baaren, R. Janssen, T.L. Chartrand, & A. Dijksterhuis "Where is the love? The social aspects of mimicry." *Philosophical Transactions of the Royal Society Biological Sciences* 364 (2009) pp. 2381-2389

34. ^

D. Grühn D, et al. ["Empathy across the adult lifespan: Longitudinal and experience-sampling findings"](#) *Emotion* 8 (2008) pp. 753-765

35. ^

Jason Kottke ["Giving Is a Form of Attention"](#) kottke.org 8 April 2022

36. ^

D. Grühn D, et al. ["Empathy across the adult lifespan: Longitudinal and experience-sampling findings"](#) *Emotion* 8 (2008) pp. 753-765

37. ^

B. Bernard *Resiliency: What we have learned* (2004)

38. ^

Bruce P. Doré et al. "Helping Others Regulate Emotion Predicts Increased Regulation of One's Own Emotions and Decreased Symptoms of Depression" *Personality & Social Psychology Bulletin* 43 (2017) pp. 729-739

39. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019), Epilogue

40. ^

Paul Bloom *Against Empathy: The Case for Rational Compassion* (2016) chapter 6

41. ^

Leslie Jamison [The Empathy Exams](#) (2014), title essay

42. ^

Kaitlin Ugolik Phillips *The Future of Feeling: Building Empathy in a Tech-Obsessed World* (2020), pp. 101-102

43. ^

Jon Kolko "A Process for Empathetic Product Design" in *Empathy* Harvard Business Review Press (2007)

44. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014) pp.xi-xii

45. ^

Adam Waytz "The Limits of Empathy" in *Empathy* Harvard Business Review Press (2007)

46. ^

["Fidelity exploits virtual reality for 'empathy training'"](#) Finextra 16 October 2017

47. ^

Rolfe Winkler, Daisuke Wakabayashi, & Elizabeth Dwoskin ["Apple Buys Artificial-Intelligence Startup Emotient"](#) Wall Street Journal 7 January 2016. Quote: "...amassing an enormous database of human emotions using technology that relies on algorithms to analyze people's faces and potentially discover their deepest feelings."

48. ^

Harma Meffert, et al. "Reduced Spontaneous but Relatively Normal Deliberate Vicarious Representations in Psychopathy" *Brain* 136 (2013) pp. 2550-2562

49. ^

Jon Sutton, Peter K. Smith, & John Swettenham "Social Cognition and Bullying: Social Inadequacy or Skilled Manipulation?" *British Journal of Developmental Psychology* 17 (1999) pp. 435-450

50. ^

Eric W. Dolan ["Psychopathic tendencies linked to reduced susceptibility to contagious yawning"](#) Psypost 29 April 2022

51. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014) p.xvii

52. ^

Roman Krznaric *Empathy: Why It Matters, and How to Get It* (2014) pp.xvii-xviii

53. ^

Simon Baron-Cohen *The Science of Evil: On Empathy and the Origins of Cruelty* (2012) p. 194

54. ^

Leslie Jamison [*The Empathy Exams*](#) (2014) pp. 4-5, 12

55. ^

Ezekiel Gleichgerrcht & Jean Decety "The Costs of Empathy among Health Professionals" in J. Decety, ed. (2012) *Empathy: From bench to bedside* p. 255

56. ^

J. Halbesleben, et al. "Too Engaged? A conservation of Resources View of the Relationships Between Work Engagement and Work Interference with Family" *Journal of Applied Psychology* 94.6 (2009)

57. ^

Cicero, [*Tusculan Disputations*](#) (~45BC), book III "On Grief of Mind"

58. ^

for example, David D. Vachon, Donald R. Lynam, & Jarrod A. Johnson "The (Non) Relation between Empathy and Aggression: Surprising Results from a Meta-Analysis" *Psychological Bulletin* 140 (2014) p. 16

59. ^

Betsy Blaney "[Baby Jessica turns 25, gains access to trust fund](#)" Associated Press 26 March 2011

60. ^

C.D. Batson, T.R. Klein, L. Highberger, & L.L. Shaw "Immorality from empathy-induced altruism: When compassion and justice conflict" *Journal of Personality and Social Psychology* 68 (1995) pp. 1042-1054

Anneke E.K. Buffone & Michael J. Poulin "Empathy, Target Distress, and Neurohormone Genes Interact to Predict Aggression for Others—Even Without Provocation" *Personality and Social Psychology Bulletin* 40 (2014) pp. 1406-1422

61. ^

Michael N. Stagnaro & Paul Bloom "The Paradoxical Effects of Empathy on the Willingness to Punish" (unpublished, 2016)

62. ^

S.G. Young & K. Hugenberg "Mere socialization categorization modulates identification of facial expressions of emotion" *Journal of Personality and Social Psychology* 99 (2010) pp. 964-977

63. ^

description of (not quote from) J.J. Van Bavel & W.A. Cunningham "Self-categorization with a novel mixed-race group moderates automatic social and racial biases" *Personality and Social Psychology Bulletin* 35 (2009) pp. 321-335 [description from Elizabeth A. Segal, et al. *Assessing Empathy* (2017)]

64. ^

M. Cikara & S.T. Fiske "Bounded empathy: Neural responses to outgroup targets' (mis)fortunes" *Journal of Cognitive Neuroscience* 23 (2011) pp. 3791-3803

G. Bucchioni & al. "Do we feel the same empathy for loved and hated peers?" *PLoS ONE* 10 (2015) pp. 1-11

65. ^

M.W. Kraus, S. Côté, & D. Keltner "Social class, contextualism, and empathic accuracy" *Psychological Science* 2 (2010) pp. 1716-1723

66. ^

J. Hogeveen, M. Inzlicht, & S.S. Obhi "Power changes how the brain responds to others" *Journal of Experimental Psychology* 143 (2014) pp. 755-762

A.D. Galinsky et al. "Power and perspectives not taken" *Psychological Science* 17 (2006) pp. 1068-1074

G.A. Van Kleef et al. "Power gets you high: The powerful are more inspired by themselves than by others" *Social Psychology and Personality Science* 6 (2015) pp. 472-480

67. ^

Elizabeth A. Segal, et al. *Assessing Empathy* (2017) ch. 4; Stéphane Côté, et al. "Social power facilitates the effect of prosocial orientation on empathic accuracy" *Journal of Personality and Social Psychology* 101 (2011) pp. 217-232

68. ^

M. Schmid Mast & A. Darioly "[Emotion recognition accuracy in hierarchical relationships](#)" *Swiss Journal of Psychology* 73 (2014) pp. 69-75

69. ^

Christopher J. Einolf "[Empathic concern and prosocial behaviors: A test of experimental results using survey data](#)" *Social Science Research* 37 (2008) pp. 1267-1279

70. ^

J. Decety, S. Echols, & J. Correll "The blame game: The effect of responsibility and social stigma on empathy for pain" *Journal of Cognitive Neuroscience* 22 (2009) pp. 985-997

71. ^

Paul Bloom *Against Empathy: The Case for Rational Compassion* (2016) chapter 5

72. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019), chapter 5

73. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019), chapter 4

74. ^

Fritz Breithaupt *The Dark Sides of Empathy* (2019)

Friedrich Nietzsche *Beyond Good and Evil* (1886) and *The Genealogy of Morals* (1887)

75. ^

E. Teding van Berkhout & J.M. Malouff ["The efficacy of empathy training: A meta-analysis of randomized controlled trials"](#) *Journal of Counseling Psychology* 63 (2016) pp. 32-41

76. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013), chapter 8

77. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013), p. 71

78. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013), pp. 236-261

79. ^

George Eliot "The Natural History of German Life" *Westminister Review LXVI* (July 1856) pp. 28-44

80. ^

Steven Pinker *The Better Angels of our Nature: Why Violence Has Declined* (2011) p. 175

81. ^

David Dodell-Feder, et al. ["Fiction reading has a small positive impact on social cognition: A meta-analysis"](#) *Journal of Experimental Psychology: General* 147 (2018), pp. 1713-1727

82. ^

Lynn Hunt *Inventing Human Rights: A History* (2007)

83. ^

Lynn Hunt ["Inventing Human Rights"](#) (lecture, March 2008)

84. ^

Karla McLaren *The Art of Empathy: A Complete Guide to Life's Most Essential Skill* (2013) chapter 6

85. ^

Rachel Ruttan, Mary-Hunter McDonnell, & Loran Nordgren "It's Harder to Empathize with People If You've Been in Their Shoes" in *Empathy* Harvard Business Review Press (2007)

Rachel L. Ruttan, et al. "Having 'been there' doesn't mean I care: when prior experience reduces compassion for emotional distress." *Journal of Personality and Social Psychology* 108 (2015) pp. 610-622

86. ^

Elizabeth A. Segal, et al. *Assessing Empathy* (2017), chapter 4

87. ^

C.D. Batson, et al. "Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group?" *Journal of Personality and Social Psychology* 72 (1997) pp. 105-118

88. ^

T. Saguy, et al. "Awareness of intergroup help can rehumanize the out-group" *Social Psychological and Personality Science* 6 (2015) pp. 551-558

89. ^

Keith C. Barton & Alan W. McCully "Trying to 'See Things Differently': Northern Ireland Students' Struggle to Understand Alternative Historical Perspectives" *Theory & Research in Social Education* 40 (2012)

90. ^

Scott Davis e.g. ["Tennis legend Andre Agassi revealed that he learned how to beat a rival by watching his tongue on serves"](#) *Yahoo! Sports* 29 April 2021

91. ^

Liz Stinson ["Can a Card Game Teach You Empathy? This Creative Agency Thinks So"](#) *Eye on Design* 26 May 2017

92. ^

Ben Davis "Can VR Really Make Us Feel Empathy? Alejandro G. Iñárritu's 'Carne y Arena' Proves That's the Wrong Question" *Artnet News* 30 March 2018

93. ^

e.g. ["A video game can change the brain, may improve empathy in middle schoolers"](#) *ScienceDaily* 9 August 2018

94. ^

Kelly Doherty ["The gaming industry is finally teaching players empathy"](#) *i-D* 7 October 2021

95. ^

As I was putting the finishing touches on this post, Joanna Cannon's "["I feel your pain": confessions of a hyper-empath"](#)" was published in the *Guardian*; it has a section with some advice on how to cope with hyperempathy.

96. ^

J.S. Mascaro, J.K. Rilling, L.T. Negi, & C.L. Raison "[Compassion meditation enhances empathic accuracy and related neural activity](#)" *Social Cognitive and Affective Neuroscience* 8 (2013), pp. 48–55

Notes on Frugality

This post examines the virtue of **frugality** (a.k.a. thrift, economy). I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is this virtue?

Frugality concerns the efficiency of your economic choices: how you earn you living, husband resources, spend money, and the like. A frugal person is conscious of the value of their resources and of their time, and makes trade-offs wisely.

Frugality is usually contrasted with its vice-of-deficiency: extravagance, prodigality, squandering, wastefulness, being a spendthrift. In this it is allied with virtues like efficiency, [moderation](#), [prudence](#), and [simplicity](#). Benjamin Franklin paired frugality with [industry](#) as the key virtues for financial security: “[W]aste neither time nor money, but make the best use of both. Without industry and frugality nothing will do, and with them everything.”^[1]

Sometimes that same contrast is made not to the vice itself, but to the likely consequences of that vice: indebtedness, dependence on others, poverty. If you are not frugal, you may suffer for it, and this may also include having to lower your standards of [dignity](#). Franklin also had a nice *bon mot* about that: “It is hard for an empty bag to stand upright.”^[2]

Frugality sometimes is also contrasted with its vice-of-excess, which goes by names like miserliness, niggardliness, or stinginess. There is some tension between frugality and the virtues of generosity and magnificence.

Occasionally in older works you will see a virtue called “frugality” that is described in a way that makes it sound more like [temperance](#). (For example, in one translation of Cicero’s *Tusculan Disputations*, he describes frugality as combining fortitude, justice, and prudence in one package, says the cowardly man cannot be called frugal, and says frugality’s “peculiar property seems to be to govern and appease all tendencies to too eager desire after anything, to restrain lust, and to preserve a decent steadiness in everything”^[3] — all of which doesn’t really resemble how we use the term “frugal” today.)

About frugality

“Men don’t understand how great a revenue sparingness is.” —Cicero^[4]

Frugality is a facet of life optimization. A frugal person lives efficiently, in a streamlined way that enables them to accomplish more with less effort.

Time is money

Most every adult gets through life by exchanging resources they control for things they need and want. Most of us have to work for a living in order to get those

resources in the first place. In wealthier countries, the number of hours per year the typical adult worker works for a living has dropped by almost half in the last hundred and fifty years^[5] but that worker still spends about 1,500 hours per year “at work”—about a quarter of their waking hours.

That does not count time spent commuting to and from work, or other time spent on things that support one’s employability (taking licensing exams, hunting for a job, networking). It also doesn’t count the portion of time a person spends in school that is primarily for the purposes of vocational training. On the other hand, there is also a period of childhood and (if we’re lucky) retirement in which many people have their basic needs more-or-less taken care of without having to contemporaneously work to obtain the resources that make that possible.

But without splitting hairs, it’s clear that the portion of the typical person’s life that is devoted in large part to generating economic resources is considerable.^[6]

Some people are fortunate enough to make their living in a way that is inherently rewarding as well as financially compensated. But for most of us there is some degree of trade-off—to a greater or lesser extent we spend our working hours doing something we would not be doing were we not being paid for it.

The money you spend is, in this way, a sort of crystallized version of the time, effort, and attention you took away from your own immediate interests and put instead into obtaining the money. When you spend that money, it is as though you were redevoting that time, effort, and attention back to your own interests. But hardly anyone, when they spend money, asks themselves if they really want to devote X amount of time, effort, and attention on whatever it is they’re buying.

The fact that the time, effort, and attention is mediated through the more abstract and fungible medium of money, and the fact that there is a delay from the moment this time, effort, and attention is exerted to the moment its equivalent in money is spent, make it difficult for our minds to process in this way. Pulling a dollar bill from your wallet feels much the same as pulling a \$10 bill or a \$100 bill. Seeing a digit in your bank account balance change feels about the same whether that digit is in the ones’ place or the hundreds’ place. But the more visceral difference between having to stay ten minutes late one night at your stinking job and having to work through the whole damn weekend is a difference of the same magnitude.

There is also the fact that once you have earned the money, the time, effort, and attention you put into earning it is a sunk cost. So seeing that money as the exact equivalent of the sunk time, effort, and attention isn’t quite right. You have to look at it with fresh eyes: What is that money worth to me now. And that isn’t as simple as it might seem.

What is money worth?

Money seems at first to have an absolute value. It is denominated in nice, simple, natural numbers. Five bucks will buy you a venti mocha the same as it will for Elon Musk or anyone else. But because different people have to put in different amounts of time, effort, and attention to earn five dollars, what looks superficially like the same exchange of money for goods/services may really be a very different exchange of time/effort/attention for goods/services. For this reason, it can be a mistake to look for

universal standards of frugal spending — what is frugal for you might not be for someone else, and vice-versa.

In *David Copperfield*, the character Wilkins Micawber put it this way:

Annual income twenty pounds, annual expenditure nineteen, nineteen and six, result happiness. Annual income twenty pounds, annual expenditure twenty pounds ought and six, result misery.^[7]

An extra sixpence means much more to one than the other, and so they will reasonably value the same amount differently. But it's even more confusing than that: Not only is five bucks legitimately worth more or less to different people, but it is worth more or less to the same person — for instance, depending on whether those dollars are wholly insufficient, marginally sufficient, or surplus relative to their needs. For another example, if you'll be homeless tomorrow if you can't scrape up another \$400 for the rent, you probably won't be very tempted by an opportunity to work six hours today for \$50 per hour — you'd be better off spending the time packing your bags and looking for somewhere to shelter. But an opportunity to work the next 12 hours for \$35 per hour might be the difference between having a roof over your head next month or not. By being willing to work more for less, it's not that you are being irrational in how you are valuing your time, but that the value of money is a discontinuous function of its quantity in your case.

Money is time

The translation between money and time/effort/attention also runs in the opposite direction. You can [buy free time](#) with money (for example, by paying someone else to do some time-consuming chore for you). If it costs you less in time/effort/attention to earn that money than the time/effort/attention you can purchase with it, such a purchase may be a frugal bargain, and being unwilling to make such a purchase from motives of short-sighted frugality might be a case of “penny wise, pound foolish.”

Some money-saving do-it-yourself tips can be ironically unfrugal by this same measure: costing you so much extra time and effort that it would have been more frugal to whip out your wallet from the get go. (Though being able to [do-it-yourself](#) can help to build skills and understanding that go beyond the immediate need that is satisfied, so take that into account.)

The tempting reward of retail therapy

Something that further baffles people is that we seem to be able to give ourselves a reinforcing psychological reward simply by making a purchase^[8] — one that seems harder to obtain by earning or saving money, and one that can temporarily inflate the apparent value of what we purchase. The folk practice of “[retail therapy](#)” takes advantage of this reward mechanism to provide a short-term mood boost. Some people abuse this mechanism to the point where it resembles other compulsive addictions.* If your pursuit of frugality is frustrated by cravings like these, you may need to work on the virtues of [temperance](#) and/or [self control](#).

Vendors and advertisers eagerly take advantage of this, and of various other cognitive biases that can make us undervalue our time and money and overestimate the value of what we exchange them for. Although each vendor and advertiser has the specific

goal of trying to convince us to purchase their particular product or service, their collective efforts can in this way have the effect of eroding our [rational](#) frugality in general. If money is burning a hole in your pocket, consider that the fire may have been deliberately set.

What frugality communicates

One's earning and spending decisions have signalling effects as well. People often communicate what they value or are interested in — are *invested* in — by what they spend money on. Someone's more expensive tastes are likely to be among the unusual things people mention about them (or that they themselves casually let drop or post photos of). Advertisers often imply that if you purchase their products, you'll also get a boost of esteem from those around you as a bonus. Conspicuous consumption is a good way of broadcasting which economic strata you aspire to. In ways like these, as you spend money you also contribute to the construction of your persona. (And if you buy an expensive persona, you may find it also costs a lot to feed.)

Frugality is an attitude towards earning and spending that also sends a message and contributes to your persona. One possible obstacle to frugality is the sense in some circles that it's vulgar or unfashionable to be concerned about it. Being casual with money, ignorant about budgeting basics, and so forth can give one an air of devil-may-care financial security and confidence. Being calculating and frugal can give the impression that you are obsessive about money or fearful about the future.

This can work the other way, too, however. Someone who is sensibly frugal can thereby become *genuinely* confident about their financial situation in a way that is at least as attractive as being ignorantly blasé, and can also seem respectably down-to-earth and sensible.

In 1978, the teenaged autodidact Dolly Freed wrote about her family's deliberately frugal lifestyle. She noted:

[M]oney doesn't buy only goods and services, it also buys prestige and status. Being somewhat egocentric, we don't feel the need to buy prestige or status. The neat trick that Diogenes pulled was to turn the tables on those of his contemporaries who believed that "Life is a game and money is how you keep score." He didn't keep score. We don't keep score. You needn't keep score either if you don't want to. It's entirely up to you. [\[9\]](#)

Money and the hedonic treadmill

Several years back, an American wealth management firm decided to do some market research by surveying about 800 wealthy American adults (people with at least \$500,000 in investable assets).

When asked how much they needed to feel financially secure in the future, respondents consistently cited a need to approximately double their current level of assets. Those with \$10 million or more felt they needed a median of \$18.1 million; those with \$5 million or more needed \$10.4 million, and those with a half million to \$1 million said they needed \$2.4 million. [\[10\]](#)

Furthermore, a third of those surveyed said “that having enough money is a constant worry in their life.”^[10]

There seems to be some sort of bias that causes people to be more apt to look at their current financial situation from above rather than from below. Instead of saying “how financially fortunate I am now compared to how I would be if I had half as much as I do,” they say “how financially fortunate I would finally be if I only had twice as much as I do.”

This is sometimes blamed on a lifestyle-ratchet: If your financial situation improves, you may respond by increasing your expenses — moving into a nicer home, taking on a more expensive hobby, developing a taste for fine wine — in such a way that your baseline expectations change. What started as a delightful improvement in your economic fortunes becomes instead the Lifestyle-You-Are-Accustomed-To, one which it would now be a painful deprivation to do without. It seems many people guess they would prefer to live an insecurely, marginally wealthy life than a securely, comfortably ordinary one, and have to learn by experience that it’s not as fun as they expected.

If you can avoid this ratcheting temptation, through discipline and conscious deliberation, and maybe eventually by means of establishing a habitually frugal character, you can finally hope to reach a point of financial satisfaction rather than constantly seeing that point recede into the distance twice as quickly as you approach it. And if you can deratchet yourself from any unrewarding excessive expense in your current lifestyle, you may find that you’ve already reached that point.

If you notice that you are in the habit of comparing your financial situation enviously to people with a lot more than you have (or whose unfrugal habits make it appear that they have a lot more than you have), consider that maybe this is a source of bias. To correct for this you might also try comparing yourself to the average person in your surroundings, in your culture, on earth today, and over the course of human history, so as to gain a fuller perspective.

My experience with frugality

In 2003 I had just started to hit the six-figure salary range in my San Francisco area tech job, and I felt pretty high on the hog. But in reaction to the Iraq War (or at least with that as the final straw) I became a conscientious objector to federal taxes,^[11] and I decided to reduce my income to the point at which I no longer owe federal income tax. As a result, I had to adjust my lifestyle so that I could live on a much lower income. So I got some quick practical lessons in frugality.

Money magazine profiled me briefly some years ago for an article on how people avoid paying taxes.^[12] They concluded that their readers probably wouldn’t enjoy what they called the “ascetic lifestyle” that comes along with my technique. But if this is “asceticism,” asceticism is very underrated. The life I’m leading now is fuller and more enjoyable than ever, I have less anxiety (and less guilt about my taxes) and feel more integrity, and I’m genuinely living a life of abundance.

One way I measure this abundance is in the amount of free time I have. It typically takes less time to earn less money, so you can give yourself a time raise by giving yourself a pay cut. (And the fact that I no longer have income tax withheld from my wages means that although my total income went down, my hourly take-home went up.) In a typical year nowadays I devote fewer than 500 hours to making a living,

compared with maybe 2,500 hours of commute+office time back when I was making the big bucks. Having seen both sides, I now feel that much of the time I was surrendering in order to pursue a career and more money is more valuable to me than the extra money I gained in the trade.

By living frugally, even while earning much less than before I squirrel away about 40% of what I earn for retirement or for health emergencies, so I don't feel I'm being neglectful in that way, either.

When I first adopted my more frugal lifestyle, I thought I was going to be making a noble sacrifice for my principles. But it wasn't long before I realized that my life had improved through my frugal choices, and now I think of a well-paying full-time job (even with all the perks) as something that would be a sacrifice in comparison. The frugal life I took on out of conscientious necessity, I now would continue in simply because I prefer it. (Still, of all the things I don't miss spending money on since I started living more frugally, I don't miss the government the most.)

I don't mean to suggest that the specific choices I made would be the right ones for everyone. I've been very lucky in many ways. I went to college when it was still more-or-less affordable, and by the time I got my degree I had no more debt than fit on a Discover card balance. That degree that helps me command a good pay rate in a variety of work in which I can set my own hours. I've never had aspirations to have children, so have never had to cope with the economic challenges of parenthood. The rest of my family are economically self-supporting; I don't have dependents. I don't have any expensive chronic medical conditions. I live in a wealthy country that abounds in cheap second-hand cast-offs that are, absolutely if not relatively, items of delightful luxury — and in an internet era in which entertainment and education is plentifully available for next-to-nothing at the push of a button.

But I think my experience — of believing frugality would be harsh medicine, but then discovering it to be sweet nectar once I actually tried some — ought to be encouraging to people who are considering it.

Budgeting

It helps, if you want to be rational and prudent about your budget, to know what that budget is in some detail. You can better discover and correct the inefficiencies in your economic behavior if you have data to base your decisions on. But people often shrink from the task of attending to their incomes and outgoes in anything approaching a rigorous way.

Do you really know how much you earn, even? Sure, you may know your *salary* or your *wage*, but have you done the math — accounting for what gets deducted from your paycheck for this and that, for the various costs of being employed (e.g. commute expenses), but also adding in bonuses and benefits and what having that extra line on your résumé is worth to you?

Do you know how much you spend on this and that? How much of your budget goes to food, and in what ways? How much does your car cost you per mile, once you add up all of the costs — the car itself, insurance, registration, maintenance and repairs, gas, tolls, parking and traffic tickets? How much would that cost have to be to be no longer worth it to you — do you know your figure, even within an order of magnitude?

Many of us only have a foggy idea of the answers to questions like these, and yet those answers are an important way of describing the efficiency of how we live our day-to-day lives: whether what we're doing is really worth the cost.

When I started my experiment in frugality, one of the things I did was to carry around a notebook for one month each year and make a note of every time I spent money: how much, and on what. This accounting helped me to keep my lifestyle affordable on my under-the-tax-line income, and also helps me to better understand where my money is going. ([Here is an example.](#)) I sometimes catch myself ratcheting up my lifestyle in this way, and so, having noticed, can loosen the ratchet and scale down again.

Help with frugality

Personal economics can be a subtle art. It has many facets and is not always intuitive. Fortunately, if you are interested in improving your frugal living skills, there is a lot of help to be had. Here are some examples:

- The book [Your Money or Your Life](#) helps people to consciously reevaluate the way they conceptualize money and then take concrete steps to become more optimally frugal.
- The [Early Retirement Extreme](#) blog (and [wiki](#)) is full of useful tips and insights.
- Discussion fora like [r/LeanFIRE](#) (FIRE = Financially Independent, Retire Early) allow you to share experiences with and learn from others who are in a variety of life stages and situations but all have frugal living and financial security as a goal.

1. [^](#)

Benjamin Franklin, *Autobiography* (1791)

2. [^](#)

Benjamin Franklin, *The Way to Wealth* (1758)

3. [^](#)

Cicero, [Tusculan Disputations](#) (C. D. Yonge translation, 1853)

4. [^](#)

Cicero, *Paradoxa Stoicorum*

5. [^](#)

Charlie Giattino, Esteban Ortiz-Ospina and Max Roser, ["Working Hours"](#) *Our World In Data* (2020 revision)

6. [^](#)

(we have after all just been considering it, Q.E.D.)

7. [^](#)

Charles Dickens, *David Copperfield* (1849) chapter xii

8. ^

Jaimee Bell, "[The psychology of shopping addiction](#)" *Big Think* (3 March 2020)

9. ^

Dolly Freed, *Possum Living: How to Live Well Without a Job and with (Almost) No Money* (1978)

10. ^

"[PNC Advisors Survey Finds that Wealth Brings Complication, Added Responsibilities](#)" The PNC Financial Services Group, Inc. press release, 10 January 2005

11. ^

David Gross, "[How I Stopped Paying Taxes and Started Living My Values](#)"

12. ^

Marlys Harris, "How To Pay Zero Taxes" *Money* (March 2008)

Notes on Dignity

This post examines the virtue of **dignity**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is dignity?

The word “dignity” is commonly used in a few ways, only one of which describes a virtue.

One way is in phrases like [“human dignity.”](#) This is something that is supposed to reside in all people and that makes them worthy of respect. If you mistreat someone in certain ways, you might be said to be “violating their dignity.” In such a case, it’s not that you were mistaken about whether or not this particular person had exceptional dignity worthy of respect, but that you were failing to acknowledge a sort of inalienable baseline dignity that belongs to everyone and the possession of which sets certain boundaries to how you ought to behave toward them. Discussions of dignity in contexts like human rights and medical ethics are usually discussions of this sort of dignity. (Dignity in this sense has a suspicious resemblance to [phlogiston](#): a hypothesized substance that serves as a placeholder for a more complex explanation — less generously, a mere [applause light](#), or [“nonsense on stilts”](#).)

Occasionally “dignity” is used as a sort of honorific, bestowed on somebody as a way of raising them above the common crowd. A “dignitary” for example, is someone who has been officially granted dignity of a sort.

If you succeed in elevating something (an art form, a profession, a group of people) into a position of higher respectability, you might be said to have “brought dignity” to it. (For example, “Zora Neale Hurston brought dignity to African American folk tales by revealing them to be a sophisticated literature.”)

Finally, another use of “dignity” is in phrases like “that would be beneath my dignity” or “to bear trials with dignity.” This form of dignity you don’t just have by default; some people have it and some people don’t. But on the other hand, it is not something that is bestowed on you by someone else or that requires affirmation from others; it seems to be self-generated. This last meaning is the meaning of dignity-as-a-virtue that I want to explore in this post.

Related virtues

Dignity overlaps with [honor](#) somewhat — in the sense of character / pride / reputation / rectitude. Dignity also sometimes goes under the names of self-worth, self-esteem, or self-assurance. The “self-”ness of those epithets highlights that dignity is an internal resource.

Dignity is often associated with poise — a “dignified bearing.” Body language that signifies dignity (sometimes as a literary or theatrical shorthand) includes having an upright posture (not cowering or cringing), an unhurried and deliberate way of doing things (not rushed or clumsy), a proud demeanor (looking people in the eye, not fidgeting or wringing one’s hands), and a steady voice (not shouting or stammering or whining). So grace, gravitas, and decorum all seem to be outward markers of dignity or to be connected with it somehow. That said, they are neither necessary nor sufficient. Somebody who has been injured by disease or accident, for example, may lack some of these superficial markers of dignity but can still show just as much dignity in how they play the cards they are dealt.

[Endurance](#) in the face of pain, [courage](#) in the face of fear, and [temperance](#) in the face of temptation are considered dignified, while someone who wilts easily or who does something unseemly while tempted may be told that they are behaving in an undignified way.

“Quiet” is a modifier often-applied to “dignity” — which contrasts the dignified person with the person who kvetches, whines, rages, makes ostentatious declarations, puts on airs, talks a big game, or has a big hat-to-cattle ratio. So there may be some affinity between dignity and silence, reserve, [temper](#), and serenity.

There is also a somewhat ironic connection between dignity and [shame](#). To have a sense of shame requires that you think some things are beneath your dignity; to be ashamed of something you have done is to believe that it was a stain on your dignity; to be restrained by your sense of shame is to put a high price on your dignity. So shame seems to require dignity, and dignity is protected by shame. But on the other hand, feeling chronically ashamed can make dignity difficult. ([Brené Brown](#) is a popular self-help author and speaker who has made shame her bailiwick. She describes people with resilience to shame — who neither become shameless nor consumed by shame — as “Wholehearted” and gives some advice on how to get there from wherever you’re at.)

There may be some tension between dignity and [humility](#), and in particular where humility meets [piety](#) in self-abasement. On the other hand, some religious people include dignity among the virtues. Christians sometimes reason that as people were made [in God's image](#), only [a little lower than the angels](#), we ought to attempt to reflect the glory of God with our own dignified stature. But it can be hard to maintain a dignified bearing while in humble awe at just how small, weak, ephemeral, and ignorant we are in the grand scheme of things. Isn’t it ridiculous — comic and tragic at the same time — for an overdressed chimp like me to pretend he is dignified?

“[What a piece of work is a man!](#) How noble in [reason](#), how infinite in faculty! In form and moving how express and admirable! In action how like an angel, in apprehension how like a god! The beauty of the world. The paragon of animals. And yet, to me, what is this quintessence of dust?” —Hamlet

Fools’ dignity

“The well-bred are dignified but not pompous. The ill-bred are pompous, but not dignified.” —Confucius^[1]

Pomposity, braggadocio, fronting, and machismo are among the ways people without dignity try to fake it, but to the discerning such things are just evidence for an

absence of dignity. Their flaw is in part that they show the braggart to be conditioning their dignity on the respect shown to them by the sort of people who might be impressed by their bluster, whereas a person with genuine dignity would not be bothered by what easily-swayed randos might think of them.

Because such blowhards seek to impress rather than to become the sort of person they ought to admire, they become mere performers, insecure every time the applause dies down. In thrall to their audience, they cannot stand on their own feet, and real dignity eludes them.

Sometimes also people try to assert their dignity by being domineering, in a way that ultimately isn't very dignified. [Donald Trump's cringeworthily aggressive handshakes](#) are one example, and *South Park* Cartman's "[respect my authoritah!](#)" is another.

Dignity can also overshoot the mark and become a snobbish sense of superiority, or a supercilious stuffiness ("I would not lower myself to see some silly comic book movie"). Aristotle's [great-souled man](#) is proud, has tremendous self-regard, believes himself to be deserving of all honor, and has earned every bit of it.^[2] He also doesn't seem like he'd be much fun to be around. He has dignity, to be sure, but not of a sort most people would be wise to imitate.

Dignity and “Pride”

For people in low-status castes or oppressed groups, assertions of dignity can be political statements. “Gay Pride,” “Black & Proud,” and so forth try to defeat or invert the effect of being in the politically disadvantaged group by refusing to go along with the idea that there is any lesser dignity to being in the group.

For people in higher-status groups, superficially similar assertions of pride (e.g. “White Pride”) operate differently. Some people in those groups may use “pride” as an attempt to leverage the extra privilege and status that the group has obtained as though it were something that could give them dignity as individuals.

Groups will sometimes use this borrowed-dignity temptation to cheaply buy off people who don’t have much dignity of their own — by offering them low-value tokens that ostensibly confer “official” dignity: titles, ranks, insignia, awards, positions, and things of that sort.

Behaving in an undignified way can sometimes reflect badly on groups you belong to, and this can also become tangled in questions of group status and politics. For example, a character like [Stepin Fetchit](#), who played up degrading stereotypes for laughs, came to be seen as not only an undignified role for the actor to play, but as something that was harmful to the dignity of black people in general and so a sort of betrayal or lapse of solidarity on the part of the actor.

(People who attempt to display dignity and fail are stars of many comic tropes. Such scenarios are put to good use by many comedic actors, consider Charlie Chaplin or Peter Sellers for example. The struggle to maintain dignity in trying circumstances seems to strike a nerve with people, and seeing the pompous brought down to earth is also popular.)

Becoming more dignified

A dignified person puts a high value on their character and stature, and so they will not stoop to scrape for less valuable things. This suggests that in order to be dignified, it helps if you have your basic needs (in a [Maslow's hierarchy](#) sense) well in hand, so you don't have to "swallow your pride" to make ends meet.

Dignity can be threatened by being in another person's debt or by being reliant on somebody else. This makes self-reliance an important ingredient of dignity. If you live paycheck to paycheck, you may feel less able to tell your employer to take a hike if they ask you to do something beneath your dignity. Having a financial cushion — in this context, sometimes described as "fuck you money" — enables you to value your dignity higher than your job. If you place a high value on your dignity, you'll be more motivated to save up money to better insure it in this way.

If your [ambitions](#) require you to flatter, kow-tow, or suck-up to people with power, money, or authority, or to be a crowd-pleaser, you may find your dignity and your ambitions at cross-purposes. It's a good idea to consider the potential cost in dignity of the ambitions you decide to pursue.

The influence of intoxicating drugs, alcohol in particular, can cause people to shed their dignity. And an addiction can tempt a person to do undignified things to feed the monkey on their back. So if you want to be more dignified, [caution](#) around drugs is warranted.

If you catch yourself behaving in an undignified way, see if you can find a pattern. There may be another virtue you can strengthen that will help your dignity as a side effect. If you lose your dignity when you are tired or uncomfortable, work on your [endurance](#). If you lose your dignity when you are tempted by something you want, work on your [temperance](#). If you lose your dignity when you are frightened, work on your [courage](#). If you lose your dignity when you are angry, work on your [temper](#).

People sometimes behave in an undignified way under the sway of strong emotions. Knowing how to feel an emotion fully and genuinely without letting it push you beyond respectable bounds is an art that takes practice to get right. If you are in the habit of suppressing strong emotions, you may find yourself out of your depth when emotion breaks through your defenses; on the other hand, if you use strong emotion as an excuse to throw off all restraint, you'll likely develop the bad habit of sacrificing your dignity in the process.

1. ^

Analects of Confucius, XIII.xxvi

2. ^

Aristotle, *Nicomachean Ethics*, book IV, section 3

Notes on Courtesy

This post examines the virtue of **courtesy**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is courtesy?

The best concise definition of courtesy I have found is this one, from H.E. Norton, honorary secretary of the charmingly-named "Children's National Guild of Courtesy": "kindly and thoughtful consideration for others."^[1]

Courtesy is closely related to things like manners, etiquette, politeness, decorum, and propriety. If you go hunting for advice about courtesy (as I did), much of what you find will be guides to polite manners for various occasions. Some of these are more nuanced than others, but many seem to be catering to a hope that courtesy can be reduced to a set of rules that, once you learn them, will keep you out of trouble. But heuristics, norms, and rules of thumb will only get you so far. "Kindly and thoughtful consideration for others" requires attentiveness, discernment, and creativity.

Etiquette often has more to do with behaving conventionally than thoughtfully. That said, such conventions sometimes have courtesy as a goal, and often seem meant to be safe defaults that approximate kindly and thoughtful consideration for others in common situations. Other conventions have other purposes, such as to show submission to people in dominant classes ("always doff your hat to a duke"), or to signal your membership in (or aspiration to) a class ("a well-bred gentleman never..."). The origins of some conventions are obscure, but they can still be useful signals of kindly-and-thoughtful intentions. For some conventions, even if they have an element of the arbitrary to them, just that they *are* conventions makes it kindly and thoughtful to follow them ("slow drivers stay in the right lane"). However, you should be prepared to sacrifice conventional etiquette for actual kindly and thoughtful consideration ("I apologize for not shaking your hand; I'm trying to be especially cautious during the pandemic").

"Honorifics and formal politeness provide lubrication where people rub together. Often the very young, the untraveled, the naïve, the unsophisticated deplore these formalities as 'empty,' 'meaningless,' or 'dishonest,' and scorn to use them. No matter how 'pure' their motives, they thereby throw sand into machinery that does not work too well at best." —Robert Heinlein^[2]

Because of this element of the arbitrary, different dialects of manners emerge in different cultures and subcultures. A wise person will not mistake the dialect he or she knows most fluently for the universal standard by which others are judged.

The purpose of courtesy is often to make other people feel at ease. The urbane host and the tactful conversationalist gracefully and unobtrusively smooth over any rough spots that might interfere with social interactions going well.

Other times, courtesy is a form of [compassion](#) — opening the door for someone carrying a package, offering to let someone go ahead of you in line at the grocery store if they seem in a hurry. Courtesy is usually mild and gentle, always kind.

[Civility](#) is a form of courtesy, along with charity (in the sense of interpreting others' actions in the best plausible light, or of [steelmanning](#) their arguments).

Courtesy and other virtues

Many other social virtues are themselves courteous to exhibit, or their absence is discourteous: things like [appreciation](#), sympathy, remembrance, concern, *muditā*, [goodwill](#), agreeableness, geniality, graciousness, modesty, warmth, [sincerity](#), [honesty](#), [gratitude](#), [respect-for-others](#), magnanimity, hospitality, cooperation, conciliation, [fairness](#), forgiveness, willingness to accept fault, openness, understanding, friendship, fashion sense (in certain contexts, e.g. a funeral), reciprocity, dependability, [loyalty](#), unpretentiousness, discretion, tact, and conversational competence in general, including good listening.

In my own experience, I've noted that my lapses in courtesy have often been a result of a failure of [attention](#): An occasion to be courteous arose, but because I was absent-minded, I didn't notice until too late. I've been able to improve on this slowly by doing regular mindfulness meditation, and also by getting in the habit of periodically reminding myself to come back down to earth when I'm lost in thought.

The trouble with being polite

Norms of civility can sometimes be polite masks for attitudes that [defend the status quo against unpopular ideas and low-status people](#). This is especially true of norms of etiquette that are designed to help people to pass as being members of high-status groups. For this reason, [beware isolated demands for courtesy](#).

Being polite requires that you accommodate yourself to other people somewhat, but it can be tricky to know how far to bend without breaking. It's probably courteous not to insult another person's religion, but ought you to go further and avoid violating certain of their religious taboos in their presence? must you join them in their rituals to avoid giving offense? It's probably a courteous [kindness](#) to laugh at another person's jokes whether or not you find them all that funny... but what if they tell a joke that demeans others or is offensive to your dignity? Courtesy can run into boundaries defended by other virtues, and it requires [wisdom](#) and discernment to decide how accommodating you can be when that happens. How to say "no" politely, and how to be courteously assertive, are skills that can make courtesy easier to practice.

Some forms of "political correctness" are complex dialects of etiquette that act as status markers. If you are well-educated and keep up to date on intellectual trends in the right circles, you'll know which phrases (e.g. "all lives matter") and ideas (e.g. "sexual preference") have gone out of fashion and now mark you as politically backward or as disrespectful. This is another balancing act. On the one hand, it is good to learn ways in which the language you use may embed assumptions that inadvertently offend, so you can correct for this; on the other hand, the roiling fashions of what is and isn't correct language can become detached from concerns of kindness and respect and can become mere shibboleths with the more sinister purpose of tarring the unfashionable out-group with a broad brush.

Politeness can be mockingly ironic, merely [insincere](#), or more self-consciously [deceptive](#). Such things are not merely harmful in the immediate context, but can degrade the value of politeness more generally by making politeness a less reliable indicator of genuine courtesy. Someone who gets a reputation for being insincere in their displays of politeness may find that it becomes increasingly expensive for them to convincingly signal courtesy and goodwill.

Politeness, if it has mercenary motives, is just a sort of flattery or obsequiousness: kissing up to those above you on the ladder in the hopes of getting a boost or joining the winning team. It is one of the hazards of being in situations where political ambition or cutthroat social climbing rule the roost that politeness becomes less sincere or more obviously transactional.

What good is courtesy?

“The small courtesies sweeten life, the greater enoble it.” —Christian Nestell Bovee^[3]

Courtesy, done right, is nice to people. It makes them feel more comfortable, more respected. It tends to evoke courtesy in others, which can rebound on you, which can make you feel more comfortable and respected in turn, so wins all around. Courtesy also tends to evoke a like response: one flame can kindle a flame war.

Courtesy suggests goodwill and invites goodwill from others. It communicates the absence of threats or hostile tensions in the environment. In such ways, courtesy helps social interactions go smoothly. If you are courteous to others, they will be more likely to enjoy and to seek out your company, which is nice if you like company. Even if you don't much like company, it's usually easier to tolerate in a genuinely courteous atmosphere.

If you are courteous, this signals that you are well-brought-up and well-socialized. It also marks you as self-sufficient and doing well (if you can concentrate thoughtfully on kindly consideration for others, you likely are not overly-concerned about your own needs). It is a signal to others that you have your shit together. This can be attractive to others and also suggests that you won't be a burden on them, which can make people more eager to cooperate with you. On the other hand, being discourteous can make you an object of ridicule and contempt (the recent [“Karen” caricature](#) is an example).

How to become more courteous

To strengthen the virtue of courtesy, I suggest the following: First, there are certain prerequisites that make courtesy easier to practice. Self-sufficiency and contentment give you the slack that makes it easier to pay attention to others. That attention also requires mindfulness, the ability to see what's in front of your eyes. Along with [empathy](#) and [respect-for-others](#), attention can help you identify opportunities to be considerate. Strengthening any of those virtues will help you to be more courteous in general.

Remember also that list of social virtues I gave earlier. Any of those may come into play on a specific occasion of courteousness. There are also a multitude of even more specific guidelines that don't rise to the level of virtues but are important habits of

politeness — things like being on time, knowing how to apologize, not monopolizing conversation, taking good care of things you borrow, lending a hand when one is needed, and so forth... more than I can list. [Books of etiquette](#) can help with some of this.

It's good to learn the local etiquette dialect of your place and time, and some adjacent ones too. A more cosmopolitan curiosity can help you learn which parts of your local dialects are just local, which might save you from some embarrassment when you encounter someone from another culture.

You can also get hints from observing what other people do, and what they approve and disapprove of in others. Carefully identify people whose manners you admire, who tend to make people around them comfortable without sacrificing their dignity. Observe and learn. What would you have done in their situation, and what did they do instead, and how do you explain the difference? Every once in a while I stumble on an old book that has a character who is exceptionally good at genuine courtesy, and that can be another good example that I can use to compare my instincts with.

And while I don't want to put too much emphasis on this, I feel I should mention that it might help to get stoned once in a while. I made some leaps-and-bounds changes in my personality as a young stoner when, while good and high, I was able to see my own actions from an outside perspective while at the same time introspecting the hell out of the triggers and habits that led to those actions. I saw some things I didn't admire and so I put in the work to change them. I don't know that marijuana is any sort of panacea for this sort of thing, but I think in my case I was able to use it to make progress that would have taken me a lot longer to reach otherwise.

1. ^

H.E. Norton, [*Courtesy: A Reader for Older Boys and Girls*](#) (1913 ed., p. 15)

2. ^

Robert Heinlein, *Time Enough For Love* (1973)

3. ^

Christian Nestell Bovee, *Intuitions and Summaries of Thought* (1862)

Notes on Chastity

This post examines the virtue of **chastity**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

What is chastity?

Chastity is usually defined as abstention from sex, or restriction of sex to certain moderate bounds, combined with a [temperate](#) regulation of the sexual appetite.

[G.E.M. Anscombe](#) asserted in 1972 that “chastity is simply the virtue whose topic is sex, just as [courage](#) is the virtue whose topic is danger and difficulty.”^[1] But by 1972, and certainly by now, people seem much more concerned with other sexual virtues: is your sexuality *authentic* to your desires? is it *generous* to your partners? is it [respectful](#) of other people? is it [skillful](#)? is it sufficiently [cautious](#)? Chastity almost never makes the list anymore.

Chastity is related to temperance, [continence](#), modesty (“the outpost and safeguard of chastity,”^[2] according to *The Catholic Encyclopedia*), purity, decency, and [dignity](#). It is sometimes confused with “celibacy,” which is usually defined as complete abstinence from sex and marriage, for instance as part of the qualifications for joining a monastic or priestly order.

What happened to chastity?

From the looks of things, chastity has long been a virtue more honored in the breach than in the breeches. But for a long time people at least paid lip service to it and thought that it was a virtue. Nowadays (except in the case of schoolchildren, who are sometimes asked in abstinence-only sex education classes to pretend as though adults value chastity), it's rare to see chastity mentioned as an important character trait. So what happened? The whirlwind version of [the just-so story](#) I've heard about Euro-American culture goes something like this:

Even when chastity was held in esteem, there has always been a libertine subculture who did not buy it, believing that those who urged chastity were essentially killjoys who had an irrational disgust for sex and thought everyone else should too. This subculture waxed and waned in its influence.

The collapse of the hegemonic Catholic church as the gatekeeper to legal, officially chaste sexuality, meant that people were more encouraged to think for themselves about what virtuous sexuality meant, and different protestant sects could use their interpretations of what chastity consists of as ways of distinguishing themselves in the marketplace.

The importance of strategic considerations in marriage (social climbing, family alliances) has been declining, and now the cultural norm is to marry for [love](#), based on mutual romantic (and sexual) attraction. Marriage is seen less as an alliance or partnership, and more of an ongoing romantic adventure, in which sexual satisfaction is vital. In that context, premarital sex is an important part of courtship as it tests sexual compatibility.

Freudian psychology pathologized sexual inhibition as the cause of maladies, and painted societal and institutional guardrails of chastity as things that were interfering with people maturing into healthy sexual adults. The libertine subculture used this as propaganda to gain more respect and to advance their creed, until it became the culture and chastity was relegated to a traditionalist subculture.

The automobile helped to normalize incorporating sexual trysts into dating and courtship. Mass media noticed that sex sells and kept pushing the boundaries of what could be said in public in order to gain eyeball-share. Civil libertarians sometimes added substantive defenses (of e.g. pornography, homosexuality, contraception, divorce) to their libertarian legal defenses, often using psychologized terminology, to the effect that ever-more promiscuous and polymorphous forms of sexuality became normalized and chastity and sexual modesty became further pathologized and relegated to the realm of the prude and the censor. As people discussed sexuality in all of its variety more publicly, they became more aware of its range and contours, and a sexual connoisseurship developed, [\[3\]](#) which further expanded the market for ever-more-expansive descriptions of human sexuality, until what was depicted as normal had gone well beyond what would traditionally have been considered chaste.

The Pill (and convenient contraception and abortion more generally) mostly eliminated one of the most tangible reasons why the virtue of chastity might be important to a flourishing life, while meanwhile reductions in childhood mortality made contraception seem more desperately vital to not being swamped by a deluge of human larva. Feminists identified traditional sexual mores as having been part of the mechanism of the subjugation of women, and exposed chastity to critique on those grounds. People learned to identify themselves with their sexuality ("I am gay"), and so sexual identities and the freedom to express them became another front in civil rights battles. The civil libertarian argument that consenting adults ought to be able to please one another according to taste, combined with an emerging awareness of the extent and harm of non-consensual sexual behavior, refocused cultural discussion of sexual morality around consent and allowed people to remoralize sex while leaving chastity out of it. The internet made vast quantities of pornography available in every variety at the push of a button, and made it nearly impossible to maintain the pretense of chaste societal norms for the supposed benefit of children.

And so here we are. Chastity is defeated and the libertines have won. I have internalized the arguments of the libertarians and the libertines at least as much as anyone else in our society, and even the word chastity is hard for me to pronounce without feeling like a prude. I think that the attention we have given to redefining sexual morality for the modern era has been, by and large, a good thing. And yet I wonder if we lost anything important along the way.

The benefits of chastity

"Chastity does not pretend to extinguish our tender passions, or cancel one part of our nature: it only bids us not to indulge them against reason and truth; not give up the man to humor the brute, nor hurt others to please ourselves; to divert our inclinations by business, or some honest amusement, till we can gratify them lawfully, conveniently, regularly; and even then to participate of the mysteries of love with modesty, as within a veil or sacred enclosure, not with a canine impudence." —William Wollaston^[4]

I tried to hunt up some good, sophisticated, modernish defenses of chastity. Most of them assume a religious perspective in which one of the most important reasons you remain chaste is because God likes it when you do. Some others are defenses of chastity as a chosen lifestyle or even identity (people who don't have sex but wish they could now are "incels" and some people who choose abstinence call themselves "volcels" in contrast). I tried to tease out of all of them some reasons why chastity-as-a-virtue might be important to human flourishing.

There are some practical, biological reasons: unintended pregnancies and sexually-transmitted diseases are still a thing. But this seems more an argument for prudence than chastity.

Chastity is supposed to help reinforce and preserve the value of marriage. G.E.M. Anscombe put it this way: "Humanly speaking, the good and the point of a sexual act is: marriage. Sexual acts that are not true marriage acts either are mere lasciviousness, or an Ersatz, an attempt to achieve that special unitedness which only a real commitment, marriage, can promise. For we don't invent marriage, as we may invent the terms of an association or club, any more than we invent human language. It is part of the creation of humanity and if we're lucky we find it available to us and can enter into it. If we are very unlucky we may live in a society that has wrecked or deformed this human thing."^[1]

Similarly, chastity can make sex itself more special. Imagine if you only had chocolate once a year, on your birthday. That would potentially increase the value you get from that chocolate, and from your birthday. Maybe that extra value would be worth the cost of forgoing chocolate the rest of the year. Similarly, by restricting your sexuality to a particular partner in a particularly sanctified context, you add value to the sex, the context, and the partner, and conceivably that is worth the trade-off from abstaining from sex otherwise. By contrast, if you indulge in sex whenever and with whomever the fancy strikes, in a "(merely) natural" way, a part of your life that could have been extraordinary risks becoming no more remarkable than yawning or getting a haircut.

Some people insist that indulging in too much sex has harmful side effects, just as indulging in too much ice cream or alcohol does. "Chastity is one of the greatest disciplines," wrote Gandhi. "A man who is unchaste loses stamina and becomes emasculated and cowardly. If his mind is given over to animal passions, he is not capable of any great effort."^[5]

The reference to "animal passions" (or Wollaston's more colorful "canine impudence") points at another criticism of sexual indulgence: that it is the silly pandering to base instincts meant to prompt our reproductive success, at the expense of the higher callings our species is capable of. Why waste your human time questing for new opportunities to rut like a reptile when you could be devoting that energy to the more exceptional pursuits our weirdly wrinkled cerebrums allow?

Often people make terrible [decisions](#) under the influence of lust. People will degrade themselves, put themselves in compromising positions, wreck relationships, trash careers, behave shamefully, all for another few minutes of the bouncy-bouncy. The promise that we will be free from neurosis and shame if we just get rid of our hang-ups and let our sex drives guide us wherever they naturally want to go has turned out to be poor advice. Chastity can be a way of trying to reassert [rational](#) control over irrational drives, so that they serve us rather than us being served up to them.

Chastity is also sometimes recommended because it is harsh medicine. [The Catholic Encyclopedia](#) notes with approval the “real as well as etymological kinship between chastity and chastisement” and that “chastity is a thing stern and austere.”^[2] Chastity can be a sort of ongoing discipline that teaches temperance, self-control, and moderation. The sex drive can be viewed as a sort of exercise equipment: it provides the resistance we pull against in order to train ourselves in better self-discipline.

Another possible pro-chastity position would be to assert that the emerging modern consensus about sexuality *is* the new “virtue whose topic is sex” and so just is the modern version of chastity. To be chaste in this view is to be authentic, generous, open-minded, prudent, respectful, and skillful, or perhaps some other combination of things we’re still in the process of discovering. A virtue is meant to help us flourish, and as the world changes around us, what helps us flourish changes too.

How to develop chastity

Da mihi castitatem et continentiam, sed noli modo
[“Give me chastity and continence, but not just yet! ”]

—Augustine of Hippo^[6]

To become chaste seems therefore like it will require, more so than for the other virtues, spending some effort to define what chastity means for you. Traditional ideas of chastity may no longer fit modern realities, while modern ideas may be more fashionable than practical. How do you envision a healthy sexuality that helps you to lead a flourishing life? How realistic is that — how reality-based vs. based on moth-eaten ideas or pop culture fantasies?

The next step will be to develop a way to honor and respect your sexual drives, desires, fantasies, and inclinations while at the same time subordinating them to your larger project of being a flourishing person. (That ought to take at least couple of minutes, easy.) You might examine your past and present sexual behavior with an eye to discovering where it went against your better interests. Did you fail to think things through? Did you write your lust a blank check? Did your self control desert you at the last minute? Did you exaggerate how much satisfaction you would get in return? Are you trying to get something through sex that you would be more apt to find in some other way? In what way can you correct for this next time?

“Much of the battle of building better habits comes down to finding ways to reduce the friction associated with our good habits and increase the friction associated with our bad ones.” —James Clear^[7]

If you find your sex drive to be more than a match for your self control, then avoiding temptation may be important. Stay out of scenarios in which your weak self control is the only thing standing between you and a decision you will regret. Manipulate your

environment in such a way that the thing you hope you'll do is also the default, easiest, and most tempting thing to do.

See also:

- Notes on Temperance: [Can we shape our desires?](#)
- Notes on Self Control: [How do we lose self control, and how can we strengthen it?](#)
- Notes on Endurance: [How to improve endurance](#)

1. ^

G.E.M. Anscombe, "Contraception and Chastity" *The Human World* 7 (1972)

2. ^

John Melody, "Chastity" *The Catholic Encyclopedia* vol. 3. (1908)

3. ^

"Another startling aspect of [the 1948 Kinsey Report on male sexuality] for a contemporary reader is his stark elitism—the lower classes do not know how to do it and have no imagination. Only the educated can liberate themselves from mythology and can think through the differences between plain and fancy sex."—Allan Bloom *Love & Friendship* (1993) p. 17

4. ^

William Wollaston, *The Religion of Nature Delineated* (1722), chapter 9 ("Truths Belonging to a Private Man, and Respecting (Directly) Only Himself")

5. ^

M.K. Gandhi, *Hind Swaraj* (1909), chapter 17 ("Passive Resistance")

6. ^

Augustine of Hippo, *Confessions*, book VIII, chapter 7

7. ^

James Clear, *Atomic Habits* (2018), chapter 12 ("The Law of Least Effort")

Notes on Love

This post examines **love** as a virtue, as part of a sequence of posts about virtues. I mostly explore what other people have learned about love, rather than sharing my own research or opinions about it, though I've been selective about what I found interesting or credible, and my research has been scattershot. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will help people who want to know more about love and how to practice it skillfully.

What is this thing called love?

"Love" suffers from more than its share of terminological difficulty. People use the word to refer to all sorts of things—including sensations, emotions, dispositions, commitments, actions, judgments, predicaments, opinions, principles, personified or deified forces, and people. The love that people experience when they fall-in-love is very different from the Christian "love your enemies" sort of love or the love that parents have for their children or the love that a devotee has for God or the love that a gourmand has for pancetta or the love that a patriot has for her country or the love that misery has for company.

And I don't think there is a primary meaning of love, with a bunch of secondary meanings hooked to it metaphorically. As a quick sanity-check, I did an informal poll of Facebook friends in which I asked, without further context, how they would sum up "love". About half of the serious responses described love as a feeling. Another 25% described it as a sort of disposition (e.g. "the desire for the beloved to thrive," "not thinking twice about sticking around during the tough times"). A few described it as a sort of activity (e.g. "helping"). Two examples did not attempt to describe what love is, but instead described its place in your life ("love completes you", "it's all you need"). Some answers were ambiguous (e.g. is "[caring](#)" a feeling or an action?). This reinforced my impression that "love" does not have an obvious central definition with other secondary meanings, but that it is a scrambled mishmash of a lot of things.^[1]

While love is not well-defined it is nonetheless a very popular topic of inquiry in just about every literary variety. It is nigh impossible to do even a cursory literature survey of the topic. As a result, mine has been eccentric and there's a good chance I would have come up with some very different conclusions if I had begun my unraveling by pulling some other of the many threads I had to choose from.

I would be tempted to chuck love off the list and discuss more well-defined virtues instead, except that in the Western tradition in the Christian era, "love" tends to show up on lists of virtues often, and sometimes tops the list. So, though wise men say only fools rush in, I can't help flailing in love for you.

Virtues of love

Of the many uses of the word "love," several might plausibly describe virtues (that is, habits characteristic of human flourishing), including:

1. the indiscriminate attitude of love Jesus talked about in phrases like "love your enemy" and "love your neighbor" and "love one another"

2. the more discriminating and particular sort of love you might have for a close friend
3. romantic “falling in love”-style love, and romantic partner-bond love (maybe [“courtly love”](#) fits in here too, but I’m going to ignore it for now)
4. love of God
5. self-love
6. love for life, love of one’s calling, *amor fati*, and things of that nature
7. love of one’s country
8. love of beauty, love of nature
9. the nurturing love that a parent shows for a child and the reciprocal love of the child for the parent, and the more horizontal love between peer family members (e.g. siblings, adult children and parents, maybe pets?)

The last four of these I will save for possible future write-ups about enthusiasm/zest, patriotism, aesthetic appreciation, parenting/family-making, and filial piety. Love of God I touched on in my [Notes on Piety](#). Self-love strikes me as a metaphorical use of the term, but I covered some of its aspects in my [Notes on Dignity](#); I might also take on “self esteem”/“pride” at some point. That leaves the first three, which I will try to briefly examine in this post.

Christian love: *agape*

“Why love among the virtues is not known
Is, that love is them all contract in one.”
—John Donne^[2]

Jesus made “love”^[3] central to the ethics that he taught. For example:

[A Pharisee asked] “Teacher, which is the greatest commandment in the Law?” Jesus replied: “ ‘Love the Lord your God with all your heart and with all your soul and with all your mind.’ This is the first and greatest commandment. And the second is like it: ‘Love your neighbor as yourself.’ All the Law and the Prophets hang on these two commandments.” —[Matthew 22:36-40](#)

“Love your enemies, do good to those who hate you, bless those who curse you, pray for those who mistreat you.... If you love those who love you, what credit is that to you? Even sinners love those who love them.... But love your enemies, do good to them, and lend to them without expecting to get anything back.” —[Luke 6:27-35](#)

“A new command I give you: Love one another. As I have loved you, so you must love one another. By this everyone will know that you are my disciples, if you love one another.” —[John 13:34-35](#)

What did the Christians mean by this *agape*-style love? Luckily for us, Paul gives a very quotable definition:

“Love [suffers long](#) and is [kind](#); love [does not envy](#); love [does not parade itself, is not puffed up; does not behave rudely, does not seek its own, is not provoked, thinks no evil; does not rejoice in iniquity, but rejoices in the truth; bears all things, believes all things, hopes all things, endures all things.](#)” —[1 Corinthians 13:4-7](#)

However, from this, it seems that “love” isn’t a virtue, but an amalgam of more than a dozen. I see at least good temper, kindness, *muditā*, modesty, humility, courtesy, respect for others, forbearance, benevolence, justice, rationality, piety, hope, and endurance in those verses. You can find even more virtues to add to that list in certain [other passages](#).

A tempting interpretation is that “love” is a shorthand way of referring to the package of Christian social virtues. For example, perhaps it would make sense to say something like “you could use a little more modesty in your love, brother.”

Another possibility is that love is a master-virtue that unlocks these many others: St. Jerome, for example, called love “the mother of all the virtues.”^[4] Paul suggested that “love” is a distinct thing—something fundamental that forms a foundation for other virtues,^[5] or maybe an essential ingredient that makes the other virtues really virtuous:

“If I speak in the tongues of men or of angels, but do not have love, I am only a resounding gong or a clanging cymbal. If I have the gift of prophecy and can fathom all mysteries and all knowledge, and if I have a faith that can move mountains, but do not have love, I am nothing. If I give all I possess to the poor and give over my body to hardship that I may boast, but do not have love, I gain nothing.” —[1 Corinthians 13:1-3](#)

One way to look at this is that love *aims* the other virtues correctly.^[6] For example, an intrepid jewelry thief might exhibit plenty of “courage,” but if the goal of her courageous act is not coherent with love, she is only practicing a sort of misfire of the real virtue. In this interpretation, Christian love isn’t the introduction of a new virtue, but a redefinition of what virtue *is*.^[7] In the old Aristotelian sense, a habit is a virtue if it tends to exhibit the worldly human flourishing of the person who practices it. In the new Christian sense, a habit is a virtue if it harmonizes with love (of God, primarily, and of one-another, secondarily).

Iris Murdoch came to a similar conclusion from another (not explicitly Christian) direction. In her metaphor, rather than the virtues being built on a foundation of love, they build towards love at their apex:

If we reflect upon the nature of the virtues we are constantly led to consider their relation to each other... For instance, if we reflect upon courage and ask why we think it to be a virtue, what kind of courage is the highest, what distinguishes courage from rashness, ferocity, self-assertion, and so on, we are bound, in our explanation, to use the names of other virtues. The best kind of courage (that which would make a man act unselfishly in a concentration camp) is steadfast, calm, temperate, intelligent, loving... This may not in fact be exactly the right description, but it is the right sort of description.^[8]

The way the virtues reflect and refer to each other suggests to Murdoch an underlying unity, which she identifies with love, which she further describes as “the direction of [attention](#)... outward, away from self... towards the great surprising variety of the world.”^[9]

Modern writers who attempt to explain *agape*, both from within the Christian tradition and from outside of it, continue to propose diverse, imprecise, and sometimes incompatible definitions of what it entails.^[10] I think I will simply recognize this rather

than try to add to the pile with a definition of my own, but there are some common themes that I'll consider in a broad-brush way.

Often *agape* has connotations of altruism and selflessness. It may be a self-serving virtue, but only incidentally, not intentionally.^[11] You practice *agape* for them (or for Him), not for yourself.

Christian *agape* is indiscriminate. You are to love one another, love your neighbor, love your enemy... nobody seems to be left out. The Christian is supposed to disconnect love from evaluation and from relation: [Love 'em all and let God sort 'em out.](#)

This sort of love is not a two-way-street the way other varieties are. Romantic love or friendship love is paradigmatically a love *shared between* two people, where each person in the relationship participates lovingly. Christian *agape*, on the other hand, seems to come entirely from the Christian (or from God with the Christian as a conduit), and radiates onto the recipient without the recipient's active participation. This is how you can love your enemies even as they fail to love you.

Agape is gratuitous: The recipient of *agape* does not *earn* it or *deserve* it. That can make being on the receiving end of *agape* ambiguous. C.S. Lewis wrote:

We want to be loved for our cleverness, beauty, generosity, fairness, usefulness. The first hint that anyone is offering us the highest love of all is a terrible shock. This is so well recognised that spiteful people will pretend to be loving us with Charity precisely because they know that it will wound us. To say to one who expects a renewal of Affection, Friendship, or Eros, "I [forgive](#) you as a Christian" is merely a way of continuing the quarrel. Those who say it are of course lying. But the thing would not be falsely said in order to wound unless, if it were true, it would be wounding.^[12]

This is kind of paradoxical. On the one hand we may long to be loved unconditionally, so that we would be secure in the love no matter who we become or what we do. But we *really* long to be loved conditionally on the condition that we're marvelous, so as to confirm our marvelousness. We don't really want to be loved "anyway," in spite of ourselves.

The gratuitousness of *agape* suggests one way it can be evidence of a flourishing life. Erich Fromm:

For the productive character, giving... is the highest expression of potency. In the very act of giving, I experience my strength, my wealth, my power. This experience of heightened vitality and potency fills me with joy. I experience myself as overflowing, spending, alive, hence as joyous. Giving is more joyous than receiving, not because it is a deprivation, but because in the act of giving lies the expression of my aliveness.^[13]

It is not clear to me what *agape* entails. What is different about someone with it from someone without it? There seems to be no consensus about this, and sometimes Christians have what might charitably be described as counterintuitive notions: The Roman Catholic chaplain who blessed the mission that culminated in the atomic bombings of Japan, or the Russian Orthodox patriarchs who chose a patron saint of nuclear weapons, apparently believed Christians ought to love their enemies into smithereens by the thousands, for example.^[14] It does not seem to be helpful to

define *agape* as “that thing Christians do to their neighbors and enemies” in the same way that you might describe the [Hajj](#) as “that thing Muslims do when they go to Mecca.” If someone says they are full of Christian love for me, I cannot help but wish they could be more specific.

One possible explanation is that *agape* is a god-like love—the sort of love God gives to His creatures—and that is why it can seem strange and unnatural to us, and why even Christians have a hard time approaching it. When Jesus instructed his disciples to practice *agape*, this was perhaps an extension of his [“Thy will be done”](#) teaching: our *agape* towards each other (our neighbors and our enemies) is ideally God’s *agape* expressed through us, unfiltered by our sinful mortal discriminations and limitations. In this interpretation, by channeling God’s love to his creatures in an unimpeded way, we can more closely approach God.[\[15\]](#)

Although precisely what *agape* entails is difficult to pin down, it seems to suggest goodwill, kindness, and that sort of thing: the benevolent basket of goods Paul aphorized. That Christians are to behave in such a way towards everyone, even their enemies, gives Christianity a utopian appeal. It invites you to imagine yourself filled to overflowing with such attractive qualities, unable to hate anyone because of all the love getting in the way, and working side-by-side with other such delightful people to spread this gospel. It invites you to imagine a world in which you are surrounded by your loved ones all the time, because you love everyone.

I remember as a Jesus-curious boy at YMCA camp singing [“they will know we are Christians by our love”](#) and thinking that sounded delightful. It is an audacious doctrine, and (at least in a generous imagination) has appealingly revolutionary, world-changing implications. In my subsequent experience with Christians, however, it has proven difficult for me to discriminate them from others by their love. I have met Christians who both demonstrate extraordinary goodwill, benevolence, kindness, and such and who say they are inspired by their faith to do so; it’s just that this is more the exception than the rule, in my experience, and they don’t seem to turn up any more typically than their non-Christian counterparts who find some other motivation for being that way. That is just my unsystematic observation. However I think if Christianity were indeed a reliable way of provoking *agape* of the sort Christ describes, I probably would have noticed by now.

But Christ didn’t say his path was an easy one that most anyone could find their way to.[\[16\]](#) That most Christians fail to practice *agape* in the way he taught may just indicate that it’s hard to do. It still might be worth the attempt. But it’s daunting to note that even people who believe their eternal felicity depends on it apparently find it impractical.

The love of friendship

“Friendship is unnecessary... It has no survival value; rather it is one of those things which give value to survival.” —C.S. Lewis[\[12\]](#)

Before Jesus made *agape* such a big deal, there was *philia*, the love between close friends.[\[17\]](#)

Aristotle, whose *Nicomachean Ethics* is otherwise remarkably self-focused, devotes two books of that work to friendship, and describes the best variety of friendship in explicitly other-focused terms. In his telling, true friendship exists when two people

each wish for each other's good, and are both aware of this mutual relationship of goodwill.[\[18\]](#)

Unlike Christian *agape*, *philia* is discriminate: you only practice it towards a few people, and only towards those with whom you have developed an appropriate relationship. You can no more love everybody (in this way) than you can make every flavor of ice cream your favorite.[\[19\]](#)

Aristotle considered true friendship to include a regard for one another that is like the regard for oneself that a virtuous person has. A virtuous person has integrity, and wishes what is actually good for himself, for his own sake. In this way, he is like a friend to himself, for a friend will be in harmony with a friend, and will wish for that friend what is actually good for the friend's sake.[\[20\]](#) People deficient in virtues, on the other hand, are in conflict even with themselves (their appetites conflict with their reason, and so forth), and so they don't have a good foundation on which to build good friendships.[\[21\]](#)

Practicing the other virtues makes friendship easier to establish in other ways, too. For example, if you are trustworthy, another person can more confidently put trust in you, which is a prerequisite for this sort of friendship.[\[22\]](#) This indeed is one of the reasons to practice the virtues: because this helps you to become the sort of person who can have true friendships, and having such friendships is vital to human thriving.[\[23\]](#)

Another reason friendship harmonizes with the other virtues well is that the virtues are the ingredients of a good life. If friendship is (in part) wishing for the good for your friend, this will mean (in part) wishing that your friend may practice the virtues. If you do not have a good grasp of the virtues, you may wish "the good" for your friend but be mistaken about what that is, and so may harm your friend accidentally. Take for instance a "friend" who enables the other's addiction (I love my friend, so I'll make sure to buy enough meth for both of us).[\[24\]](#)

On the other hand, C.S. Lewis thought genuine friendship was just as possible between vicious people as between virtuous people, and that this makes it potentially dangerous. "It makes good men better and bad men worse," he thought, because similar people cluster in friendship and reinforce each others' characteristics.[\[12\]](#) It might even be the case that people with a lot of vices are especially eager to form friendships with others because they find other people's characters a welcome distraction from their own.[\[21\]](#)

That said, I've noticed that even among vicious people who form friendships together, they may be attracted to each other's virtues as much as or more than their vices. For example, among cruel friends whose friendship is partially based on indulging their cruelty together, it is still not necessarily the cruelest who will be most esteemed, but the one whose cruelty is boldest or wittiest.

While true friends in Aristotle's sense wish for each other's good, vice-friends in Lewis's sense may want the opposite, since wallowing in vices together is part of how those friendships work. If you trade in your drunkenness, sloth, and cynicism for sobriety, industriousness, and hope, the vice-friends you've been slacking off with may feel abandoned and judged, and they try to prevent your improvement (unless they are bright enough to be inspired by your example). This is another way friendship can work against your interests if you don't have a good foundation of other virtues or if you choose your friends poorly.

Aristotle drew a distinction between “philesis” (translated as “love” below), a perhaps fleeting feeling, and “philia” (“friendship”), a durable and characteristic disposition marked by habitual action:

[I]t seems that while love is a feeling, friendship is a habit or trained faculty. For inanimate things can equally well be the object of love, but the love of friends for one another implies purpose, and purpose proceeds from a habit or trained faculty. And in wishing well for their sakes to those they love, they are swayed not by feeling, but by habit. Again, in loving a friend they love what is good for themselves; for he who gains a good man for his friend gains something that is good for himself. Each then, loves what is good for himself, and what he gives in good wishes and pleasure is equal to what he gets; for love and equality, which are joined in the popular saying φιλότης ἴσοτης [“in amity, equality”], are found in the highest degree in the friendship of good men.[\[25\]](#)

This highlights another advantage of friendship among the virtuous: that vicious people tend to pursue their perceived self-interest in a zero-sum way that has deleterious effects on those around them (including their “friends”) whereas virtuous people tend to pursue their perceived self-interest in ways that also advantage those around them (such as their friends).

True friendship, said Aristotle, is a two-way street, but is characterized by loving rather than by being-loved.[\[26\]](#) This differentiates it from something like honor, where it seems better to receive than to give, or from flattery, which is a sort of false friendship that some people like to receive. I suppose the two-way-street aspect of it also distinguishes it from the sort of regard people may have for heroes and celebrities, in which they might value (even love?) such a person, but that person might not know they exist.

There are some practical advantages to having a friend: they’re someone you can count on in a pinch, for example; you can bounce ideas off of them; and they can give you the help you need even when it’s not the help you want. But commentators on the love of friendship are more likely to describe it not as something of practical value, but as an end in itself—as in the C.S. Lewis quote at the top of this section. You have friends because having friends is one of life’s delights, and you love your friends because that is how friendship works.

Erotic love

“Love is a snowmobile racing across the tundra and then suddenly it flips over, pinning you underneath. At night, the ice weasels come.” —Matt Groening

Popular culture mentions of “love” are dominated by the romantic, erotic variety.[\[27\]](#) Falling in love and being in love is an important part of life for many people; stories and [songs about this sort of love](#) are ubiquitous.

In popular Hollywood-style myth people will risk everything and will go on arduous quests in pursuit of “true love” of this sort. Part of what makes these myths work is that it is seen as plausible that “true love” is worth just about any sacrifice. A back-up device is to suggest that the feeling of falling-in-love is so utterly compelling in the delights it promises or the desire it provokes that other values must submit to it so long as it lasts.

Be that as it may, I'm not sure how to fit this sort of love into the scheme of virtues. Is the pursuit of love a habit that characterizes a thriving person? Or is it more of a remarkable event or a particularly overwhelming emotion or an enthusiasm? It is such a big deal (and can have such big consequences) that it seems likely there are better or worse ways to go about it, and maybe there are learnable skills involved. But you could argue that love itself isn't the virtue, but that a lover deploys other virtues more-or-less well in the course of their erotic quest.

If you were to habitually, characteristically, *fall* in love you'd be something of a laughable Casanova, so that can't be quite right. Indeed infatuation is associated with [fatuousness](#) and a variety of vices. You could imagine an anti-Paul composing an aphorism about this sort of love:

"Love is impatient, love is cruel. It is jealous and boastful. It slanders competitors and thinks only of itself. It is quick to fury and remembers every slight. It will stop at nothing and will clutch at straws. It always doubts, always pleads for reassurance, always despairs..."

Eros seems to be a *demanding* variety of love, both in the sense that it presents itself as an imperative demand to the lover, and in the sense that the person tapped by it cannot be satisfied by *loving* but demands to be *loved* in return. If you love in this way and are not loved back, you may pine or plead. Such love is more focused on *desire* (I want you, I need you) than is friendship or *agape*. It sometimes is described as a "hunger" in a way that other forms of love are not.

Cupid is notoriously volatile and unreliable. " 'I will be ever true' are almost the first words he utters,"[\[12\]](#) but he requires novelty, surprise, uncertainty, ambiguity, to fuel the flames of passion. As those things diminish, eros loses the fuel it needs to burn. To the extent that we demand that our romantic love affairs keep us perpetually twitterpated, we doom them to be short and to end in disappointment.

But maybe the falling-in-love part is not where the virtue lies anyway; maybe it's the more relationship-building being-in-love or staying-in-love instead. Erich Fromm contrasted the passive "passion" of falling in love from the active "action" of loving.[\[28\]](#) When you are under the influence of a passion, it's driving the car; when you take the wheel, you can take action, and it is then that you can exercise virtue.

There is a lot of advice out there about how to sustain and develop love once it gets past the twitterpation stage and into relationship-building. It's difficult for me to tell which of this advice is reliable. Most of it is presented confidently, but without much information about how that confidence was earned. I don't think I can come up with anything better, so I'll just link to *LessWrong's own Relationship Advice Repository*.

Erotic love as a template for better loves

I was struck by how a number of authors suggested that erotic love is a kind of hint or training meant to prepare us for more mature and important varieties of love.

Socrates, for example, suggested that when someone falls in love with some attractive person, they are really attracted to Beauty itself, but are being confusedly entranced by a frail worldly exemplar of it. But such a person can "use the beauties of the earth as steps along which he mounts upwards for the sake of that other beauty"

and eventually satisfy his desire correctly by fixing it on absolute Beauty, rather than on its disappointing mortal understudies.[\[29\]](#)

Erich Fromm thought that “the basis for our need to love” in general “lies in the experience of separateness and the resulting need to overcome the anxiety of separateness by the experience of union.”[\[28\]](#) Our erotic quests for union with some specific person, then, are stand-ins for a larger quest for reunion with the cosmos. If you get stuck at the erotic love stage, you will merely go from being an anxious separate individual to being one part of an anxious separate couple. You have to continue past erotic love to a deeper unity in order to really scratch the itch that drove you to fall in love.

C.S. Lewis thought *eros* was a sort of metaphor for *agape*: that its intensity and unanswerable insistence gives Christians an idea of how they ought to stoke the fires of their love for God, for their neighbor, and so forth:

[Eros's] total commitment is a paradigm or example, built into our natures, of the love we ought to exercise towards God and Man.... It is as if Christ said to us through Eros, “Thus—just like this—with this prodigality—not counting the cost—you are to love me and the least of my brethren.”...

Spontaneously and without effort we have fulfilled the law (towards one person) by loving our neighbour as ourselves. It is an image, a foretaste, of what we must become to all if Love Himself rules in us without a rival. It is even (well used) a preparation for that.[\[12\]](#)

If these authors are on to something, then maybe one of the ways the virtue of love operates vis-a-vis erotic love is to not take erotic love too seriously, not get hung up on it, but be willing to leave it behind as though it were a lower stage of your rocket ship.

Erotic love as union

Descriptions of this sort of love often depict it as ideally [culminating in a union](#). Informal phrases like “you complete me,” or “my other half” imply this, as does the legal institution of marriage, which unites a couple into a legal corporate entity that for example owns property jointly, files a single tax return, cannot be compelled to testify against itself (i.e. each-other), and so forth. Sexual intercourse can seem designed to imply this joining of two-into-one (so much so that even our [mechanical connectors and fasteners have genitalia](#)).

Romantic love may begin in a flurry of irresponsibility, but if it lasts long enough it seems to buzz around an attractor formed of things like [duty](#), [commitment](#), and [loyalty](#).

One way of looking at this is that when two people come together into the chrysalis of marriage, what emerges is a third person: a different kind of human creature, composed of the original two but greater than the sum of its parts. There is an echo of conception, pregnancy, childbirth in this, which is probably not coincidental given the strong historical connection between marriage and childrearing. Creating the third person of a marriage can be seen as the first step in establishing the institution of a family.

If romantic love begins with infatuation and desire, and then ripens into something that includes a *philia*-like love of the other for-their-sake, what then happens at the point of union? One interpretation is that the antithesis of desire for-my-sake and love for-their-sake synthesizes into a love for-our-sake. The lovers reorient their focus from each-other to us-both or to what-we-have-together. Their rational self interest from that point forward becomes a joint pursuit of the interest of this third-party composed of both.

Erich Fromm was concerned that this could reintroduce narcissism to the lovers who had to set narcissism aside in order to love. Once this new relationship with the beloved forms, it can become what he called “*egotism à deux*” in which the lovers care deeply for their shared union, but tune out everyone else.^[30] This can then reintroduce the original problem of the anxiety of separateness, though now instead of feeling that *I* am separate, cut off, standing outside, it’s *we* who are.

Sometimes it is implied that this united-couple third-person is what each of us is meant to become. In this telling, our solitary selves are only half-people, incomplete until we find our other half: Bachelors and spinsters are like caterpillars who failed to become butterflies, or like Peter Pans unwilling to venture forth from immaturity. In the *Symposium*, Aristophanes told this in the form of a myth in which people were originally whole-doubles, but then were split in half by an angry Zeus.^[29] ([You might prefer the *Hedwig and the Angry Inch* musical version.](#)) As a result we spend our lives trying to find our other halves, and cannot be satisfied until we do.

Critics of this way of looking at it point out the disadvantages of considering yourself to be half-a-person, hoping that someone else will complete you. For one thing, such an idea puts the focus of love on *finding that person*—who are they, where are they, how might I attract them? This can mislead people about how love works. Erich Fromm wrote that “People think that to love is simple, but that to find the right object to love—or to be loved by—is difficult.”^[31] He compared this to someone who wants to paint a beautiful portrait, but neglects to practice painting, waiting instead to find the perfect model. In the course of this argument, he described love in a way that brings it back into the category of the virtues:

Love is not primarily a relationship to a specific person: it is an *attitude*, an *orientation of character* which determines the relatedness of a person to the world as a whole, not towards one “object” of love.^[32]

Fromm’s “Art of Loving”

And what does Fromm think this virtue (“orientation of character”) consists of? Like Paul, like John Donne, like Iris Murdoch, he looks closely and finds that love consists of other virtues. In his case, he identifies the main ones as care, responsibility, respect, and knowledge. By “care” he means the active sense of [taking care of](#) (as opposed to the more passive caring-about). By “responsibility” he means something less like “duty” and more like “responsiveness”—feeling the other person is your concern. By [“respect”](#) he means that you love the other person for-their-sake, and that you do not try to dominate or possess them. And by “knowledge” he means in part a genuine and penetrating curiosity about the other person, in which you want to know them in the same spirit as the Oracle counseled [“know thyself.”](#)

Fromm calls this the “Art” of loving, and says that to master any art you need [discipline](#), [concentration](#), and [patience](#).^[33] So those might be thought of as additional

prerequisite virtues.

He thinks “the main condition for the achievement of love is the *overcoming* of one’s *narcissism*” and that the “opposite pole to narcissism is objectivity.”^[33] Narcissism interprets reality according to what-it-means-to-me; objectivity according to what-it-is. We’re all somewhere on the objective-to-narcissistic spectrum at any particular time; some of us tend more toward one end than the other. To become a loving person, we can’t spend too much of our time gazing lovingly at our face in the shallow end of that pool. “The facility to think objectively is reason; the emotional attitude behind reason is that of *humility*.”^[33] So: “love being dependent on the relative absence of narcissism, it requires the development of humility, objectivity, and reason. One’s whole life must be devoted to this aim.”^[33]

He further finds crucial something he calls “rational faith,” which he defines in a way that I interpret as being mostly optimism, hope, and trust with a soupçon of confidence, integrity, and industry. If that weren’t enough, he thinks courage is important, too.

Attachment theory

I noticed that the positive psychology book *Character Strengths and Virtues: A Handbook and Classification* quickly pivoted from talking about the virtue of love to discussing the research about “attachment.”^[34] I’ll give a thumbnail sketch of attachment theory here.^[35]

Attachment theory says that different children develop different ways of relating to novel situations and to other people, and that these persist into adulthood and shape the way we form intimate (most typically, romantic) relationships. Most adults practice one of three forms of attachment:

1. **anxiously attached** people put a lot of effort and attention into relationships, which are very important to them. They work hard to read cues, but tend to interpret them in overly-insecure ways. They often exhibit jealousy.
2. **securely attached** people are warm, loving, unconcerned about their relationships, trusting, and understanding. They tend to form long loving bonds that work fairly well.
3. **avoidant** people value autonomy and worry about losing their freedom to relationships. They tend to try to keep people at arm’s length. They may explain this as being “free spirits” who “aren’t ready to commit” or who are waiting for the right person.

Attachment theorists and researchers are sometimes careful to say that these styles are merely *different* rather than some being normal and others abnormal, but the field as a whole is usually not too subtle about pathologizing the avoidant and anxious styles in favor of the securely attached style. People in securely-attached relationships apparently tend to have longer lifespans and better outlooks on life, so maybe there’s something to this.

I don’t want to go too deep into this rabbit hole. There has been a lot written about attachment theory; it seems to resonate with a lot of people’s experiences and with the experiences of couples’ therapists. Its simple categories facilitate research and make some facets of love theoretically amenable to scientific investigation. There’s also a lot of pop-sci-type writing about it that seems barely distinguishable from

astrology. But if you see yourself mirrored in a description of the anxious or avoidant attachment style and you think this is interfering with your exercise of the virtue of love, attachment theorists may be able to explain how you got that way and how you can work with it.

The book [Attached](#)^[36] was recommended to me as a good introduction to this theory and how to use it.

Is romantic love unconditional love?

As I mentioned in my discussion of *agape*, people can have mixed feelings about unconditional love. On the one hand, you may yearn to be loved unconditionally—just for being you—not for any impermanent accidents *about* you like your attributes and actions. You don't want to have to continually earn love and be insecure about whether you remain worthy.

On the other hand, you would like confirmation that your attributes and actions are indeed lovable, that you are loved for the kind of person you are, not just out of habit. You don't want to be loved with no regard to your specific character: as though you were an interchangeable part that could be frictionlessly swapped out for another. Furthermore, if your love for your partner includes a strong for-their-sake component, you probably wouldn't want them to continue to love you if you were to become abusive, degrading, or otherwise morally repulsive.

[W]hile you seem to want it to be true that, were you to become a schmuck, your lover would continue to love you, ...you also want it to be the case that your lover would never love a schmuck.^[37]

When you love some person in particular, this in part means that you *value* them in particular. You single them out from the crowd as being special and of extra worth. There is [some debate](#) about how this works, fundamentally: Do you evaluate that person's characteristics, find those valuable, and therefore find the person lovable and begin to love them? Or do you find the person lovable and begin to love them, and as a result their characteristics become valuable ones to you?

Lovers themselves often appear confused about this. They may describe qualities of their beloved that to a less-bewitched observer seem utterly ordinary, but with stars in their eyes: "The way you wear your hat, the way you sip your tea..."^[38] Do they see something we can't see because we haven't looked closely enough? Or do they see something that isn't really there, but that they are projecting from their hearts?

There may be some feedback between love as appraisal of value and love as bestowal of value. You may come to love someone and then confirmation bias leads you to selectively sift for evidence that verifies they're wonderful. This evidence in turn enhances the love you feel for them and how exceptional they seem. And that then makes additional evidence of their marvelousness seem immediately credible.

Sometimes in popular stories of romance, the key to love is that person A somehow gets a glimpse at something deep in the core of person B and discovers there some intensely lovable hidden quality beyond what is superficially obvious. In such a case, A is *astute* or *observant* or maybe *trustworthy* enough to be granted access to this hidden quality. And this excuses why A can value B so exceptionally and in a way that seems to defy consensus. If we assume this theme of folklore represents real wisdom

about this variety of love, then having the ability to discern non-obvious founts of value in other people may be important in practicing it.

It's easy to adore someone if you can only see them through rose-colored lenses and heart-shaped pupils. It may be that "true love" doesn't come on the scene until you stop that nonsense. Anais Nin wrote: "Where the myth fails, human love begins. Then we love a human being, not our dream, a human being with flaws."^[39] But she also thought that the lover's myth can be their vision of the best potential of the beloved, part of the gift that the lover brings, and it can be as much the beloved's responsibility to accept this gift and live up to this vision as it is for the lover to stop hallucinating and come back down to earth.^[40] The lover can see the prince in the frog, even if the frog doesn't.

Love and vulnerability

"To love at all is to be vulnerable. Love anything, and your heart will certainly be wrung and possibly be broken. If you want to make sure of keeping it intact, you must give your heart to no one, not even to an animal. Wrap it carefully round with hobbies and little luxuries; avoid all entanglements; lock it up safe in the casket or coffin of your selfishness. But in that casket—safe, dark, motionless, airless—it will change. It will not be broken; it will become unbreakable, impenetrable, irredeemable. The alternative to tragedy, or at least to the risk of tragedy, is damnation. The only place outside Heaven where you can be perfectly safe from all the dangers and perturbations of love is Hell."—C.S. Lewis^[12]

Love, romantic love in particular, is associated with vulnerability. This is in at least three ways:

1. the process of building love requires some vulnerability in order to foster intimacy (for example, you share information about yourself that it is risky to allow someone else to know, or you fall asleep with your head in their lap)
2. when you love someone, you rejigger your identity and priorities to accommodate this new extraordinary value in your life; but your beloved like all mortal things is impermanent and changing, and so may be painfully wrested from you
3. you become affected by what befalls your beloved, so you can be hurt by what hurts them as much as by what hurts you directly; your surface area of vulnerability increases by that much

This suggests that the virtues of courage and moderation may be important. Courage so that you can feel the fear and love anyway. Moderation so that you stop short of an "I can't live without you" sort of utterly-dependent love.

Searching for a common denominator

So far I've seen a lot of indications that the varieties of "love" are composites of other things, including packages of virtues, but also including things like feelings and values. But I wonder if there is something extra in the glue that holds all of that together that belongs more exclusively to love itself. Maybe there's some X-factor such that you might say *agape* is X+devotion+altruism, *philia* is X+camaraderie+loyalty, *eros* is X+desire+intimacy, parental love is X+nurturance+care, or something along those lines.

Here are some possible characteristics of this X-factor:

1. Things that you love tend to rank highly as priorities relative to other similar things that you do not love.
2. Love is intensely-felt and extraordinary (even in the case of Christian *agape*, I think Christians aren't meant to make love mundane when they make it unexceptional).
3. Love is often a non-egocentric valuation of something. You don't love something because of what it brings to you, or in the sense that you would like to possess it or turn it to your own uses, but you value it for what it is. (Though you may also love someone or something appreciatively and gratefully, and may strongly value what a loved one brings to your life, so this is not cut-and-dried.)^[41]
4. And yet it does not pretend to be an *objective* valuation of something. Instead, what you love is loved in particular by you and you do not expect others to share this evaluation as a matter of course. Whether you love someone has a lot to do with accidents of relation, proximity, history, luck. Reason has a weaker role (maybe it can point out people who might be worthy of our love, or soberly determine when our love is misplaced).
5. Love can develop (or evaporate) accidentally and as an afterthought, rather than deliberately. You are more likely to "come to love" or to "realize you love" somebody than you are to "decide to love" them.
6. Love is distinguished by strong emotion. You can imagine being dispassionately kind, loyal, friendly, respectful, and so forth, but being dispassionately loving seems like a contradiction. This emotion is a positive one—you feel something like a warm joy for the things you love (not merely, for instance, strong jealousy, or strong anxiety for their safety).
7. Love for something extends in time. It typically doesn't come-and-go or show up briefly and vanish. It may not last ['til the end of time](#), but it lasts a while and puts up a fight before it goes away. If you say "I love you" on day one, and you say "oh, that was yesterday" on day two, it probably wasn't the real deal.
8. Love can change you^[42] and it forms part of your self image (you define yourself in part by what and who you love). It can entangle you with another person such that you change. If love fails or is disrupted, this can also mess up your identity; you may have to pick up the pieces and reconstruct your person on a new foundation.
9. Love necessarily has two poles: the lover and the thing that is loved. Sometimes people will complain that they are full of love but have nobody to bestow their love upon, or some formula along those lines, but that sort of pining loneliness is not itself love, but more of a hunger for the opportunity to practice love.
10. Love includes not just passive sentiments, but actions meant to promote what one loves.

I'm sure I'm leaving a lot out. And some of these things might not apply to all examples of love. But I think I'm at least in the ballpark. Now I want to try to imagine how to translate this X-factor into something like a virtue: a characteristic, habitual disposition. It might go something like this:

People with love idiosyncratically, subjectively value some people or things especially highly (in an [intrinsic rather than instrumental](#) way) and in a way that is accompanied by strong positive emotion. They are at least somewhat constant in this intense feeling, not merely momentarily enthusiastic. They recognize and respect the evidence of this love in themselves, even if they did not deliberately choose it, and even if they cannot plausibly defend it as an objective evaluation. They define themselves in part by what it is that they value in this extraordinary way. And they do

things (or will do them as the occasion arises) to nurture, defend, and enrich that which they love.

If this is to describe a virtue, this package of characteristic habits must also promote or be evidence of the human flourishing of the person who exhibits them. Love is such a commonly-valued (if often sketchily-defined) human trait, that it's tempting just to cite popular acclaim: no further explanation necessary.

But I think we can do better than that: Consider just the part of love that involves intensely and emotionally valuing certain things (much of the rest seems to follow more-or-less naturally from this). If you live in a world in which there are such subjectively valuable things, you are enriched relative to a person who lives in a world in which such things are absent. Catchphrases of the depressed like "there's nothing to live for" or "nothing really matters" are almost synonyms for "I can find nothing to love" in this sense.

Intimacy

The VIA Institute includes [love as one of its "Character Strengths"](#) (their preferred term for virtues). They describe it this way:

Love as a character strength, rather than as an emotion, refers to the degree to which you value close relationships with people, and contribute to that closeness in a warm and genuine way. ...love as a character strength really refers to the way you approach your closest and warmest relationships. Love is reciprocal, referring to both loving others and the willingness to accept love from others.

They put a stress on "closeness" that is absent from the definition I came up with. This is at least in part because I was trying to accommodate Christian *agape*, in which you are supposed to be able to love even those you are not close to (even those you'd like to keep at a distance, like your enemies).

The love of friendship, of family, and of romance, however, usually implies intimacy: mutual vulnerability and disclosure, unusually profound understanding, regular contact or proximity, and things of that nature. If intimacy failed to make the common-denominator cut, it didn't miss by much.

The Passionate Man

Simone de Beauvoir identified a subtle failure mode of love. To oversimplify, she believed that people necessarily must themselves create the ultimate values they live by, but that this power (and responsibility) frightens us and so we are perennially on the hunt for mythical objective values we can latch onto instead. One possible way to do this is to love someone or something and then to imagine that this subjective love has become fused into the love-object itself, giving it objective value. She calls the person who falls for this "the passionate man."^[43] The passionate man *almost* gets it right—he correctly sees himself as what creates values, but he incorrectly overshoots the mark: deifying those values and locking himself inside of them, excluding everyone else from a weird, private world of obsession.

The passionate man is unpleasantly monomaniacal and can be dangerous: "If the object of his passion concerns the world in general, this tyranny becomes fanaticism." The passionate man is in the final analysis less in love with what he values than with

his falsely-objectified valuation of it, and so this becomes another sort of narcissism after all.

Conclusion

When I began to research love I was vaguely aware of all of the terminological confusion and I was expecting to find that on close inspection love would either vanish into a vague applause-lit miasma or just decompose into a set of component sub-virtues.

I think now I have found something that belongs to love more particularly and that can fit (a little awkwardly) into the scheme of virtues. I think I understand why that crazy little thing called love is valuable to a flourishing life, for instance.

But I'm also more aware than before of just how much speculation about the nature of love there is out there and how difficult it is to find the gems in the pile. I don't feel at all confident that I've done justice to the topic here.

1. ^

cf. H. Jones "What Is Love?" (1983); J.M. Leslie "I Want to Know What Love Is" (1984); C. Porter "What Is This Thing Called Love?" (1929)

2. ^

John Donne, *To the Countess of Huntingdon*

3. ^

There has been a lot of hair-splitting about the various Greek terms for love in the Bible. (The books of the New Testament were written in Koine Greek; Jesus himself spoke Galilean Aramaic.) You will typically hear it said that *agape* is the Christian love Jesus was discussing, *philia* is a more casual fondness, *storge* is the sort of love you have for your children or relations, and *eros* is the sort of love romantic lovers have for one another. While I am envious of this supposedly more-particular Greek terminology, this may be a bit oversimplified. Sometimes in Biblical Greek, *agape* seems to be used as a sort of emphatic fondness, much in the same way that we might say "I love χ" in English to mean "I really really like χ." Also, the verb *philéō* had at that time taken on an additional meaning of "to kiss" ([Judas philéō Jesus to betray him in Luke 22:47](#), for example) and so sometimes writers would substitute *agapaō* to avoid that when it would be ambiguous or perhaps cause snickering. But other times the writers who quoted Jesus could have chosen to use *philia* to mean something more like fondness or affection, or *agape* to mean something more serious, and they pointedly use the latter. These Jesus quotes all use *agape*, for example.

To make the terminological difficulties just that much worse, Augustine insisted on Latinizing this variety of love as *caritas* which then became "charity" in English, though this has only a dim resemblance to how we usually use the word "charity" today. Because of this, depending on which Bible translation you prefer, the greatest Christian virtue may be [either "love" or "charity"](#).

4. ^

Jerome, *Epistles* “To Theophilus Bishop of Alexandria” (399): “*Cunctarum virtutum mater est caritas.*”

5. ^

See also [Ephesians 3:17—“being rooted and grounded in love”](#)

6. ^

See for example Thomas Aquinas *Summa Theologica* (1485) II.2 question #23, article #7: “Can any genuine virtue exist without charity?”

7. ^

See for example Thomas Aquinas *Summa Theologica* (1485) II.2 question #23, article #8: “Is charity the form of the virtues?”

8. ^

Iris Murdoch, “On ‘God’ and ‘Good’ ” (1969) in *Existentialists and Mystics*, pp. 337-362.

9. ^

Iris Murdoch, “On ‘God’ and ‘Good’ ” (1969) in *Existentialists and Mystics*, pp. 337-362. In “The Sovereignty of Good over Other Concepts” (1967) she suggests that there is one more step past love in this unification, which leads you to a Platonic “Good.” Love “is the energy and passion of the soul in its search for Good, the force that joins us to Good and joins us to the world through Good.”

10. ^

Thomas Jay Oord’s *Defining Love* (2010) gives several examples of this and concludes that this “illustrates well Gene Outka’s observation that ‘the meaning ascribed in the literature to love, in general, and to *agape*, in particular, is often characterized by both variance and ambiguity.’ [Agape: An Ethical Analysis, 1972] Robert Adams notes the diverse understandings of *agape* that scholars offer, and he concludes that ‘*agape* is a blank canvas on which one can paint whatever ideal of Christian love one favors.’ [Finite and Infinite Goods: A Framework for Ethics, 1999] I suggest that this variance arises from the theological, ethical, anthropological, scientific, and metaphysical commitments of those who use *agape* to identify something distinctive when compared with other forms of love. These diverse understandings suggest that we should not think that *agape* has a uniform or obvious meaning.”

11. ^

Jesus highlighted the extremely altruistic nature of his *agape* in [John 15:12-13](#): “This is My commandment, that you love one another as I have loved you. Greater love has no one than this, than to lay down one’s life for his friends.”

12. ^

C.S. Lewis *The Four Loves* (1960)

13. ^

Erich Fromm *The Art of Loving* (1956), chapter I

14. ^

Though the former, Lt. Col. Rev. George Zabelka, [later had a change of heart](#).

15. ^

See, e.g. [1 John 4:7-21](#)—“Dear friends, let us love one another, for love comes from God. Everyone who loves has been born of God and knows God. Whoever does not love does not know God, because God is love.... This is love: not that we loved God, but that he loved us and sent his Son as an atoning sacrifice for our sins. Dear friends, since God so loved us, we also ought to love one another. No one has ever seen God; but if we love one another, God lives in us and his love is made complete in us. This is how we know that we live in him and he in us: He has given us of his Spirit.... God is love. Whoever lives in love lives in God, and God in them....”

16. ^

“Enter through the narrow gate. For wide is the gate and broad is the road that leads to destruction, and many enter through it. But small is the gate and narrow the road that leads to life, and only a few find it.” (Matthew 7:13-14)

17. ^

In this case, Latin is in our corner: *amicitia*/friendship is more-closely related to *amor*/love.

18. ^

Aristotle also recognizes lesser forms of friendship between people who simply enjoy each other's company or who are cooperating to mutual advantage.

19. ^

Aristotle, *Nichomachean Ethics* IX.10

20. ^

See also Cicero, *de Amicitia*: “For everyone loves himself, not with a view of acquiring some profit himself from his self-love, but because he is dear to himself on his own account; and unless this same feeling were transferred to friendship, the real friend would never be found; for he is, as it were, another self.”

21. ^

Aristotle, *Nicomachean Ethics* IX.4

22. ^

Aristotle, *Nicomachean Ethics* VIII.4

23. ^

Aristotle, *Nicomachean Ethics* VIII.1

See also Seneca, in one of his letters to Lucilius: “I want your friendship, and it cannot fall to my lot unless you proceed, as you have begun, with the task of developing yourself. For now, although you love me, you are not yet my friend.... A friend loves you, of course; but one who loves you is not in every case your friend. Friendship, accordingly, is always helpful, but love sometimes even does harm. Try to perfect yourself, if for no other reason, in order that you may learn how to love.”

And Cicero, in *de Amicitia*: “I am inclined to think that, with the exception of wisdom, no better thing [than friendship] has been given to man by the immortal gods... [T]here are those who place the ‘chief good’ in virtue and that is really a noble view; but this very virtue is the parent and preserver of friendship and without virtue friendship cannot exist at all.”

24. ^

This is also how “selfishness” can be a virtue in the virtuous and a vice in the vicious: for example, a virtuous person may selfishly want even more temperance, courage, and honesty than they currently have; a vicious person may selfishly want even more self-indulgence, flattery, and sloth than they currently have.

25. ^

Aristotle, *Nichomachean Ethics* VIII.5

26. ^

Aristotle, *Nichomachean Ethics* VIII.8

27. ^

“Erotic” is now often used as a synonym or euphemism for “pornographic” or “sexy,” but I mean it in a broader sense. C.S. Lewis (*The Four Loves*, 1960) usefully differentiated Eros from Venus: Eros is the “in love” love; Venus is the “love you all night long” kind.

28. ^

Erich Fromm *The Art of Loving* (1956), chapter II

29. ^

Plato, *Symposium*

30. ^

Erich Fromm *The Art of Loving* (1956), chapter III

31. ^

Erich Fromm *The Art of Loving* (1956), chapter I.

I also like the way C.S. Lewis put this when he wrote about "...the odiousness of nearly all those treacly tunes and saccharine poems in which popular art expresses Affection. They are odious because of their falsity. They represent as a ready-made recipe for bliss (and even for goodness) what is in fact only an opportunity. There is no hint that we shall have to do anything: only let Affection pour over us like a warm shower-bath and all, it is implied, will be well." —*The Four Loves* (1960)

32. ^

Erich Fromm *The Art of Loving* (1956), chapter II §3

33. ^

Erich Fromm *The Art of Loving* (1956), chapter IV

34. ^

Christopher Peterson & Martin E.P. Seligman, *Character Strengths and Virtues* (2004), chapter 13

35. ^

Based on my notes from talks led by Dr. Erin Elfant at "The Order of the Oryx", 22 February and 1 March, 2017

36. ^

Amar Levine & Rachel S.F. Heller, *Attached* (2010)

37. ^

N. Delaney, "Romantic Love and Loving Commitment: Articulating a Modern Ideal", *American Philosophical Quarterly* 33 (1996): p. 347

38. ^

Ira Gershwin "They Can't Take That Away from Me" (1937)

39. ^

The Diary of Anaïs Nin, 1939–1944 November 1941

40. ^

The Diary of Anaïs Nin, 1966–1974 letter (p. 152)

41. ^

C.S. Lewis distinguished "need-love" (e.g. the love of a child for her caregiver, also erotic desire), "gift-love" (gratuitous, for-their-sake, e.g. the love of a parent for a child), and "appreciative love" (love of what we judge as good); these are

not necessarily completely distinct, but can alternate or be combined in the same relationship. (*The Four Loves*, 1960)

42. ^

"the way a parent can change a baby—awkwardly, and often with a great deal of mess." —Lemony Snicket, *Horseradish: Bitter Truths You Can't Avoid* (2007)

43. ^

Simone de Beauvoir *The Ethics of Ambiguity* (1947), chapter II

Notes on Resolve

This post examines **resolve** and **decisiveness** (and things in that bailiwick like determination, commitment, dedication, steadfastness, and firmness), as part of a sequence of posts about virtues. I mostly explore what other people have learned about these virtues, rather than sharing my own research or opinions about them, though I've been selective about what I found interesting or credible. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will help people who want to know more about these virtues and how to practice them skillfully.

What are these virtues?

In short, a person with these virtues makes conscious, deliberate commitments when it would be advantageous to do so, such that these commitments reliably govern their future actions.

These virtues concern a crucial pivot point in the process of deliberately taking a course of action. For the purposes of this discussion, I divide this process into the following stages:

1. Prepare for making decisions:

What sorts of decisions do you want to make? What do you want to do with your life? Do you have a “calling” or a sense of purpose? Answers to broad questions like those allow you to weigh your possible decisions against your big-picture goals and your ethos. You need [wisdom](#) to understand what sorts of decisions you want to make, [temperance](#) so that your desires prompt you to make decisions that align with your wisdom, [know-how](#) so that you have the skills you need to act on your decisions, [ambition](#) so that you have high-impact decisions to choose from, devotion to help motivate your follow-through, and [courage](#) so that you won’t flinch from hard decisions.

2. Discern that a decision needs to be made:

You need [attention](#) to notice that the opportunity for a decision has arisen, while virtues like [duty](#), [social responsibility](#), [honor](#), [care](#), valor, helpfulness, and concern help you understand which decisions are your responsibility.

3. Weigh the cases for different options:

Virtues like [phrónēsis](#) (practical wisdom), [rationality](#), curiosity, and imagination help you discover and consider different possible courses of action and to anticipate their outcomes.

4. Make a decision:

The point where you stop weighing the case and finally commit to a course of action is where virtues like decisiveness, resolve, and so forth, come into play. (In addition, [optimism](#), confidence, [earnestness](#), and enthusiasm may play a role.) If you have the virtue of resolve, you make decisions in an effective way: your decisions are really decisive, in that they indeed launch your earnest action.

5. Follow-through:

You then need [self-control](#) (discipline, continence), [industry](#), [integrity](#), consistency, reliability, and conscientiousness to keep on task. If your decision involves multiple stakeholders, you will also find cooperation, teamwork, and leadership helpful.

6. Overcome obstacles:

When you encounter setbacks or difficulties, virtues like [endurance](#), [perseverance](#), [patience](#), and [forbearance](#) help you push through anyway, while virtues like carefulness, creativity, flexibility, prioritization, and ingenuity help you find solutions and workarounds.

7. Accomplish your goal (or not):

If you accomplish your goal, you can take pride in this. If not, there are useful techniques of resignation/acceptance and graciousness that help you cut your losses. If

a failure was regrettable, a productive process of [shame](#) might help you to course-correct.

This post concerns the fourth of these stages: that crucial pivot point between indecision and determined action. Some discussions of resolve and decisiveness incorporate elements of each of the first six stages, particularly stages three through five. I won't be able to avoid this entirely either—particularly because the confidence you need to commit to a decision depends so much on the quality of the process by which you arrived at the decision. But also, determining the right choice to make, and making that choice, is not enough if your deliberate choices do not actually guide your behavior. Such choices have a fools' gold shine to them. I discuss that in some depth in my [Notes on Self Control](#), however, so I will try not to gnaw on that bone too much here.

Another model to consider

For what it's worth, Pete Davis, in his investigation of the similar virtue of "dedication," identified a somewhat different set of what he called "dedicatory virtues" and described them this way:

Dedication... requires *imagination*—the ability to envision what isn't there just yet. It requires *synthesis*—the ability to make connections. *Focus* (the ability to concentrate) and *doggedness* (so you can return to the same task again and again, even if there's nothing new about it) are key. So is *passion*—the enthusiasm required to sustain engagement. And there can be no passion without *reverence*—the ability to be awed by something. Above all, dedication requires *commitment*—the ability to stick with something, despite there being other available options.^[1]

Associated vices

The vice of deficiency associated with these virtues goes by names like "indecisiveness," "dithering," "waffling," or "hedging." Sometimes people avoid being decisive by trying to add plausible deniability to their decisions: being "ironic" or "half-hearted," or perhaps they have "divided loyalties" or are "trying to have it both ways."

Decisiveness is in tension with the virtues of flexibility and spontaneity. It is not always clear when it is better to be resolute and when it is better to keep your options open. When the time for deciding arrives, it does not necessarily announce itself with an unmistakable alarm. A frequent complaint these days is what Pete Davis calls "Infinite Browsing Mode"^[2]—epitomized by sitting down to watch a movie but instead spending the next hour scrolling through the options on Netflix unable to make up your mind.

["The Copenhagen Interpretation of Ethics"](#) is a peculiar failure mode of decision-making. If you fall prey to it, you may subject choices that involve action on your part to intense scrutiny, while allowing choices that involve *inaction* on your part to happen unscrutinized. This is perhaps another form of plausibly-deniable indecisiveness. Someone who labors under this delusion may mistake deciding-to-do-nothing for not-deciding-to-do-anything. If they allow themselves to acknowledge that they have made a decision, it is as though they have tasted the forbidden fruit: they also feel the unwelcome weight of responsibility for the results of their decision. "This is why people flip a coin to make decisions—so they can blame the coin instead of themselves."^[3]

Hannah Arendt was of the opinion that a sort of indecisiveness, rather than a decisive malevolence, is at the root of most evil: "The sad truth of the matter is that most evil is done by people who never made up their minds to be either bad or good."^[4] One way people sometimes try to disengage from decision-making and responsibility is by relinquishing their authority over themselves to some other authority: they make themselves agents of

someone else's decisions—they "obey". "Superfluousness represents a temptation: it holds the promise of an existence devoid of (enacted) human agency, hence free of the burdens of responsibility and guilt, as well as hurt and loss."^[5] This temptation rests on a lie: that you can foist off your burden of deciding and of bearing the responsibility for your decision.

The vices of excess associated with this virtue include those that involve deciding before wisely weighing the options (making "snap judgements" or being "impetuous," "rash," or "arbitrary"), and those that involve cementing your resolution too firmly, such that you are unable to deviate from it as new evidence comes in or better options open up (being "obstinate" or "inflexible").

How does a decision happen?

Not everything you do is something you decide to do. For example: some things you do involuntarily, other things you do habitually as if by default. In many cases, though, it at least seems as though you deliberate, decide, and then act based on your decision. But it can be difficult to pinpoint the moment of decision, or what characterizes the decisive point, or the mechanism by which a decision to do something transforms into the action of actually doing it.

You may remember a time when you became aware of having decided, and maybe you even mentally vocalized to yourself "now I shall..." But was that the moment of decision, or was that more of a herald formally announcing that a decision had earlier been arrived at?

Philosophers and neuroscientists have attempted to trace decision-making back to its origins.

Aristotle, of course

Aristotle believed that decision takes place where intellect/reason and desire/appetite meet. ^[6] Desire says "I want this outcome" and reason says "this is how I might act to make such an outcome come about," and when the two combine, a decision is born of the union. For the decision to be a good one, the reasoning must be sound and the desire wise.

It takes desire to pick a goal in the first place, and then intellect to pick a course of action that brings that goal closer to realization. Neither reason nor desire is capable of initiating action by itself. If you trace any deliberate human action back to the spring it came from, you reach a starting point where reason and desire combine, and you can trace the path no further.

Does reason yoke desire and start to plow, or does desire enlist reason to seek its ends? To Aristotle, it's all a matter of how you look at it: neither is more accurate—does the sperm penetrate the egg, or does the egg absorb the sperm? There's no fact of the matter; it's just how you choose to describe it.

In Aristotle's view, in order to decide well we need to train our reason so we know *how* to behave well, and we need to mold our characters so that we *want* to behave well, and that's the end of the story. If we do not actually desire what is good, there is no third-thing inside of us (such as "conscience") that can step in and contradict the other two.

Western philosophers who followed Aristotle but in the Christian tradition were more likely to distrust the ability of people to mold their own characters in this way. In their view, our desires and our reason might conspire to suggest a course of action, but both our desires and our reason are too corrupt to be trusted. Fortunately there is a trustable third-party—such as Divine Law or "natural law", or the "still small voice of God" or "conscience"—that can overrule this conspiracy with a superior decision.

Neuroscience

Experimental evidence

Some intriguing experiments have measured brain activity while a decision is being contemplated. In a typical example,^[7] subjects were asked to—"when they felt the urge to do so"—make an arbitrary decision to press either a left-hand-operated button or right-hand-operated button and then to do so "immediately". While contemplating this, they were shown a series of letters on a screen. After they pressed a button, they were then asked which letter was on the screen when they decided which button to push.

The subjects' brain activity was measured while they were doing this, and examined for signals that would predict which choice they would make. Researchers found that they could confidently predict the subjects' choices based on measurable brain activity that occurred several seconds before the subjects reported they were aware that they had made a decision.

One possible implication of studies like this is that the part of decision-making that we are conscious of may be something of an afterthought, at least in these sorts of strange, artificial, arbitrary decisions.^[8]

Evidence revealed by pathologies

Another way to investigate the brain's role in decision-making is to examine brain pathologies associated with deficits in decisiveness.

Neuroscientist Antonio Damasio described one example.^[9] A patient had a tumor removed from his brain in a surgery that necessarily damaged some of his ventromedial prefrontal cortex (VMPFC). The surgery left the patient intellectually unimpaired, but pathologically unable to make decisions—"the structures destroyed... happened to be those necessary for reasoning to culminate in decision making." He described the patient deliberating about two possible next-appointment dates:

For the better part of a half hour, the patient enumerated reasons for and against each of the two dates... [culminating in] a tiresome cost-benefit analysis, an endless outlining and fruitless comparison of options and possible consequences. It took enormous discipline to listen to all of this without pounding on the table and telling him to stop...

It is as though the patient were again and again trying to prompt the part of his brain that usually responds to weighing options with an eventual "yes, that's the one!", but without success. Intriguingly, the patient had also become emotionally flat: "I never saw a tinge of emotion in my many hours of conversation with him: no sadness, no impatience, no frustration." Damasio says this suggests that emotion is crucial to decision-making (perhaps the conclusion "yes, that's the one" is an emotional one rather than a rational one).

The role of emotion in decision

This brings up a disagreement over the role of emotion in decision-making. While Aristotle was content to describe decision-making as an equal partnership between reason and desire, the Western philosophical tradition that followed has tended to emphasize the importance of conscious, rational deliberation in *governing* or *overruling* decisions that would otherwise more-unwisely be made by subconscious, emotional drives.

More recently, Chip Heath & Dan Heath, in their book *Decisive: How to Make Better Choices in Life and Work*, identified "short-term emotion [which] will often tempt you" as the main

obstacle to good decision-making.^[10]

When people share the worst decisions they've made in life, they are often recalling choices made in the grip of visceral emotion: anger, lust, anxiety, greed. Our lives would be very different if we had a dozen "undo" buttons to use in the aftermath of these choices.^[11]

For this reason, the Heaths recommend that you get some distance from whatever context of the decision may be causing emotional responses, and that you use a variety of mind-hacks to reduce the effect of emotion on your decision-making. An example of these mind hacks is one called the "10/10/10 Rule." You consider a potential decision and then spend some time trying to imagine how you will feel—having made that decision—in ten minutes, ten months, and ten years. This gives you some distance from your immediate emotional reaction to *making* the decision by instead considering how you expect to feel about *having made* the decision.

But, on the other hand, and in part inspired by neurological insights from Damasio and others, Jonah Lehrer (*How We Decide*) thinks that it is a mistake to overrely on our rational minds to make decisions.^[12] Our conscious, rational brains have access to a certain subset of information, which we can use to make our rational evaluations. But our decision-making brains have access to a richer subset of information, including what our rational brains see but also including richer input from our emotions and from other subconscious processes than those we are conscious of and can easily make part of our rational evaluations. Our decision-making brain has been evolving longer and more painstakingly than our conscious/rational brain (which has been slapped together more recently), so maybe we should trust our "gut response" a bit less skeptically.^[13]

Particularly when snap-decisions are called for, you can't rely on a cumbersome process for coming up with an ideal solution; you have to rely on intuition. "But... intuition is only accurate in domains where it has been carefully trained. To train intuition requires a predictable environment where you get lots of repetition and quick feedback on your choices."^[14] Some domains aren't good at generating good intuition (because the feedback is ambiguous, difficult to get at, or distant from your decision; or because you can't reliably practice under real-world conditions).^[15] This suggests that you may be able to improve your decision-making gut-feeling in a particular domain by practicing making decisions in that domain, and by doing what you can to make the decision-consequences feedback loop in that domain more salient and immediate.

What good is decisiveness?

Deciding is, from one point of view, simply unavoidable. Your journey through life is one set of crossroads after another. Just sitting down and not choosing among the possible paths is either not an option or, if it is an option, is one that you must choose.

But *deliberate* decision can take extra effort. We often pick our path at the crossroads not by explicitly deciding on it, but by letting ourselves be pushed in one direction or another, or by letting our momentum carry us further along in our present path. What do we gain from exerting this extra effort?

If Aristotle was correct, all virtues are habits of *choosing* in particular ways.^[16] A brave person is one who makes courageous choices, an honest person chooses to tell the truth, and so forth. So the willingness to choose is broadly important to being a flourishing human being. But there are also some more specific reasons why decisiveness is helpful:

Decisiveness enables inter-temporal intra-personal economic exchange

When you decide to do something, you are in effect restricting your future self's options to your present self's commitment. What do you gain from selling off your future freedom in this way rather than leaving your options open?

One way of looking at this is that decisiveness allows you to import value from the future into the present. If you commit today that you *will* do X, you can take a promissory note for X to the bank today and begin earning interest on it. If you have a reputation for reliably turning your resolutions into reality, you can exchange today's promises for things of more tangible value.

Resolve can also be a way of exporting value from the present to the future. If I am to have the value of knowing how to play *The Battle Hymn of the Republic* on the harmonica by next week, I need to resolve to begin practicing at some point before that. I trade the present value of my time and effort for the future value of my skill.

This gets more complicated when you make resolutions that may have the effect of changing the way your future self evaluates things.^[17] For example, if you decide to become a parent, you are basing your decision on values you have as a childless person and on your educated guesses about what you will find valuable as a parent. But you have good reason to expect that your values and your criteria for evaluation may change in unanticipated ways over the course of parenthood. You're aiming at a moving target.

Decisiveness brings an additional class of rewards within reach

“[W]e thought we were looking for smart people, but it turned out that intelligence was not as important as we expected. If you imagine someone with 100% determination and 100% intelligence, you can discard a lot of intelligence before they stop succeeding. But if you start discarding determination, you very quickly get an ineffectual and perpetual grad student.” —Paul Graham ([explaining how he decides which startups to fund](#))

Commitment is sometimes described as restricting your choices, but it also is a way of making certain sorts of choices possible. It trades a vast landscape of superficial potentialities for “the novelty of depth.”^[18]

Some rewards demand effort over a long period, where that effort is mostly wasted unless you sustain it to the finish line. So to begin you need to be confident that you have the ability to follow-through. Resolve or decisiveness is this confidence-in-embarking.

Pete Davis described our choice-filled modernity with the metaphor of a hallway lined with many doors. The breadth of options is so intoxicatingly delightful that you can forget that none of the many options is really yours unless you take it. The “Culture of Open Options” encourages you to stay in the hallway where all of these possibilities remain in reach, while those in a “Counterculture of Commitment” instead pick a door and leave the hallway for what lies beyond.^[19]

Decisiveness clarifies and makes future decisions easier

When you commit to something, you can clear your desk of the options you discarded and whatever considerations accompanied them. This allows you to spend more of your effort accomplishing goals and less weighing options.^[18] Your future decisions may also become clearer, as, while you won't always know for sure how to reach your goal, at least you know what that goal is.

The proliferation of choices (such as lifestyle choices and consumer choices) makes decisiveness both more crucial and more difficult. Increased choice can lead to more anxiety about decisions, as the grass is potentially greener in so many different locations now. A theorized phenomenon of [choice overload](#) (a.k.a. "the paradox of choice") makes it more difficult for people to decide well the more options they have to choose from. It may be that our historically extraordinary period demands an extraordinary aptitude for decisiveness.

What bad is decisiveness?

"A foolish consistency is the hobgoblin of little minds... With consistency a great soul has nothing to do. He may as well concern himself with his shadow on the wall." —Ralph Waldo Emerson^[20]

Decisiveness can go awry when it is done "not wisely but too well." It appears to be one of those Aristotelian virtues which are best practiced at a golden mean: when you are decisive at the right times, in the right ways, and to the right extent—rather than always, absolutely, and as much as possible.

Decisiveness cuts off other options

"Marry, and you will regret it; don't marry, you will also regret it; marry or don't marry, you will regret it either way. Laugh at the world's foolishness, you will regret it; weep over it, you will regret that too; laugh at the world's foolishness or weep over it, you will regret both. Believe a woman, you will regret it; believe her not, you will also regret it... Hang yourself, you will regret it; do not hang yourself, and you will regret that too; hang yourself or don't hang yourself, you'll regret it either way; whether you hang yourself or do not hang yourself, you will regret both. This, gentlemen, is the essence of all philosophy." —Søren Kierkegaard^[21]

When you resolve on some course of action, you necessarily abandon any mutually-exclusive alternatives. This can lead to fear-of-missing-out, or to regret if you later suspect you could have chosen better.^[22] On the other hand, if you refuse to ever commit, you risk missing out on things like depth, expertise, and mastery that are only available to the committed.^[23] So maybe it's a wash.

It can be hard to commit to some particular cause (e.g. AI safety, climate change, political reform) because so many other causes also seem crucial that it can seem negligent to fail to give them as much attention. Something similar leads to the altruism anti-pattern of donating modest amounts to a diverse bunch of well-meaning charities that cover the bases of what you believe are worthy concerns rather than a large amount to one particularly potent charity that you believe will make the marginally most effective use of your money.

[Taking a stand or making a judgement](#) is a type of decision. A permissive who-am-I-to-judge attitude can hide a sort of fear of commitment (to a particular moral stance): timid indifference masquerading as tolerance. On the other hand, a rush to judgement or a prudish eagerness to cast aspersions can result from being too eager to resolve life's gray areas into decisive black-and-white.

If you commit to a group or an institution or a cause or a leader/guru, you risk not only the alternatives that you have foregone, but also the possibility that what you have committed to will transform in unanticipated and unsavory ways. When something you have committed to disappoints you or turns on you, you confront questions of [loyalty](#): do you abandon your commitment or redouble your effort in order to save what you have committed to from itself? [\[24\]](#)

Decisiveness as posturing

Resolve can be a pose. It is especially notorious as a sort of macho posturing.

A leader's decisiveness can be alluring and can tempt followers to try to piggy-back on their decisions—believing that by doing so they can reap the benefits of decisiveness while foisting the responsibility for having decided off on the leader. But all too often this means that the followers are not dedicating themselves to a wise course of action, but to the vainglory of someone who has learned a parlor trick of social engineering, or someone who has decided not wisely but impetuously, leading with the chin.

Some people make a rash gesture of decisiveness as a way to bolster a weak or unconfident personal identity. By committing to a cause, taking an oath, or vowing fidelity to some authority figure, such a person can lash their leaky raft to what seems like a more seaworthy vessel. (Before, I wasn't sure who I was, but now that I've signed on the dotted line I know I'm a man with a mission.) But this rash discrete show of resolve can hide an attempt to have future decisions made-for-you by the cause, the institution, the authority figure. The implicit promise is that if you make one big decision now, the rest of your decisions will be made for you from here on out, and you'll be off the hook. [\[25\]](#)

Decisiveness as stubbornness in disguise

"If you set out to teach someone *how to not turn little mistakes into big mistakes*, it's nearly the same art whether in hedge funds or romance, and one of the keys is this: Be ready to admit you lost." —Eliezer Yudkowsky [\[26\]](#)

Some people who seek a reputation for decisiveness are just ashamed to admit they were wrong or that the time has come to cut their losses. This can result in a mulish stubbornness, in which you stick with your resolutions even when the underlying facts change or when it turns out you chose poorly. A desire to appear decisive can lead you to stick with bad decisions when a wiser person would have been more flexible.

Are you being virtuously resolute or merely obstinate? They can look awfully similar, and if there are any hard-and-fast rules that can reliably distinguish between them, I'm not aware of them. People who stick with something long after most anyone else would have given up—some are fools, others are visionaries.

In his discussion of the virtue of self-control, Aristotle tried to come up with a way of disentangling stubbornness from continence. [\[27\]](#) He emphasized that a good choice involves reasonably picking effective means to wise ends. If you later discover that you were mistaken or that the situation has changed such that your chosen means are no longer apt to lead to wise ends, it is no longer a virtue to stick "resolutely" to those means.

Decisiveness can cement a new and unwise status quo bias

We can be slow to reassess a decision, as its quality gradually degrades over time, when we are convinced that decision was a wise one when it was made. Providers of consumer products and services can capitalize on this effect. They may introduce a product or service to market in a high quality form at a bargain price, wait for people to internalize that the choice of that product or service is a wise choice, and then slowly ratchet down its quality or economy.

We may be tempted to over-fit choices to those that require a decisive resolution

How often is it *really* true that you have to make a grand commitment rather than just making the next small sensible decision?

You may feel pressure to declare yourself resolutely in some way when in fact there is little or no advantage to doing so. Such a thing can punish you twice: by putting you firmly on a course of action that you would be wiser to approach tentatively, and by eroding the effectiveness of your resolve if you later realize you would be better off backing out of your resolution.

Sticking with resolutions you've outgrown

You may have made commitments in your youth that no longer serve much purpose in your life. But having staked your character on your decision—having become “the sort of person who” has that commitment—you find it hard to surrender what is now a white elephant of a resolution. Commitments that lock you in to being a certain sort of person, with a certain set of values and interests, over the potentially long span of your life, can unwisely fail to account for the way people and their contexts change. Maybe put off getting that favorite-band or clever-meme tattoo for a few more years.

Trying to resolve the world instead of yourself

Resolution or determination is only the pivot point; after the resolution or determination comes the work of carrying out whatever you have resolved to. But sometimes people are tempted to view the pivot point magically as though it had much more power than that.



When I was a boy, my father went through Erhard Seminars Training (EST), an ancestor of today's [Landmark Forum](#). He joined an EST-related group called "The Hunger Project" which was to end world hunger by 1990. Werner Erhard was of the opinion that the key to ending world hunger was to make the end of world hunger "an idea whose time has come."^[28] So this group did not focus on providing food to the hungry or on improving food-providing networks, but instead on convincing people "to take personal responsibility for making the end of starvation an idea whose time has come." The project had a relentless focus not on practical steps that would alleviate world hunger, but on a "common stand, a commitment, a declaration... a worldwide, grass-roots commitment... [a] growing expression of commitment... [of the] boldly determined... expressing their commitment... because they themselves are committed... and [have] the commitment to realize their vision."^[29] Needless to say, world hunger did not end by 1990, and no amount of "commitment" could wish that away.^[30]

If you have the virtue of resolve, you can resolve that *you* will do something and be confident that you will in fact do it. But what you cannot do is to resolve that *someone else or something else* will turn out a certain way. That may be largely up to forces beyond your control, and to pretend otherwise is to leave the realm of resolution for the realm of magic.

How to enhance your resolve

The two major components to resolution are 1) to explicitly choose one option of those available, and 2) for your choice to be a potent one that actually guides your actions. A variety of methods may help you improve in one or both of these components.

Balance the explorer and exploiter mindsets

Curiosity, imagination, and flexibility are virtues of the “explorer” who is seeking out the best options; resolve, persistence, and tenacity are virtues of the “exploiter” who is trying to get the most out of an option they have chosen. A skillful balance of both modes will get you the most bang for your buck. Some people seem to overweight the explorer mode, others the exploiter mode. If you already tend to overweight the exploiter mode, it may be that advice on how to practice more determination or resolve is the opposite of what you need.

Make decisions well

Perhaps it goes without saying, but you will be more confident when you decide, and more apt to follow-through on your decision, if you have made your decision well.

This does not necessarily mean that you know *for sure* that your decision is the best one. You may not be able to know that. The future is full of surprises, and we often make decisions in unavoidable ignorance and uncertainty. Sometimes we reassess our goals and values in a way that calls into question our past commitments. And it's often impractical to thoroughly vet all of our possible options.

However, if you know you used a [rational, prudent](#) process in the face of these constraints, you can at least be confident that you are making the best decision you know how to make. [\[31\]](#)

It is important that you be able to decide even though you can't know all the ramifications of your choice. If you wait until you can see the future perfectly you'll never stop waiting. “Grand commitments do not need grand blueprints.”[\[13\]](#) Sometimes you need to commit just to begin drawing up the blueprints.

The “non resolve” technique

Among the obstacles to decisiveness are “fear of regret,” “fear of association,” and “fear of missing out.”[\[19\]](#) If you wed yourself to one course of action, you may mourn the loss of your other options, or have buyer's remorse over the one you chose.

Linda Linsefors of the Center for Applied Rationality suggested a workaround for this that she calls “Non Resolve.”[\[32\]](#) It has two main elements:

1. When you are deciding, carefully consider the arguments for and against your decision. That way, if you are later tempted to deviate from your decision, you can't use “newly discovered” counterarguments that were there all along as an excuse.
2. Give yourself explicit permission to later consider any *genuinely* new arguments against your decision, and to abandon your decision if (and only if) those new arguments are substantial enough to mean the decision is no longer a good one.

Ironically, in order to commit, it can help to know that you have the option to quit or change your mind.[\[13\]](#) Otherwise you risk being paralyzed by the fear that your choice will lock you in.

In a similar spirit, you may be *conditionally* decisive. For example, you might say “I resolve to do X, *unless* Y or Z happen, in which case to hell with it.” Another form of conditional

decision is the [assurance contract](#) in which you resolve to do X if some condition is met (e.g. if someone else, or some minimum number of someone-elses, make a similar resolution).

What would Brian Boitano do?

When you are conflicted or indecisive, it may help to imaginatively recast your decision as though it were being confronted by someone you especially admire—even if, or maybe especially if, that person is largely a speculative anthropomorphized projection of your values and aspirations. What would a person whose judgement you most admire do in your situation? [\[13\]](#)



“What Would Jesus Do?” bracelets are meant to prompt Christians to make better decisions

Imagine counterfactual restrictions on your choice

There are several “mind hacks” that are meant to improve your decision-making by asking you to imagine a scenario in which your range of choices were restricted in some way.

For example, when you find yourself unable to decide between two possible alternatives, you may be counseled to flip a coin—but not to decide based on the results of the coin-flip, but on how, once the coin is in the air, you come to realize you *hope* the coin-flip turns out.

Another example is to change your perspective on the decision you are contemplating: Try to imagine a future in which you have already made the decision and you are looking back on it. [\[13\]](#) In one version of this technique, known as a [“premortem.”](#) you imagine that your favored decision has turned out disastrously, and you try to preconstruct how that disaster happened.

If some decision begins to seem inevitable, try to imagine a world in which that decision were impossible: what then would you do instead? [\[33\]](#)

These hacks, and others like them, are designed to help you more vividly and thoroughly explore your space of options, so that when you do make a decision, you do so from a more well-informed place and so can do so more confidently.

Burning your boats and other “commitment mechanisms”

One way to better guarantee that you follow through on your resolutions is to make them more difficult to reverse.^[13] For example: once you’ve tattooed “MS-13” across your face, you become much less likely to deviate from your chosen career path.

[Rosabeth Moss Kanter](#), in her study of utopian intentional communities, found that the ones that required their members to make strong sacrifices in order to join (such as surrendering their assets to the group) had more staying power.^[34]

For some decisions, there are built-in disincentives to backtrack. For others, the [sunk-cost fallacy](#) can (arguably) work in your favor in this way. The virtue of [shame](#) can also discourage you from abandoning your resolutions. Where these are of no help, you may need to [artificially induce this sort of commitment](#).

The idea of “burning your boats” originates from the decision of conquistador Hernan Cortés to scuttle his ships off Mexico so as to eliminate the option for his crew to mutiny and return to Cuba rather than continue to battle toward Tenochtitlán. Another legendary maritime example is the [Ulysses pact](#). Ulysses, knowing that as his ship passed the Sirens they would convincingly persuade him to abandon his rational decisions and instead doom himself and his crew, had himself strapped to the mast so that he could not act on their directives, and his crews ears stopped up so that they could neither hear the Sirens nor his inevitably deranged commands.

The burn-your-boats gambit is an extreme version of a more mundane strategy, which is simply to arrange your environment as best you can so that it supports the resolution you have made. If you resolve to exercise every day, then put your exercise equipment in a convenient place, free from any clutter you will need to clear away first, etc. If you resolve to stick to some diet, make sure that diet’s recommended foods are the easiest ones to grab when you’re hungry.

Brinksmanship

Brinksmanship uses displays of resolve as a signalling device. In the game of “[chicken](#)”, the drivers of two cars drive at speed directly towards each other. The first person to deviate from this collision-course in order to avoid the fatal catastrophe is declared the “chicken” and by convention loses the game. A commitment mechanism that would work as a legible signalling device in such a game would be for a driver to remove the steering wheel from the car and throw it out the window, thereby convincing the other driver that their only alternative to intolerable calamity is concession.

Artificial incentives

Applications like [Beeminder](#) and [stikK](#) allow you to add financial costs to your failure to follow-through with a variety of decisions. For example, if you resolve to lose weight, you can rig up Beeminder such that it takes a certain amount of money from you if you deviate too far from your weight target. Possibly even more motivating, you can configure stikK to donate your money to your least-favorite political cause if you fail to follow through on your decision.

Positive incentives can work, too. You might sweeten the pot of a decision you make by adding a superfluous reward to it. “Not only do I resolve to lose fifteen pounds by Summer, but if I do I’m going to treat myself to a weekend at favorite-place.” @sarahconstantin suggested that you reconfigure your social network to surround yourself with people who give you positive social feedback for following-through on your commitments.^[35]

As a motivational hack towards any kind of project, it really helps to set yourself up to have recurrent social interactions with people who support you in that project.... Actually select for people who *like* the thing you’re into, and it’s astonishing how much it’ll feel like the “world” supports you!

Exercise [caution](#) around commitment mechanisms

Commitment mechanisms can be abused. Used carefully, they help you make decisions and stick to them. But used unwisely, they can be a way of “tying your hands” so that you can excuse yourself from having to make decisions.

Remind yourself of your big-picture goals

Often, our difficult decisions involve means to ends that are themselves means to greater ends. The intermediate ends may not be very motivating, but while we are focused on accomplishing on them we may lose sight of more attractive and more distant goals. It can help you motivate your decision-making and your follow-through to “keep your eyes on the prize.”

@Emiya uses a version of this technique called [“The Evil Master Plan File”](#):

I feel pretty silly explaining it to others, and I can say this works only based on my experience.

I keep a file that has my Master Plan to “take over the world” in it (basically your end goal, what you want to obtain in the real long term).

I go look at it once in a while, to remember what I’m trying to do and sticking with it. I keep it as organised as possible, with partial objectives and the steps I need to do to succeed in these objectives....

I might occasionally try to gloat and laugh evilly as I work on it. I find super villain mentality to be highly motivating, which is why I refer to as “my Master Plan to take over the world” rather than using a more realistic and precise name. I feel that without the super villain related stuff I wouldn’t find updating and working on it as interesting or amusing, and I’d risk losing track of it.

Memorialize your decision

This may also be a good reason to make your decision iconic, momentous, charged with portent, full of oaths and totems—if only because these things make the moment of your decisive conclusion easier to recall.^[23]

Sometimes people mark their commitments with a ceremony (like a housewarming, a baby shower, a wedding, taking an oath of office, writing down New Year’s resolutions). Another way you may mark your commitments is by incorporating them into your identity (e.g. as your “purpose,” “calling,” or “sense of vocation”).^[36]

If a resolve represents an *ongoing* commitment, without any obvious finish-line or discrete accomplishment, it may be helpful to celebrate milestones. [Sobriety chips](#) that mark a

certain amount of time that a person has remained sober are one example. “I’d really like a drink right now, but I’m so close to my six-month chip” may seem like an unlikely motivator, but it can work. This is an example of the strategy of [gamification](#)—taking something that you *wish* you wanted to do and turning it into something you genuinely want to do by adding artificial elements to it that make it fun and intrinsically motivating.



Take care not to make too many values purely instrumental

Modern economic life is manic with possibility, and encourages us to be ever-ready for opportunity. As a result, we tend to view much of what we are, do, and own in terms of its exchange value. That way we can always trade away what we have today for a better offer tomorrow.

Money itself is a way of earning, gaining, accumulating, but at the same time keeping our options open about what it is exactly that we have earned, gained, or accumulated. Choosing something and taking stewardship of it—committing to taking care of it and learning how to use it well—or choosing something because it is special and intrinsically valuable to you, because you love it or find it sacred—is less common in an age of mass-produced,

disposable, obsolescing consumer artifacts and an economy in which most of us have to market ourselves to employers and so have to be concerned about our own exchange value. [\[37\]](#)

This sort of thing can retard decisive commitment. It is hard to commit to making your house a home, for example, if in the back of your mind you think of your house primarily as an investment whose market value you want to preserve or enhance. It is hard to commit to living your best life when a nagging inner-voice asks you “but how’s that going to look on my résumé?”

Incremental micro-decisions in the place of commencement-decisions

“If you correct course at a high enough frequency, you can be simultaneously decisive at a micro scale and tentative at a macro scale. The result is a somewhat winding path, but executed very rapidly, like the path a running back takes downfield. And in practice there’s less backtracking than you might expect.” —Paul Graham[\[38\]](#)

We may daydream about pivotal moments when we can bring about the right result all of a sudden in a dramatic and decisive way. But more typically it takes long-term, persistent, incremental work to move the needle.

When I meditate, I’ve often begun with a resolution (usually, in this context, called an “intention” for obscure cultural reasons) like “I’m going to keep my conscious mind attending to the sensations of breathing for the duration of the meditation session.” And, to put it plainly: a hell of a lot of good that does me. Ten minutes in and I’m lost in a daydream with my breath unattended to.

@moridinamael points out that there’s another way to go about this. [\[39\]](#) Rather than making a big commitment up-front and relying on your willpower to keep your hand steady on the rudder, make little commitments as you go, by building those commitments in—recursively—to the task you’re committing to. This works in meditation (it’s improved mine), and you may be able to extend it to decisions off-the-cushion as well.

A similar technique is @Unreal’s “policy-based” (as opposed to “willpower-based”) decisions. [\[40\]](#) By establishing a *policy-for-deciding* rather than a *decision*, you don’t make your initial resolution do all the work:

It basically costs no willpower to implement the policy. I’m not having to nudge myself, “Now remember I decided I’d do X in these situations.” I’m not having to consciously hold the intention in my mind. It’s more like I changed the underlying code—the old, default behavior—and now it just runs the new script automatically.

The “one day at a time” slogan that is common in addiction recovery is another example of this. Sincere vows to be sober henceforth and forever aren’t worth a whole lot, as the addict often discovers. Instead, the daily work of staying sober for the rest of today is where it’s at.

But there may be a class of decisions for which a decisive initial gambit is the only way to make them happen—where the only micro-intentions available are those that delay or evade the decision. You don’t always notice when you’re making a decision if that decision is merely to stick with the status quo or with already-established habits.

A woman from Alabama dreams of visiting Italy. One year she has the chance to go but postpones the trip because of responsibilities at work. Time slips by, and she thinks often of Italy, but years turn into decades, and eventually her health deteriorates to the point where she can’t make the trip. When, exactly, did she “choose” not to visit Italy? Was it

every day? Or never? She surely never expected that her first decision, to postpone the trip, would become a permanent one.^[41]

Insights about [near mode and far mode thinking](#) may be important in understanding how to make effective resolutions. If you make your resolutions in far mode, but ultimately decide on your actions in near mode, they may get out of sync. One way this problem is addressed in the addiction recovery context is for the addict to rehearse real-world near-mode scenarios (either in the imagination, or by doing role-playing with others) in which they might be tempted to deviate from their far-mode resolution to stop using, and how they will successfully respond to such temptations. This makes it easier for them to translate their far-mode resolution into near-mode behaviors that implement the resolution.^[42] You may recognize this as a variety of [trigger-action planning](#).

Ignatian discernment

Ignatius of Loyola developed an influential technique for making important decisions that, while it is tightly coupled to Christian metaphysical assumptions, may have some hints of broader applicability.^[43] Here's my attempt at a probably overly-simplified and overly-secularized paraphrase:

To make a good choice, make sure that your most important, primary end is first in mind. Don't get lost in the details before you begin, and make sure that you fit the means to the end and not the end to the means.

When you are making an important choice—for instance a “how shall I live?” sort of decision—there are three ways you might go about it. First, if you are lucky, God will make your path clear to you in an unmistakable and incontrovertible way, and then all you have to do is hop to it. Second, your attentive “experience of desolations and consolations” (spiritual struggles in which you zig-zag towards and away from God) “and discernment of diverse spirits” (spirits of different sorts may urge you to either good or bad decisions) may throw light on your conundrum and show you the way forward.

However, sometimes you don't have either of those to work with. In such cases, get yourself some breathing room and try the third method, for which there are two techniques:

1. Bring to mind whatever it is you are deliberating about: whatever it is that depends on your choice. Keeping your primary end in mind, begin your deliberation in equilibrium, without a bias toward deciding one way or another. Ask God for help in deciding correctly, then use your understanding to weigh the advantages and disadvantages of the various alternatives. After pondering in this way every aspect of the matter, use your reason to choose the best alternative.
2. If you are already leaning towards making a particular choice, double-check your motivations to make sure this is for the right reasons. Consider an imaginary third-party confronting this choice, and imagine that you want the best for them. What advice would you give them? Consider what choice you would make if you were on death's door—if this were the last decision you were going to make. Consider what choice would you wish you had made if you were defending your choice before the Divine Judge. Then, having considered these things, pick the choice that stands up best to these considerations.

Then, either way, offer your choice up in prayer and listen for God's opinion.

Set a tripwire

If you find yourself indecisively muddling and you suspect you would be better off making a decision, but you are having a hard time choosing, you might consider setting a “tripwire” that will compel a decision. This can be an arbitrary thing like a self-imposed deadline (“If I haven’t made up my mind by Thursday, I’ll flip a coin”) or it can be a condition (“if one more person quits, I’ll start sending my résumé around”).^[44]

Sleep on it

People can be influenced in their decision-making by their emotional state, by characteristics of their environments, and by [“decision fatigue”](#) (people tend to make poorer decisions, or simpler decisions whether or not they are good ones, if they have recently had to make a lot of other decisions).

So one way to have more confidence in your decision-making is to make a tentative decision, sleep on it, and then see if that decision still seems like a good idea in the morning.^[45] This evens out some of the transient emotional and environmental effects.

Simplify

Another way of addressing decision fatigue is through [simplification](#). I have heard it said^[46] that executives like Steve Jobs and Mark Zuckerberg who wear monotonously consistent wardrobes do so in order to reduce the number of decisions they have to make. This way they can conserve their decision-making mojo for decisions that really count.

[Concentration](#) is a sort of decisiveness—you resolve to commit your attention span to one thing rather than letting it ricochet among the many distractions trying to claw it away. People who have difficulty focusing their attention may be suffering from the same sort of fear-of-missing-out that can make decisiveness in general difficult.^[23]

Give up

People who begin the Alcoholics Anonymous program are typically people who have decided to stop drinking... over and over again. AA begins with the counterintuitive idea that the key to not drinking is to give up on deciding not to drink.^[47] The alcoholic *cannot* effectively decide to stop drinking—such is the nature of the beast—and so should stop trying. Instead alcoholics need to hand the steering wheel over to “a Power greater than ourselves” and let this Power do the steering.^[48]

It sounds fishy to me, but it seems to work for many people when nothing else does. While the canonical Power is, of course, God (“as we understood Him”^[48]), AA members have invented a vast and varied pantheon of “higher powers” to rely on, many of which aren’t very god-like at all.

Giving up can also be an anti-pattern that interferes with decisiveness. Giving up is sometimes a coping strategy for dealing with stress. If you over-rely on it, the solution might be to strengthen some of the many other possible coping strategies. But this shades into other virtues like endurance, perseverance, and patience, so I’ll stop there.

Be Attentive

Iris Murdoch believed that decisiveness was not so much a particular activity that happens only once you reach the crossroads, but is a sort of natural side-effect of being habitually, skillfully [attentive](#):

[I]f we consider what the work of attention is like, how continuously it goes on, and how imperceptibly it builds up structures of value round about us, we shall not be surprised that at crucial moments of choice most of the business of choosing is already over. This does not imply that we are not free, certainly not. But it implies that the exercise of our freedom is a piecemeal business which goes on all the time and not a grandiose leaping about unimpeded at important moments. The moral life, on this view, is something that goes on continually...[\[49\]](#)

Conclusion

If life seems to happen to you, and you find yourself wondering “when did I decide to live this way?” it may be time for you to develop some resolve. Resolve seems to be more of an art than a science. How do you know when the moment to decide has arrived? How do you know which decisions require discrete resolutions, and which would work best with ongoing micro-resolutions, and which would be better addressed through adaptability rather than decisiveness? How do you balance being flexible enough to cut your losses with being resolute enough to weather misfortune?

I had some hope when I started the research for this post that I would find some good answers, or at least rules-of-thumb, for questions like these. No such luck. This seems to remain one of those areas of life where you hone a fuzzy intuition by trial and error, and where you may never be able to fully articulate your heuristics.

I hope that some of what I have dug up will help in making this process a little susceptible to conscious awareness and to rational honing.

1. [^](#)

Pete Davis *Dedicated: the case for commitment in an age of infinite browsing* (2021)
chapter 4

2. [^](#)

“the defining characteristic of my generation”—Pete Davis *Dedicated* (2021)

3. [^](#)

Pete Davis *Dedicated* (2021) p. 33

4. [^](#)

Hannah Arendt “Thinking and Moral Considerations” *Social Research* 38 (1971) p. 438

5. [^](#)

Arne Johan Vetlesen, summarizing Hannah Arendt’s point of view, in *Evil and Human Agency: Understanding Collective EvilDoing* (2005)

6. [^](#)

Aristotle *Nicomachean Ethics* VI.2

7. [^](#)

Chun Siong Soon, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes
[“Unconscious determinants of free decisions in the human brain”](#) *Nature Neuroscience* 11 (2008) pp. 543–45.

Other examples of this sort of thing include Benjamin Libet, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl "[Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity \(Readiness-Potential\)—The Unconscious Initiation of a Freely Voluntary Act](#)" *Brain* 106 (1983) pp. 623–642; Chun Siong Soon, Anna Hanxi He, Stefan Bode, and John-Dylan Haynes "[Predicting free choices for abstract intentions](#)" *Proceedings of the National Academy of Sciences USA* 110 (2013) pp. 6217–22; and Roger Koenig-Robert, and Joel Pearson, "[Decoding the contents and strength of imagery before volitional engagement](#)" *Scientific Reports* 9 (2019) #3504

8. ^

The jury is still out about less-arbitrary decisions. See for example Uri Maoz, Gideon Yaffe, Christof Koch, and Liad Mudrik "[Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice](#)" *eLife* 8 (2019).

9. ^

Antonio Damasio *Descartes' Error: Emotion, Reason, and the Human Brain* (1994)

10. ^

Chip Heath & Dan Heath *Decisive: How to Make Better Choices in Life and Work* Crown Publishing (2013) p. 23

11. ^

Chip Heath & Dan Heath *Decisive* (2013) p. 160

12. ^

Jonah Lehrer *How We Decide* Houghton Mifflin Harcourt (2009). This book is mostly about how to make better decisions, rather than how to be decisive or resolute, however, so I don't refer to it much here. Also, there are [some questions](#) about Lehrer's scholarship.

13. ^

Pete Davis *Dedicated* (2021) chapter 7

14. ^

Chip Heath & Dan Heath *Decisive* (2013) p. 25

15. ^

See e.g. Robin Hogarth *Educating Intuition* (2001).

16. ^

Koturski, Joseph "The Ethics of Aristotle" The Teaching Company: The Great Courses

17. ^

See also the discussion of "[ambition and aspiration](#)" in my [Notes on Ambition](#)

18. ^

Pete Davis *Dedicated* (2021) chapter 13

19. [^](#)

Pete Davis *Dedicated* (2021) chapter 1
20. [^](#)

Ralph Waldo Emerson “Self-Reliance” *Essays* (1841)
21. [^](#)

Søren Kierkegaard *Either/Or* (1843)
22. [^](#)

bryjnar [“Choice begets regret”](#) LessWrong 4 January 2018
23. [^](#)

Pete Davis *Dedicated* (2021) chapter 9
24. [^](#)

Pete Davis *Dedicated* (2021) chapter 11
25. [^](#)

An excellent meditation on how people use “serious” decisions and loyalties as a way of trying to evade the burden of choice is in part II of Simone de Beauvoir’s *The Ethics of Ambiguity* (1947)
26. [^](#)

Eliezer Yudkowsky [“Just Lose Hope Already”](#) LessWrong 24 February 2007
27. [^](#)

Aristotle, [Nicomachean Ethics VII.9](#)
28. [^](#)

Werner Erhard *The Hunger Project: An Idea Whose Time Has Come* (1977) p. 28
29. [^](#)

Hunger Project *Ending Hunger: An Idea Whose Time Has Come* (1985) (these quotes all come from page 3)
30. [^](#)

The global hunger situation certainly improved between 1977 and 1990, but not appreciably more impressively than it had between 1970 and 1977.
31. [^](#)

Chip Heath & Dan Heath *Decisive* (2013) p. 252
32. [^](#)

Linda Linsefors [“Non Resolve as Resolve”](#) LessWrong 10 July 2018

33. ^

Steven Johnson *Farsighted: How We Make the Decisions That Matter the Most* (2018)
p. 68

34. ^

Rosabeth Moss Kanter *Commitment and Community: Communes and Utopias in Sociological Perspective* (1972)

35. ^

sarahconstantin "['Cheat to Win': Engineering Positive Social Feedback](#)" LessWrong 5
February 2018

36. ^

Pete Davis *Dedicated* (2021) chapter 7; however in chapter 8 he also notes that this can be another factor in making people reluctant to commit: new commitments can threaten old identities. Commitments are also often package deals: you have to take the good with the bad, and this includes the effect on your identity.

37. ^

Pete Davis *Dedicated* (2021) chapter 10

38. ^

Paul Graham "[What I've Learned from Users](#)" September 2022

39. ^

moridinamael "[Spamming Micro-Intentions to Generate Willpower](#)" LessWrong 13
February 2018

40. ^

Unreal "[Policy-Based vs Willpower-Based Intentions](#)" LessWrong 27 February 2019

41. ^

Chip Heath & Dan Heath *Decisive* (2013) pp. 221-22

42. ^

Michael W. Otto, et al. "[Attending to emotional cues for drug abuse: bridging the gap between clinic and home behaviors.](#)" *Science & Practice Perspectives* (2007) pp. 48-56

43. ^

Ignatius of Loyola [Spiritual Exercises](#) (1548) §169-188

44. ^

Chip Heath & Dan Heath *Decisive* (2013) p. 227

45. ^

Steven Johnson *Farsighted* (2018) pp. 142-43

46. ^

e.g. Vincent Carlos ["Why Successful People Wear The Same Thing Every Day"](#) *The Startup* 31 October 2019

47. ^

Step one of the Twelve Steps: “We admitted we were powerless over alcohol—that our lives had become unmanageable.”

48. ^

Steps two and three of the Twelve Steps: “Came to believe that a Power greater than ourselves could restore us to sanity,” and “Made a decision to turn our will and our lives over to the care of God as we understood Him.”

49. ^

Iris Murdoch, “The Idea of Perfection” (1962) in *Existentialists and Mystics*, pp. 299-336.

Notes on Caution

This post examines the virtue of **caution**. As with my other posts in this sequence, I'm less interested in breaking new ground and more in gathering and synthesizing whatever wisdom I could find on the subject. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will help people who want to know more about this virtue and how to nurture it.

What is caution?

To a first approximation, someone with the virtue of caution is habitually alert to potential risks, weighs those risks well when making decisions, and deploys strategies to mitigate them.

I want to distinguish this definition from another one that suggests itself: Caution as the habit of making choices that minimize possible negative consequences. That is "risk aversion" and can be part of caution, or in an unbalanced form can become an unwise timidity.

However, unlike [phrónēsis](#) (prudence, practical wisdom), caution does focus on negative outcomes.

Is it a mistake to call the virtue "caution"?

Does framing this virtue as "caution"—rather than just subsuming caution under "practical wisdom"—put a thumb on the scale rather than weighing things wisely? How can you know ahead of time whether the best decision is more cautious rather than more bold? There is something to that criticism, but here are some reasons why this framing might nonetheless make sense:

For one thing, a virtue is a habitual variety of choice-making. There may well be advantages to making cautious choices by habit (and therefore by default) and incautious ones as an exception that requires a more deliberate override. Perhaps in any particular instance there is no good reason to prejudge it as being best met by caution, before you have a chance to deliberate, but for those many choices we make without deliberating much, a cautious default mode may save us a lot of trouble.

For another, the virtue of caution—properly understood—has to do less with risk-averse *results* of your [decision-making](#) than with a risk-aware *process* of decision-making. Caution does not insist that you make risk-minimizing decisions, but that you be alert for risks and carefully weigh those risks during your deliberation.

For another, in human lives risk/reward is not symmetrical. For example, death is a sort of risk-singularity for which there is no reward counterpart (unless you are among the believers in the possibility of eternal life). People do not see the range of possible future up-sides and down-sides as equivalent opposites, and this is not merely an irrational bias, but reflects

certain asymmetries in the human condition.

Jordan Chase-Yog-Sothoth 
@jachaseyoung

...

(a) Would you endure 100 years of the most intense pain possible if you got to enjoy 1 million years of the most intense joy afterward? (2) Would you agree to 1 million years of max joy if you then had to endure 100 years of max pain?

(a) Yes, (b) Yes	22.9%
(a) Yes, (b) No	13.1%
(a) No, (b) Yes	4.8%
(a) No, (b) No	59.2%

1,400 votes · Final results

3:37 AM · Sep 25, 2022 · Twitter for Android

A more mundane possibility is that the virtue is called “caution” because recklessness is more common or more problematic than timidity, and so if you’re counseling someone to aim at [the golden mean](#), you’ll more likely be telling them to be more cautious. This would be similar to how we call the virtue [“patience”](#) even though there are a minority of people who are too long-suffering and would be better off losing their cool and not putting up with so much.

Caution and “prudence”

In my [Notes on Prudence](#) I noted that “prudence” was the traditional translation for *phrónēsis* (Greek) and *prudentia* (Latin) as the name of the virtue of practical wisdom, but that the word “prudence” has since become more associated with caution specifically. For example, the VIA Institute on Character, which takes a modern virtue-based approach, summarizes prudence in this way: “I act carefully and cautiously, looking to avoid unnecessary risks and planning with the future in mind.”^[1]

I chose to stick with the more traditional and expansive definition of prudence-as-practical-wisdom when I discussed that virtue, and I’ll discuss this more restricted sense of prudence-as-caution in this post.

That said, it is difficult to discern where caution ends and practical wisdom begins. If you act or deliberate to avoid a danger or threat, you’re acting cautiously. But what if you do so in order to avoid something that is merely suboptimal: if you are careful not so much to avoid a danger but to avoid missing an opportunity? Is the consideration “I don’t want to look back on my life and realize I never took any big chances” also a *cautious* consideration?

Risks are complicated

The risks you are subject to at any time include a shifting variety of potential threats to your health, reputation, property, comfort, plans, family & friends, values, things you hold sacred, and so forth. In order to evaluate and respond to such diverse risks well, you must also at least implicitly have some way of ranking these many things or of evaluating trade-offs between them. If you value certain of these components in a superstitiously or disproportionately imbalanced way, your evaluation of the portfolio of risks will also be imbalanced, and your response to those risks will suffer.

You may find that you have a large-scale risk strategy that is subject to change, with effects that ripple throughout your decisions and habits. For example, at some point in your life you may have adopted a roughly “[maximin](#)” outlook toward life: trying to make decisions that protect you from the worst down-sides, but perhaps at the cost of taking risks that could gain you some of the better up-sides. Then something in your life or outlook changes and you decide that you would be better off sticking your neck out and reaching for the brass ring. This is one interpretation of the midlife-crisis: someone has arranged their life on conservative, playing-it-safe principles, and then abruptly asks “is this all there is?” and decides to switch strategies. What makes the midlife-crisis sometimes comic is the awkward way in which such a strategy tweak ripples out in the form of uncharacteristic decisions and value-rankings and risk tolerances—ones that can seem incongruous and ridiculous, at least until the pendulums resettle.

All of your possible options have consequences that in turn present their own sets of risks, for many of which you can only roughly estimate the probability distribution of the various possible degrees of severity. Each of those risks in turn invites a set of additional options in the form of possible mitigation strategies or other responses, and so *ad infinitum*. Each option also has potential *benefits*, and among the risks of each option is the set of opportunity costs of not having chosen differently.

For that matter, there’s an opportunity cost paid for putting in the effort to be better-informed about risk, better prepared, more thorough in evaluating options, and so forth. This too is a risk, and a chicken-and-egg problem: how do you know if this cost is worth paying or is too high until you’ve already paid it?

Given all this, it is probably beyond our grasp to choose optimally. We need shortcuts.

Unfortunately, many common heuristics are demonstrably faulty. Using them, people often estimate risks poorly and plan for them badly. Lists of typical human cognitive biases show few that are not also ways risk-assessment can go awry.^[2] We seem to have a variety of contradictory devices that are good enough to help us make the day-to-day quick decisions we need to muddle through life, but that reveal themselves to be shockingly absurd when examined closely.

The popularity of casino gambling, and its addictiveness in some people, suggests that even when we gamify simple scenarios of risk management and provide prompt negative feedback for poor risk assessment, people can fail to correct appropriately.

Certainly if the stakes are high enough and we have enough time to think about it, we would be wise to insist on more rational methods than “just eyeballing it” with our ramshackle instincts. This is especially true in circumstances in which we are exposed to risks very different from those our ancestors would have faced—such as driving on the freeway, starting a course of chemotherapy, or sharing an unguarded opinion on an internet forum. In such cases we can expect even less reliable help from our instinctual heuristics.

One way we may be able to improve the reliability of our caution is to be better aware of cognitive biases so that we are more apt to notice when they lead us astray, or so that we can correct for their effects. For example, people are often tripped up by risks that are

individually negligible but either cumulatively (cigarette smoking) or occasionally (fatigued driving) tremendous. Because these choices only rarely result in immediately threatening consequences, we may not instinctively regard them as risks to be taken into account. We can adjust for this by more deliberately and rationally looking on them as real risks.

It is unintuitive to judge tiny probabilities of catastrophic outcomes, but these can be really important. Plenty of people die from catastrophes that were themselves improbable in their specifics, but that are not at all exceptional when you see them as making up part of the large class of ordinary accidents to which mortals are liable. You are assailed on all sides at all times by a swarm of tiny risks, no one of which is at all likely to be instantly threatening, but each of which could happen to be the one that has your number. A habit of carefulness (that is, the virtue of caution) helps you to better your odds against the whole swarm, whereas a painstakingly rational calculation to counteract a habitual incautiousness on a risk-by-risk basis is comparatively expensive and ineffective.

Related virtues and vices

The vice of deficiency goes by names like recklessness, carelessness, incautiousness. The vice of excess is timidity.

There are other failure modes of cautiousness that don't map well to linear deficiency/excess. For example, phobias and superstitions can cause carefulness to be poorly-targeted at the wrong sorts of threats. People often also exaggerate certain sources of risk in irrational ways that don't rise to the level of phobia (things like air travel feeling riskier than road travel). Advertisers and political manipulators may drum up fears or promote safety-mimicking responses that are really only in their own interests. When risks are difficult to understand, or whenever "demonstrating safety" becomes more important than being cautious, [rational](#) action can give way to rituals and totems or to "safety theater."

#yolo (see below) can go beyond recklessness to be intentional risk-seeking for thrills or for show.

Some people expect the universe to be just, to right the scales, to protect the righteous, to deal a fair hand. That turns out not to be a practical substitute for caution. We may not realize we're counting on the cosmos to take our side until it fails us, and then we catch ourselves saying things like "how could they?" or "why me?"

Virtues in tension with caution include [optimism](#) (which may obscure caution's concern for downsides with its focus on the upsides) and boldness/daring (which encourages you to take high-reward gambles now and again). A common failure of caution is when you over-represent the possible rewards of risks in your calculations and then choose unwisely based on this exaggeration. This seems to be at least part of what's going on in casino gambling and lotteries-wins are big and flashy (though infrequent); losses are small and subdued (but cumulatively costly).

Virtues that can come to the assistance of caution include [attention](#), [focus](#), foresight, and curiosity (how could this go wrong?). Willful ignorance can interfere with risk analysis. It is easy to get in the habit of substituting the most convenient of either "here there be dragons" or "nothing to see here" as placeholders for things you haven't investigated well. If you are already motivated to either engage in or refrain from some course of action, you may be tempted to use placeholders like these to justify such a decision.

[Subject-matter expertise](#) can help you to better intuit when things don't smell right and it's time to get up your guard. ["Wisdom"](#) more generally helps you anticipate the variety of potential dangers, estimate their likelihoods, and mitigate them most efficiently.

If we lack [courage](#), some sorts of frightening risks will seem more dangerous than they really are because we “fear fear itself.” Lack of courage can masquerade as caution and can hide the fact that we are deficient in caution. Much of the challenge of courage has to do with mastering our emotional response to fear, whereas much of the challenge of caution has to do with the cognitive challenge of assessing risk well. Still, there is some overlap, and some people who think of themselves as overly risk-averse may need to work on courage as much as or more than on risk-assessment.

Caution is a component or ingredient in virtues like [frugality](#), [care](#), prudence, preparedness, [fitness](#), and know-how.

#yolo

Sometimes (and always, to some extent) danger envelops us such that all of our paths forward are dangerous ones. And sometimes we may rationally decide that the promise of exceptional gains merits taking on exceptional risks. But there are also occasions on which some people appear to be intentionally incautious, as if recklessness was itself desirable—they flout caution in order to take on what seem to be gratuitous risks. Why might this be, and can this be justified? Here are some possibilities:

- It is an excusable folly of youth.
 - There may be some [survivorship bias](#) here, but that probably can’t explain all of it.
 - Youth are less experienced, don’t always have a good grasp of consequences, and haven’t yet been burned-once in order to be twice-shy.
 - Perhaps the rewards of risk-taking are higher at a younger age, or the costs are lower? When you are young, you may have more of a support network to pick you up when you fall, for one thing. And maybe a higher percentage of rewards are available through one-off feats of derring-do during youth, while adults are better equipped to cash in on slow-and-steady long-term plans. I’m skeptical and would want to see the math before I’d accept this argument, though.
 - It could be [peacocking](#), in a sexual selection sense.^[3] That’s a plausible just-so-story, anyway.^[4] This might also explain the stereotypical sexual dimorphism in #yolo behavior. Men might be more punished than women for being mediocre or more rewarded for being exceptional, in terms of reproductive success, and this could encourage young men to take more chances.
- It is an impressive demonstration of courage. People often think of those who court danger, at least in some ways, as being formidable, dashing, and vigorous (not merely endangered, vain, or foolhardy).
- To the person who confronts danger, this can be thrilling, either from the adrenaline rush or from the more cerebral sense that one is dancing with death, taking one’s destiny into one’s hands, living fully in the moment, and that sort of thing. For example, a researcher who studied skydiving enthusiasts said such a person “passes from a state of fear to a condition of happiness and excitement characterized by hyperrealism and perception of time which is focused on the present. Once the risk behaviour has come to a happy end, a very pleasant feeling of self-realization develops. [They] talk about having felt a kind of purification, an amplification of the self and of feeling a higher level of self-determination.”^[5]
- Curiosity may also be a motive for incautious behavior (what will *really* happen if I stick a fork in the light socket?).
- Deliberately putting yourself into a dangerous situation can be a way of testing your limits and learning more about your capacity or [resilience](#).
- You might put yourself in “danger” because you idiosyncratically do not believe in the danger, and you hope to debunk invalid warnings.
- Some dangers are only apparent (for example, those in certain stunts or magic tricks).

- You might deliberately tempt disaster in order to learn from mistakes. When the stakes are low, it can be useful to practice being innovative to overcome deliberately invited challenges. For example, a strong player in chess may offer to handicap themselves against a weaker opponent (e.g. by removing certain of their pieces before gameplay begins) in order to better learn from the game.

On reflection, some of these excuses for incautious-seeming behavior strike me as inadequate. In particular, the excuse that you “feel more alive” when dancing with death I think deserves more scrutiny than it’s usually given. We may nod or shrug when we hear this excuse from, say, “free solo” rock climbers, but it should raise our suspicions that we hear similar words from people who deliberately cut or burn themselves as a symptom of depression or certain personality disorders. “Depend upon it, sir, when a man knows he is to be hanged in a fortnight, it concentrates his mind wonderfully,” said Samuel Johnson,^[6] but such wonders aren’t worth getting hanged over. (There are [healthier ways to improve your focus](#).) If you feel you need to brush up against the final curtain in order to feel alive, maybe that’s more an issue to confront than an urge to indulge.

That said, it’s true that times of danger can, in retrospect, be charged with significance and feel like meaningful, valuable landmarks in your life: times when you were put to the test and your character developed a new polish. What distinguishes these from mere #yolo stunts, though, is that typically they involve real stakes. Surviving a “[hold my beer](#)” moment may be thrilling, but is unlikely to be life-defining (except in a bad way, e.g. if you are maimed).

Sometimes you hear it said that “it’s better to regret something you have done than to regret something you haven’t done”^[7] or you hear that people, as they approach the end of their lives, regret not having taken more chances. I think there is good reason to be suspicious of this line of thought. For one thing, it’s too easy to cast the glow of the best possible outcome onto risks you regret not having taken. Really, what you regret is missing out on that rosy-colored fantasy-outcome. The real-life distribution of likely outcomes was probably far different. For another, I don’t think end-of-life regrets of this sort deserve uncritical respect, even if they turn out to be more than anecdotal. Survivorship bias is one reason (those who live long enough to have regrets about their caution may have been more well-served by their caution than they realize). But also, at the end of life you may sensibly discount mere survival: caution to preserve your few waning years may no longer seem as valuable as it once did. Meanwhile, risks for great rewards may seem more out-of-reach than before. When you cast your mind back to when you had an opportunity to roll the dice you may wish you had another chance since the wager now seems more in your favor.

However, it can make sense to be cautious about being overcautious. You don’t want to be so afraid of taking risks that you lose out on opportunities for worthwhile rewards, or to be so obsessed about preserving your life that you are afraid to live it. But I think this, properly understood, is not in opposition to the virtue of caution but is incorporated in it.

How to develop the virtue of caution

Caution incorporates several distinguishable strategies:

1. Anticipate possible risks.
2. Avoid exposing yourself to risk unnecessarily.
3. Take preventative steps to reduce negative consequences.
4. Prepare to cope with such consequences.
5. Act with care and attention.

This sort of subdivision can be helpful because each component can go awry in its own way. You may be able to most effectively improve your caution by concentrating on one component of it.

For example, if you are frequently blindsided by things that other people seem to more-easily anticipate, you may need to work on the first of these components. If for you “what’s the worst that could happen?” is always merely a rhetorical question, you may want to apply yourself here. People who have irrational phobias also may recognize that there is something awry in the way they anticipate possible sources of risk.

People who take unnecessary risks, or who do not seem to take risks into account when they choose among options, would be wise to attend to the second subcomponent.

There are some risks that can and should be avoided, but others must be faced. For those especially you need the next two components of caution: mitigation and preparedness. You cannot prevent the rain if the storm comes, but you can bring an umbrella and an extra pair of dry socks.

Finally, although cautious *choice* involves imagining, thinking ahead, and anticipating possibilities, cautious *action* requires focus on the present and actual: taking care, not getting sloppy or skipping steps, staying alert. “Watch what you’re doing.”

Know the limits of instinct

Particularly when you are dealing with things that are very high, very deep, very sharp, very fast, very hot, very cold, very bright, very powerful, very heavy, very big, and so forth: don’t just eyeball it. Your instincts are liable to mislead you when they confront something outside of the parameters in which they evolved.

Consider instead (when these are available) the hard-won risk mitigation strategies embodied in the practices of experts, institutions, and traditions.

Don’t be too eager to override or skip safety features that seem at first glance to be unnecessary or excessive. Such things may be a small price to pay to avoid uncommon disasters. If you do things extra-cautiously by default, you’ll be doing things extra-cautiously that one time when it really matters.

Another way instinct misleads is by exaggerating the danger of things that may emotionally feel scary (harmless spiders, graveyards) and discounting objectively dangerous things that feel reassuringly comforting (climate-controlled car interiors, cigarette breaks, cocktails).

Attend to good advice, [#rtfm](#), [#lifemf](#)

Don’t waste time learning from your own mistakes when there are so many excellent prefab mistakes out there you could learn from instead. When you embark on something unfamiliar, there’s no shame in asking for advice from those with experience, or in reading the manual first.

Meditate on others’ misfortunes. Indulge your pity or your schadenfreude if you must, but use the opportunity to learn from their fail. If you [lend a sympathetic ear to those who are suffering](#), one beneficial side effect is that you may learn more about paths that lead to suffering and about choices people wish they had made differently. Where did they make a wrong turn or fail to prepare? How might you have known better?

Reporting, books, and literature of all sorts can also tell stories of disasters and of missed opportunities to avoid them (or nick-of-time decisions that did). However such stories are selected for how gripping they are rather than for how representative or accurate they are, so *caveat lector*.

Dreams and daydreams

Maybe this is just a personal foible, but I've noticed that in idle moments (or during meditation when I'm being especially alert to what spray my mind tosses up from its turbulence), my thoughts often turn to improbable horror stories. What if I were accused of a murder I didn't (or did!) commit? What if I were walking along the Golden Gate Bridge with my mother and a sudden gust of wind blew her into the deep? What if a passenger jet crashed into the parking lot of my apartment complex? (These aren't very *good* horror stories, typically, nor very realistic.) I used to get frustrated at my brain for dangling this sort of cheap macabre entertainment into my consciousness. Now I wonder whether maybe it's trying to prepare me for the sort of rare, long-tail sorts of crises that do pop up from time to time. My mind seems to take crisis elements in odd combinations, almost improv-theater style, and then present them to me as problems to solve, in my otherwise free time.

It does seem to me that some of my real-world precautions have been prompted by side-effects of such daydreams. Is it an abstract actuarial understanding or some wandering fantasy of an unlikely conflagration that actually gets me to check my smoke alarm batteries or buy a fire extinguisher?

This sort of thing seems to happen in dreams, too. They often involve me having some agenda that keeps getting frustrated by the warped Escher landscape of the dream-world, so that I have to continually improvise new solutions. Is this a kind of training for how to handle novel challenges?

Dystopian literature might also be seen in this light: as exploring certain possible-if-unlikely disaster modes as a way of preparing for or preventing them.

"Catastrophizing" is a pathological version of this. People who catastrophize exaggerate the likelihood and the danger of the worst possible outcomes of scenarios. This can result in anxiety and in excessive timidity.

Drugs and risk-taking

One way you can promote your caution is to beware of decisions you make under the influence of drugs. Alcohol, of course, is a notorious #yolo promoter. But any drug that causes altered states of consciousness as part of its menu of effects may also alter the way you detect, evaluate, and react to risk.

[Stimulants, medications and recreational drugs that tweak dopamine](#), and even [acetaminophen/paracetamol](#) can demonstrably and predictably incline people to more risky choices.

Consider the advantages of making decisions about risk by using similar brain chemistry to that with which you calibrated your risk evaluation. Or, more plainly: maybe sleep on it, and see if it still seems like a good idea tomorrow when you're sober.

More generally, your state of mind is important to your ability to practice caution. The more well-rested you are, the less stress you are under, the less clutter or distraction there is in your environment, the more capable you are of focusing, the more you will be able to make cautious decisions and take careful actions.

Recreational drug use can itself be an incautious behavior. It is in some cases an unwise risk that comes pre-packaged with additional incentives to take the risk or rewards for having taken it (the high) along with neurochemical changes that degrade or distort caution or that indeed promote risk-taking. When you look at that package as a whole, it ought at the very least to get your guard up.

Institutional help

In this post I concentrate on caution as a personal virtue, rather than on safety generally. However: Safety-oriented conventions as embodied in institutions, authorities, and traditions can supplement personal caution, and it can be a component of personal caution to regard such things wisely. However, ostensibly safety-oriented conventions can also be dangerous—even (especially?) well-meaning ones. One way to exercise [social responsibility](#) is by improving and promoting the best such conventions, while helping to usher the worst to their deserved obsolescence.

Adding safety features and procedures to some process may make that process more safe but at the same time more costly, and thereby incentivize the use of alternate processes that are even less safe, such as Paul Graham asserts happened with nuclear power in the United States:

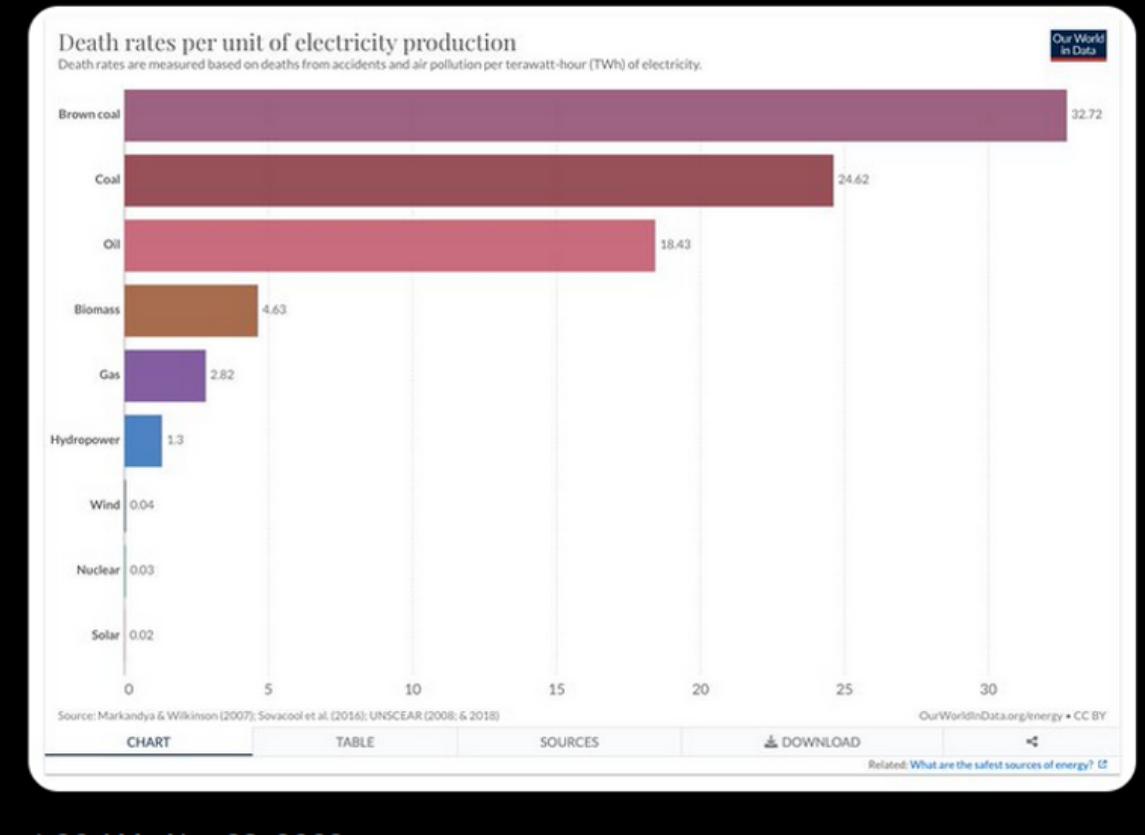


Paul Graham ✅

@paulg

...

No one realized it at the time, but the accident at Three Mile Island cost thousands of lives. Not directly, but by making nuclear power unpopular in the US, and thus causing us to switch to more dangerous power sources.



4:36 AM · Nov 28, 2022

People sometimes respond to the presence of safety features by [using them to compensate for additional risk](#) rather than allowing them to reduce risk. In particular, safety interventions that reduce the harm or likelihood of otherwise frequently-encountered negative consequences of a risk can increase people's willingness to undertake that risk and increase their exposure to otherwise less-frequently encountered negative consequences. For example, a frequently-encountered risk of speeding is traffic tickets; a less frequently-encountered risk is a fatal traffic accident. Some feature that makes you more safe from the frequently-encountered risk (a radar detector, a "back the blue" bumper sticker) may thereby encourage you to speed and expose yourself to the less frequently-encountered but more severe danger.

(One economist recommended, tongue-in-cheek I think, that [vehicles have sharp spikes protruding from the center of their steering wheels](#) such that accidents would be *more*

dangerous to drivers, in order to counteract this sort of risk compensation and thereby make driving more safe.)

Similarly, some safety features—because they are meant to supplement fallible human attention—can cause people to compensate with increased inattention. Traffic engineer [Hans Monderman](#) made a name for himself by demonstrating how removing lane markers, crosswalks, advisory signage, traffic lights, and other such safety features can make people more cautious and alert and as a result can reduce traffic accidents (a strategy called “[shared space](#)”).

You can count me among those who tend to believe that our society is over-regulated, choked with laws, and stiflingly litigious. And yet I have to acknowledge that it is difficult to come up with an objective answer about whether or not such a belief is really true. Regulations and lawsuit-preventing precautions, when they *do* successfully prevent disaster, typically do so without fanfare: Nobody notices the catastrophe that didn’t happen. It’s possible that [Ralph Nader](#) saved my life at some point, and in my obliviousness I never even knew to thank him. I bristle at the paternalism of “nudges” like taxing alcohol and subsidizing vaccines, but will grudgingly admit that, at least considered in some isolation, they can probably save lives. At the same time, the victims of overzealous safety regulation can also be hidden: people whose lives could have been saved by a medical device that was red-taped out of existence, people who turned to dangerous black-market drugs when legal alternatives were prohibited, or the many diverse benefits people have had to forego in order that the expensive demands (or, yes, nudges) of the safety bureaucracies may be satisfied.

Sometimes institutions adopt “safety” procedures, but do so under the influence of incentives that do not actually prioritize the safety of those people who will be following those procedures. It can require some sophisticated understanding of these institutions and incentives to understand when this might be the case. (Other times, with only a little experience of how the world works, you can recognize that a three-page, fine-print “important safety advisory” is really only an “our lawyers made us say something that should be obvious to anyone.”)

Regulators may come to understand that their only job is to reduce direct risk, and so they lose all sense of cost-benefit proportion. Would more lives be saved if air traffic were made cheaper by being allowed to become more dangerous, if this also meant fewer people traveling by car? Maybe so, but those who regulate airline safety are unlikely to find that saving those lives is part of their job description.

As the covid pandemic emerged, safety guidance from some of the most authoritative sources was disturbingly hit-or-miss. Public health authorities, whose incentives at times seemed to be dominated by concern about optics and ass-covering, too frequently gave advice (or promoted/hindered mitigation strategies) in a way that was sub-optimal or indeed harmful from a safety perspective. And unfortunately, the sensible alternatives to this struggled to be heard over a tsunami of superstition and snake-oil and disinformation. Anyone who wanted to behave with an appropriate level of caution has had to put in extraordinary effort to learn how to do so.

Jason Crawford’s recent post, [“Towards a Philosophy of Safety”](#) has some good thoughts on how we should think about safety in these institutional contexts, what biases we should be vigilant about, and what techniques might help us better optimize safety trade-offs. Crawford also considers how our traditional institutional safety culture, which typically springs into action to prevent repeats of disasters that have already happened, might be retooled for an era of rapid technological change in which we would be wise to prevent novel disasters from happening the first time.

The Precautionary Principle

“The Precautionary Principle” has been formulated in a variety of ways.^[8] The lack of a consensus definition means that sometimes people form strong opinions about different conceptions of it and then agree or disagree with one another without being clear about what they’re agreeing or disagreeing about.

In broad outline, this principle goes something like this: If a sufficiently bad consequence is a foreseeably possible (even though not certainly probable) outcome of some endeavor, then before we pursue that endeavor, we ought to take steps to remedy that possible consequence (even in this absence of certainty).

(There is also a closely-related “Catastrophe Principle” in which if the foreseeably possible consequence is [sufficiently apocalyptic](#), you must certainly act to prevent that consequence, even if it seems unlikely. There are some [Pascal’s Mugging](#)-like edge cases of this to beware of, but it can be a persuasive argument.^[8])

Sometimes the Precautionary Principle is deployed to put the burden-of-proof on those who propose a change from the status quo: demonstrate that you have thought through the possible consequences and that you are prepared to meet them before you go through with your plan. This is a conservative, status-quo-preserving, [Chesterton’s Fence](#)-upholding interpretation. Critics point out that sometimes bad consequences may plausibly but unprovably result not only from an action, but also from preventing that action, or from failure to act. [What standard justifies](#) holding action to a higher burden of proof than inaction?

It seems possible to interpret the Precautionary Principle in such a way that it similarly treats action and inaction, risks and remedies, advances and retreats. But then it seems to become simply the more mundane advice to anticipate and prepare for the possible consequences of whatever comes to pass, quantify risks carefully, and choose your course in life wisely based on these possible consequences: in other words, to practice caution.

Appendix: We’re Beginning Our Descent

It may be helpful, when assessing your risk of death in particular, to remind yourself of the statistically most common causes of death, so that you aren’t overly-fearful of rare but flashy sources of doom, nor astonished by the creepingly common assassins. To this end, I have composed a mnemonic poem:

Were you warned? [of March? the Ides?](#)
I’ve been warned too: triglycerides.
Chronic obstructive lung disease,
Blocked coronary arteries.
None would marvel if I croak
From diabetes or a stroke.
I miss a step; head-first I fall.
That lump was cancer after all.
A full-grown man, could it be true:
Laid on the slab just from the flu?
Before they had a good vaccine,
I rubbed my nose: Covid-19.
“It’s like he never saw the red”
(The driver explains why I’m dead).
A brain rot nobody can cure,
A “hunting accident” they’re sure.
Charts of mortality provide
These common ends and more beside.
The Steward said as we embarked:
“The exit rows are clearly marked.”

1. [Character Strengths: Prudence](#) VIA Institute on Character: *The 24 Character Strengths*
2. [Logical Fallacies when Assessing Risks](#) scip
3. [George Leybourne “The Daring Young Man on the Flying Trapeze”](#) (1867)
4. [Ryan H. Murphy “The Rationality of Literal Tide Pod Consumption”](#) *Journal of Bioeconomics* (2019)
5. [Stephen Lyng “A social psychological analysis of voluntary risk taking”](#) *American Journal of Sociology* (1990) as described in Fiorenzo Ranieri “Extreme Risk Seeking Addiction: Theory and Treatment” *British Journal of Psychotherapy* (2011)
6. [James Boswell, Life of Samuel Johnson](#) (1791), entry for September 19, 1777
7. [Butthole Surfers, “Sweat Loaf” Locust Abortion Technician](#) (1987)
8. [Neal A. Manson “Formulating the Precautionary Principle”](#) *Environmental Ethics* (2002)

Notes on Judgment and Righteous Anger

This post examines the virtues of **judgment** and **righteous anger**. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What are these virtues?

Judgment, or censure, is (at least) the ability and willingness to identify and call out something that is unjust or someone who is acting unjustly. Sometimes it goes beyond questions of justice: you can also judge someone unwise, unkind, unreasonable, and so forth.

Righteous anger is a motivating fury, usually provoked by injustice. If you are righteously angry you have noticed something (or suffered something) that violates [your sense of justice](#) in a way that provokes an emotional response that encourages you to do something about it (and often, to signal your displeasure to those around you).

We contain multitudes

I have already written about the virtues of [good temper](#), [patience & forbearance](#), and [forgiveness & clemency](#) and you might wonder if I've lost the plot by pivoting to something like "righteous anger" that seems at first to contradict those. In my defense:

1. These might not be as contradictory as they appear. You may have the virtues of forbearance, forgiveness, good temper, and the like, but still have a limit beyond which you get angry in a virtuous way. It may be that there is a time for forbearance and a time for judgment, a time for good temper and a time to be furious, and that you need skill in both modes in order to best flourish as a human being. Even forgiveness guru Jesus thought that judgment came first, forgiveness after: "If thy brother trespass against thee, rebuke him; and if he repent, forgive him."^[1]
2. I am being ecumenical in my research into virtues. Different virtues were prized by different cultures and traditions in different periods. I should not be surprised if the collection of virtues I unearth does not cohere seamlessly.
3. Some virtues are commonly given the name of one of their more imbalanced extremes, when the opposite extreme is more common or more harmful or when the pendulum is swinging against it for whatever reason. For example, the virtue concerning self-regard is sometimes called "pride" when it seems important to contrast it with poor self esteem, slavishness, etc., but is sometimes called "humility" when it seems important to contrast it with arrogance, narcissism, etc. The virtue concerning anger could be like this, where pushovers might be told they need to get mad and stand up for themselves, while hotheads might be told they need to chill out: different advice for different people, but aiming at the same [golden mean](#).

Related virtues and concepts

If judgment / righteous anger causes you to strike out against the person who provoked you, this is punishment or chastisement, which can be done more-or-less skillfully, though I don't know that punishment / chastisement rises to the level of a virtue.

Sometimes judgment / righteous anger can be a component of valor. A lot of cinema heroes are motivated by righteous anger over some villainy, as they engage in their bold, courageous acts in the name of justice.

As I mentioned, there is a tension between these virtues and forgiveness, mercy, clemency, and so forth. There is also some with equanimity / tranquility, acceptance / surrender, humility, and tolerance / charity (in the sense of being charitable towards others' foibles).

If you are deficient in the virtues of judgment and righteous anger, you might be called a pushover or schnook or someone who lets people take advantage of you (if the injustice was aimed at you), or as a ["good German,"](#) someone who is overindulgent or -lenient, or someone who is unmoved (if the injustice was aimed at others). An example is Martin Luther King's admonition about not only "the hateful words and actions of the bad people but... the appalling silence by the good people."^[2]

If you have an excess of these virtues, you might be called sanctimonious, superior, holier-than-thou, vindictive, blame-seeking, or judgmental. Critics of "callout culture," "outrage porn," and the like have noted pathological ways in which judgment and righteous anger can be harnessed to unvirtuous ends. When righteous anger is shared with others, it can feed on itself and result in a sort of lynch mob that becomes detached from clear judgment in pursuit of catharsis through scapegoating. This suggests that the virtues of judgment and righteous anger require discernment and [courage](#) (the courage to stand alone in your judgment, or to judge differently from those in your in-group).

The pros and cons of judgment and righteous anger

"I rather must confess, that I always suspect people that affect to cover all defects of others with the cloak of charity." —Baron Knigge^[3]

In a recent essay, philosopher [@AgnesCallard](#) argued that "Anger is a moral sense." In the same way that cold chills us, brightness makes us squint, and sweetness makes us salivate, injustice makes us angry. She resists "anger management"—the attempt by others to make us disbelieve our anger-senses and conform to theirs: "'If you don't stop being angry, you're irrational.' 'If you don't start being angry, you're immoral.' Neither of these speeches tends to go over well—at least not with me." She defends anger this way:

[A]t times it is only the angry who are in a position to apprehend the magnitude of some injustice. For they are the ones willing to sacrifice all their other concerns and interests so as to attend, with an almost divine focus, to some tear in the moral fabric. When I am really angry, it is not even clear to me that I can calm down—the eyes of the heart do not have eyelids—and the person making that request strikes me, to adapt a locution of Socrates', as trying to banish me from my property, the truth. They are calling me "irrational," but they seem not to see that there are *reasons to be angry*.^[4]

Anger as a warning signal

Righteous anger can be a way of drawing a line. You go from being tolerant to saying "this is no longer tolerable," and do so in an unmistakable way. Anger can be a form of raising your hackles in a way that deters: it says "don't push me," "back off," "you're going too far." It can be a good complement to tolerance by setting a limit: "I'm a tolerant person, but do not try to take advantage of that by walking all over me." The abrupt change in body language and vocal tone that (often) accompanies anger signals that a threshold has been crossed.



The character Howard Beale in the movie *Network* struck a nerve by inviting television viewers to move from tolerance to anger.

Such a signal works best when it is rare. A person who regularly uses explosions of anger to try to change other people's behavior just gets a reputation as an angry person, a hot-head, someone who "is always on about something." Their anger has a "[crying wolf](#)" quality to it, which makes it less effective when it's most needed.

The trouble with quibbles

"The surest way to work up a crusade in favor of some good cause is to promise people they will have a chance of maltreating someone. To be able to destroy with good conscience, to be able to behave badly and call your bad behavior 'righteous indignation'—this is the height of psychological luxury, the most delicious of moral treats." —Aldous Huxley^[5]

The "divine focus" that Callard says anger provides can also be a dangerous sort of tunnel vision. She notes that "the more perfectly one attends to the gravity of the wrongs done, the less sensitive one becomes to the gravity of the wrongs one is poised to commit in response."^[4]

Anger can exacerbate some cognitive biases, such as the [fundamental attribution error](#) (in which we attribute something we are angry about to the rotten *character* of whoever is to blame). On the other hand, there is [some evidence](#) that anger can motivate you to investigate the cause of your anger *more* critically.

Judgment as a way of backing your values

If you are too reluctant to judge — for instance because you overvalue tolerance, don't like to rock the boat, or don't want to draw attention to your own shortcomings — this can lead to a decay in your values, and, if this becomes a common practice, to culture-wide corruption.

"Can [love](#) be other than exacting, or loyalty refrain from admonition?" asked Confucius.
^[6] Judgment does not have to be hostile. It can be a loving thing, if done skillfully and with

good intentions. For example: “That racist joke you told, it was beneath you. Don’t be that way.” That judges the act as a poor one, but judges the actor perhaps even above their own self-regard.

Montaigne insightfully struck back against the idea that you should not judge anyone else until you have become morally pure yourself, when he said: “To censure my own faults in some other person seems to me no more incongruous than to censure, as I often do, another’s in myself. They must be denounced everywhere, and be allowed no place of sanctuary.”

And Alexander Solzhenitsyn decried the practice of pardoning corruption in high places:

When we neither punish nor reproach evildoers, we are not simply protecting their trivial old age, we are thereby ripping the foundations of justice from beneath new generations.... Young people are acquiring the conviction that foul deeds are never punished on earth, that they always bring prosperity. ¶ It is going to be uncomfortable, horrible, to live in such a country!^[2]

Summary

It seems there are a number of virtues associated with anger and judgment. In summary: Do not get angered too easily, or at the wrong things. When you do get angry, do so in a measured way that does not lead you to do something regretful or unwise. But blow your top if you can do so in a way that helps you focus your efforts on a righteous cause that defends your values, and when you do, do so in a way that is legible to others, particularly the target of your anger. Make anger temporary — do not seethe. Transition from anger to rational action and when possible to forgiveness.

1. ^

[Luke 17:3](#)

2. ^

Martin Luther King, Jr. “Letter from Birmingham Jail” (1963)

3. ^

Adolf Freiherr von Knigge, *Practical Philosophy of Social Life* (1788), Chapter I, Section XXVIII

4. ^

Agnes Callard, [“Anger Management”](#) *The Point* (27 January 2021)

5. ^

Frequently attributed to him, anyway, though I have had a hard time tracking down the source. It is often alleged to be *Crome Yellow*, but I could not find it there.

6. ^

Analects of Confucius, XIV.viii

7. ^

Alexander Solzhenitsyn, *The Gulag Archipelago* (1974)

Notes on Gratitude

This post examines the virtues of **appreciation**, **gratitude**, and **reciprocity**. It is meant mostly as an exploration of what other people have learned about these virtues, rather than as me expressing my own opinions about them, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about these virtues and how to nurture them.

What are these virtues?

The sort of “appreciation” I want to explore here is appreciation-of-others specifically (not, for example, aesthetic appreciation). When you appreciate someone for something they have done that benefits you, we sometimes call this sort of appreciation “gratitude.” However, you can appreciate someone without necessarily feeling gratitude (for example, you might appreciate someone’s sense of humor but not think that it makes sense to thank them for it).

In practice, there’s a lot of overlap between gratitude and appreciation, and the border between them isn’t very well-defined. In this post, I’ll use one or the other of “appreciation” or “gratitude” if the distinction seems to matter, and I’ll use either of them if it doesn’t seem to matter. If that’s confusing, leave a comment and I’ll try to clarify.

When we return a favor, we are doing a specific sort of gratitude-adjacent action that goes by the name “reciprocity.”

To rise to the level of *virtues*, appreciation, gratitude, and reciprocity should be habitual: in other words, you have the virtue of appreciation if it is characteristic for you to notice opportunities for appreciation, to then feel appreciative, and to follow through by skillfully performing appreciative actions.

Related virtues & vices

There are some other virtues that have to do with recognizing others, including respect, remembrance, honor, consideration, recognition, and solidarity. The ability to notice opportunities for appreciation requires attention. The skill of “savoring” can also be part of a good appreciative sense.

The skill of *accepting* appreciation or gratitude gracefully (rather than with deflection or false-modesty for example) is also useful. For one thing, it is more difficult to express appreciation or gratitude if you see the receipt of appreciation or gratitude as an occasion for embarrassment or awkwardness. For another, if you are graceful in the way that you accept gratitude, people will be more likely to model it for you and this will help you learn how to express it well.

A culture may have established rituals of gratitude (e.g. the “thank you note,” tipping), and learning to competently perform these rituals is part of the virtue of courtesy.

People informally use the word “gratitude” in both a propositional and a prepositional sense. That is to say, we sometimes feel gratitude about something, and sometimes feel gratitude towards someone for something. There’s some debate about whether the first variety is really “gratitude” or whether it’s something else (e.g. gladness).

Some vices that interfere with appreciation include narcissism (which can prevent you from noticing others enough to appreciate them), hubris (which can keep you from feeling gratitude because you assume your good fortune was inevitable), entitlement (which can make you believe that gratitude is superfluous), cynicism (which can make you assume ulterior motives are behind other people’s helpful actions), and resentment (which can make you unwilling to acknowledge having received a favor).

Spite (or maybe righteous anger) is, in a way, a sort of dark-complement to gratitude: instead of repaying [kindness](#) with thankfulness, it replies to injury with bitterness. This may suggest that [forgiveness \(or mercy\)](#) is a sort of diagonal-complement to gratitude.

What does gratitude consist of?

Gratitude is sometimes described as a feeling of appreciation that may or may not end up expressed by some sort of appropriate communication or action in response. Other times it is described in a way that includes both the feeling and the action that feeling provokes, or in a way that suggests that the gratitude is more complete or more sincere if it culminates in some sort of act of gratitude that is legible to the person who is deserving of appreciation.

I've seen a couple of attempts to break gratitude down into its component parts. One comes from the [Jubilee Center for Character and Virtues](#) ([here's a brief video in which Dr. Blaire Morgan explains it](#)). In their model, gratitude consists of (1) conceptions/understandings about gratitude, (2) grateful emotions, (3) attitudes about gratefulness (e.g. that it is important, worthy), and (4) gratitude-related behaviors. In the research associated with this model, these four components can be independently measured, and there is a positive relationship between your well-being and how many components you perform relatively well in (i.e. if you're above-average in more components, you will feel more well-being).

Another four-factor model comes from researchers working with the [Expanding the Science and Practice of Gratitude Project](#), who were focused on the development of gratitude in children. The way this was first formulated, you (1) notice that you have received something, (2) realize that this benefited you, (3) notice that the giver acted intentionally to bestow this benefit, and (4) do something to show appreciation. This has since been refined into a framework that now goes by the label “notice-think-feel-do”: (1) notice something for which you have reason to be grateful, (2) think about how it is that this came about, (3) feel the positive emotions that follow from this, (4) do something to show your appreciation.*

When people decide whether gratitude is warranted, part of this assessment involves evaluating the motives of the person who granted the appreciated favor or help. People are less likely to feel grateful to someone whose otherwise praiseworthy act was motivated by ulterior self-seeking motives or was done for pay or from duty. People also typically judge favors not absolutely but relative to the favors they expected to receive or to those given them by others.

For these reasons, while you would be being polite to thank the grocery store employee who bags your groceries, you would probably be considered to be going a little overboard by adding “that was so kind of you!” and you would likely be considered weirdly eccentric were you to write them a thank-you card. On the other hand, there’s no rule that says you cannot or should not be appreciative to people whose kindness to you is also being paid for (by you or someone else), and given how win-win gratitude seems to be, there may be good reason to overcome this bias and err instead on the side of being eccentrically-grateful.

Where did gratitude come from?

Adam Smith, in *The Theory of Moral Sentiments*, wrote that human moral emotions like gratitude and resentment are purposeful in that they help us regulate our own social behavior and the behavior of others around us in ways that benefit us.^[1] In his scheme, we are grateful to someone to reward them for doing something beneficial for us because we want to encourage them to do such things again in the future; and we appreciate expressions of gratitude because they indicate that the recipient of our favors acknowledges them as such and recognizes an obligation to reciprocate.

Smith was writing before the theory of evolution through natural selection was developed, but his speculations closely resemble those that are now put forward about how the emotions and perceptions surrounding gratitude evolved to help species like ours regulate [reciprocal altruism](#).

The ability to [trust](#) others to reciprocate reliably, and the ability to discern who is and isn’t trustable in this regard, is key to the economic behavior that enables our species to be so world-transformative.*

Is gratitude instinctive or learned? It does seem like parents have to work hard to teach children gratitude and how to properly express it. But maybe they’re just trying to rush something that would come naturally eventually, or maybe the gratitude itself is instinctive but the ways to express it legibly are culturally-specific and have to be learned.

What is gratitude good for?

“In all things we should try to make ourselves be as grateful as possible. For gratitude is a good thing for ourselves, in a manner in which justice, commonly held to belong to others, is not. Gratitude pays itself back in large measure.”

—Seneca^[2]

Gratitude is a social virtue, valuable for the way it helps to strengthen social bonds and to encourage pro-social behavior. It has also been getting increasing attention in recent years for its benefits to the individual who expresses it, in the forms of increased happiness, enhanced life satisfaction, and even improved health.

While gratitude is usually thought of in a positive light, it sometimes has a less-positive shadow. For example, people who rely on others (e.g. people with severe disabilities) may find gratitude burdensome, or that it makes them feel like a burden in the way it emphasizes the imbalanced nature of their dependence relationships. Gratitude can make you feel beholden on someone who has done you kindness or

given you gift, and sometimes people take advantage of this (it's a sales technique, for instance, to give a small gift as a foot-in-the-door). Gratitude sometimes has implications of indebtedness or obligation (and this in some cultures more than others), and some people try to avoid situations in which they would feel gratitude for this reason.*

Social benefits of gratitude

"The art of acceptance is the art of making someone who has just done you a small favor wish that he might have done you a greater one." —Russell Lynes^[3]

Gratitude helps you form, maintain, and improve your social circle. By expressing appreciation to someone, you make that person feel better and feel better-disposed towards you. By expressing appreciation for things you value, you reward people who do things that you value and reinforce those things. Expressing appreciation also models the behavior of expressing appreciation for others, and so can multiply its effects in a way which redounds on you and your circle (this is sometimes called "upstream reciprocity").

If you can't think of anything to appreciate about the people you're currently running with, that can be a good sign that you need to start running with a different crowd. So in this way, keen perceptions associated with appreciation can be valuable even in the absence of much that's worth appreciating.

Gratitude researchers have come to describe the social functions of gratitude using the "find, remind, and bind" model.* In this model, being well-attuned to gratitude helps you find people who are worth spending time with, reminds you of the value and importance of your relationships, and binds you to such people through behaviors that help to maintain those relationships.

Personal benefits of feeling & expressing gratitude

"Benevolence gladdens constantly the grateful; the ungrateful, however, but once." —Seneca^[4]

There has been a flood of research over the last twenty years about possible personal benefits of feeling and expressing gratitude. This includes improvements in subjective happiness, subjective well-being, and objective health measures.

In a typical experiment, a group of people will be divided into one subgroup that performs a gratitude-boosting exercise of some sort (e.g. listing things they're grateful for), while a second group does some similar exercise that does not have a gratitude component (e.g. listing childhood memories or recent "hassles"). The subjects will be measured in some way before and after the exercises to see if any effect can be noticed on their health or subjective well-being (hedonic or eudaimonic). Sometimes also the subjects' practices and attitudes of gratitude are themselves measured before and after the experiment to see if the experiment makes a person more apt to feel or express gratitude. Occasionally experiments will include long-term follow-ups as well.

Much of this research may suffer from some of the weaknesses that have plagued social science and psychological research in recent years. It's also dominated by research subjects from [WEIRD](#) cultures. I don't feel confident about trying to distinguish the vigorous from the hopelessly unreplicable myself. If you want to delve further, the Greater Good Science Center white paper "[The Science of Gratitude](#)" looks to be a good overview.

Some of the personal benefits of gratitude may be social benefits in disguise. If your gratitude helps to strengthen your social network, for example, you may feel more able to ask your friends for help, and this might improve your well-being. Or if you have high regard for other people, this might include both expressing gratitude toward others and respecting the advice of your doctor, which can improve your health outcomes.

But it seems intuitively sensible that gratitude might directly improve your subjective well-being. For one thing, gratitude concentrates your mind on the things in your life that you like, value, and appreciate. In that way it helps you to enjoy them all over again and adds to their benefit. In one study, people who were assigned to write about their "intensely positive experiences" for 20 minutes showed measurable positive changes in both mood and physical health thereafter, compared to a control group that wrote for 20 minutes a day on some neutral topic. [\[5\]](#) And this was without "gratitude" being an explicit part of the process.

This also helps to remind you of the things you find most valuable, memorable, and enjoyable, which can help you align your life with the pursuit of those things. (For example, people often spend a lot of resources on *stuff* but are more apt to appreciate and reminisce about *activities*. Some people interpret this as a clue that we would be wiser to reallocate our resources toward the pursuit of valuable activities.)

Appreciation is a way of short-circuiting the [hedonic treadmill](#) (people tend to quickly get used to improvements in our lives such that we take them for granted and they no longer make us happy).

The superstitious or religious feeling that good fortune is a personalized grant of the gods might be a way of sprinkling a little extra sugar on an already sweet situation. If you think your good fortune is deserved, or just random, you can feel blasé about it. But if you look at your good fortune as something that was granted to you specially, you get the warm fuzzies from being favored by benevolence.

Anthropomorphizing fate, or having a God to thank for everything, may allow you to take advantage of the positive aspects of gratitude in cases where it otherwise wouldn't make sense. A friend of mine suggested that as a (perhaps *the*) sentient species capable of feeling appreciation for the marvelous, wonderful miracle that is life, the universe, and everything, such gratitude gives us the purpose we long for: Perhaps the point of human life is to appreciate and applaud this bizarre cornucopia of astonishment and sensation.

How to improve in this virtue

In a welcome contrast to many of the other virtues I have examined, there is a wealth of advice on how one can become more appreciative, and feel or show more gratitude.

General advice and notes

It may take a deliberate, conscious act of attention to become aware of things we can be grateful for. There is a cognitive bias that has been labeled “headwind/tailwind asymmetry”^{*} in which we take more notice of challenges we have faced or overcome than we do of privileges or benefits we have taken advantage of. That is to say: “headwinds are far more salient than tailwinds.”

For any specific gratitude-boosting practice or practices you choose, if you want to develop the habit of doing that practice regularly, it can be helpful to choose some trigger to prompt the practice and thereby establish the habit. One common trigger for gratitude-expression is the evening meal. This is probably an outgrowth of the Christian tradition of [“saying grace”](#) — giving a prayer of thanks — at the commencement of a meal. Habit guru James Clear, for example, says: “When I sit down to eat dinner, I say one thing that I am grateful for happening today.”^[6]

Specific practices

“Don’t set your mind on things you don’t possess as if they were yours, but count the blessings you actually possess and think how much you would desire them if they weren’t already yours.” —Marcus Aurelius^[7]

The scientific literature on gratitude has promoted a handful of “interventions” that are meant to prompt or boost gratitude. On the one hand, these have the benefit of having been subject to scientific scrutiny and so in theory have foreseeable, measurable results. On the other hand, they are designed for ease of use in an experimental setting and so are usually brief and simple. More complex or demanding exercises that might also be more rewarding may have been overlooked because of how difficult it would be to subject a bunch of cheaply-obtained research subjects to them.

That said, these are some of the exercises that have been most experimentally investigated:

- **counting blessings** (write down five things worth being thankful for, daily or weekly; “saying grace,” as mentioned earlier, is a vocal version of this)
- **three good things** (write down three things that went well for you, and their causes)^{*}
- **mental subtraction** (imagine & maybe write about what it would be like if some positive event had not happened — see the Marcus Aurelius quote above)
 - “To feel grateful for some of these things you might have to try to vividly imagine being without them for a time. If you are deprived of some of these things for a time (or temporarily believe you are) you can also try to remember what that feels like, so that you can recapture it later when you have them again.” —[MaxCarpendale](#)
- **gratitude letters/visits** (write & deliver a thoughtful thank-you to a person you had not properly thanked before)^{*}
- **death reflection** (consider your own death in a manner similar to the “mental subtraction” exercise)
- **experiential consumption** (spend money on an experience, rather than things)

Dan Weinand's post [Gratitude: Data and Anecdote](#) from last month looks at a few of those exercises a little more closely and describes his own experiences. Other exercises I have seen recommended include:

- [Keep a gratitude journal](#) in which you regularly write about things worth being grateful for.*
 - [And don't neglect the big things](#) like "being alive at all, being alive at this time in history, having loved ones who are alive, being born a human, having functional limbs..."
 - One set of researchers claim their research shows that the ideal frequency for such journaling is about once per week.^[8]
- Every day, at least once, express genuine, specific appreciation to somebody — either someone you encounter during the course of the day, or, if such an occasion doesn't arise, by sending a letter or email or making a phone call to someone who you appreciate.
- Make your gratitude list more of a bulletin board, in a place where you see it and notice it and so are reminded of the things to be thankful for regularly.
- For practice, find ways to express appreciation to yourself ("Thank you, past me, for doing the dishes so I could wake up to a clean coffee mug this morning.")

You can show gratitude and appreciation by expressing it directly to the person who you appreciate or are thankful to, of course, but you can also do so by expressing that same gratitude to third parties, e.g. "I'm willing to pay more for X's products because I really appreciate her craftsmanship," or "I was really grateful to Y for his help." This has a somewhat different set of benefits, but benefits nonetheless: It also reminds you of positive things and admirable people in your life, and it helps to seed the social conversation with your opinions of what sorts of behavior are worthy of appreciation.

[Ben Franklin invented a curious hack of the human gratitude response](#) that seems worth mentioning here: He would jump-start a process of reciprocal gratitude in someone who was otherwise not well-disposed to him by asking that person for a very small favor (in his example, asking a colleague if he could borrow a book). He then was careful to skillfully express gratitude for that favor, and then found that the person who had granted him the favor became friendly towards him.

What makes people feel appreciated?

For gratitude or appreciation to work at its best, it must be expressed in a way that is legible by the receiver as intended. I asked several friends what makes them feel appreciated, and also looked for advice and hints in the literature I reviewed. Here is some of the advice I found:

- People feel appreciated when their contributions are enjoyed (e.g. when food they serve or bring is eaten with gusto).
- Hugs and similar relationship-affirming body contact (hand on the shoulder, handshake) can show appreciation.
- People feel appreciated when they are clearly included or explicitly invited.
- Showing people [respect](#) is also a way of showing appreciation. For example, you show you appreciate someone when you follow-through on your promises to them.
- Thank-yous (especially "genuine and un-asked-for" ones) show appreciation. If you never established the common-courtesy habit of offering thank-yous in

return for small kindnesses, that may be a habit worth establishing. That said, phony or rote rituals of appreciation aren't worth a whole lot either to the giver or receiver. It can be worth your time to put in the attention and energy it takes to be more genuine. Alas, being awkward or unpracticed can feel similar to being inauthentic, so that can make this a difficult habit to establish: you have to push past the awkward stage. In my experience, even an awkward expression of gratitude is usually positively received, so if you can get past the uncomfortable feeling of awkwardness, you can probably still get benefits from expressions of gratitude before they become more graceful from practice.

- Reciprocation is one way of showing appreciation: You scratch my back, I'll scratch yours. There's an art to getting this right. On the one hand, it's a good idea to go out of our way to show kindness to those who have done us favors. On the other hand, being quick to return a favor can ironically be interpreted as a sign of being unfriendly and of thinking of the relationship in mercenary terms. Doing this well takes some finesse.
- Simply saying to someone straightforwardly that you noticed they have done you a favor or done something exceptional and that you appreciate it, or writing them a note to that effect, can be enough. This usually works better if you appreciate something specific and extraordinary rather than something vague or generic, though a skillfully-worded "I appreciate you for who you are" or "for what you bring to my life" can work in a pinch.
- If you can disagree with someone but do so in a way that does not attack or belittle them, this shows that you appreciate them in spite of your disagreement.
- Two people mentioned "second-hand referrals" — which I interpreted as vouching for them or endorsing them to others. One said: "It's evidence that someone I taught or helped out appreciated the result enough to convince a stranger to approach me with a new need to be fulfilled."
- If you can learn what a person is proud of or feels they ought to be appreciated for, and explicitly appreciate them for that, you get a gold star. This doesn't necessarily have to be something you think is extraordinarily admirable, but can be something like their taste in clothes, their new haircut, or their cool wheels.
- It shows someone you appreciate them if you remember their special day (e.g. a birthday).

My experience with trying to develop this virtue

"The golden moments in the stream of life rush past us and we see nothing but sand; the angels come to visit us, and we only know them when they are gone." — George Eliot^[9]

I'm doing a virtue-strengthening program on the buddy system with a friend. At the end of last year, in a sort of New Year's Resolution, and feeling like I had a previously-adopted habit well-established, I decided to pick a new virtue to start 2021 with. I chose appreciation/gratitude, which was one my buddy had already been working on.

I chose it in part because of my buddy's good experiences, in part because of the wealth of evidence in its favor, and in part because it's something I don't feel very good at so there seemed to be a lot of room for improvement. I've long felt awkward around showing appreciation for people. I can "thank you" just fine, but when it comes

to a sincere, look-someone-in-the-eye, specific expression of appreciation, I've typically come to a stuttering stop before I even get started.

I think part of this is that when such shows of appreciation were modeled for me as a child, in the form of an adult showing appreciation to a child, they struck me as condescending — as something that highlighted the adult/child superior/inferior relationship. So now, when I think of showing appreciation to someone, I worry that I'd be putting on airs or presuming to be in a position of authority over them.

There's also the awkwardness factor. Not being well-practiced in how to show appreciation, it sometimes doesn't seem to come out right when I try. But I've so far found that the down-side to awkward attempts at appreciation is pretty minor, while the up-side to typical attempts is pretty great, so that's helped me to stick with it.

The technique I've chosen so far is to express explicit appreciation to some individual person, for something specific, at least once a day. Given our pandemically socially-distant times, this has sometimes meant writing an email rather than talking with someone face-to-face, but that's fine. I usually have expressed my appreciation to someone I personally know, but on a couple of occasions I have sent a note to someone I don't know personally but whose on-line generosity I appreciate. I also check in with my virtue-buddy a few times each week, which helps keep me accountable in establishing this new habit.

My impression of this exercise so far is that I'm a fool not to have started this earlier. By starting the day with a mission of identifying things I appreciate about the people around me, I have become more aware of positive things and admirable people in my life. By giving people positive feedback for things I appreciate, I help to encourage more of those things, and (assuming I have good taste) thereby encourage them to be more delightful in general. This also comes back to me: I pay more attention to behavior I admire and so am more likely to learn how to exhibit that behavior myself.

One thing I would recommend to people considering such an exercise is not to put all of your focus on the goal of *expressing* appreciation, but to attend also to improving your awareness to occasions for appreciation. If you rush to the expression of appreciation, you may try to force it in a way that comes off as insincere. However if you first notice and appreciate something, your expression of appreciation can follow more naturally. My virtue-buddy put it this way in her advice to me:

Slow down. Life comes at us so fast and we deal with what we can quickly. Instead of looking at life as a racquetball coming at you that you gotta swing at, look at it as a game on TV, slowing it down and taking it in. You can *notice* and become *aware*. Listen to your inner voice. It's commenting on people, on things, on you, and a lot of events in Life. In its quiet whispers, it says things that you appreciate in Life. Give those whispers a form in this world that everyone can hear.

1. ^

Adam Smith, *The Theory of Moral Sentiments* (1759)

2. ^

Lucius Annaeus Seneca (the Younger), *Moral Letters* 81.19

3. ^

Russell Lynes, *Life in the Slow Lane* (1991, reproducing his quote from *Reader's Digest*, December 1954)

4. ^

Lucius Annaeus Seneca (the Younger), *De Beneficiis* book III, section XVII

5. ^

C.M. Burton & L.A. King "The health benefits of writing about intensely positive experiences" *Journal of Research in Personality* (2004) pp. 150-163

6. ^

James Clear, ["Use This Simple Daily Habit to Add More Gratitude to Your Life"](#)

7. ^

Marcus Aurelius, *Meditations* VII.27

8. ^

S. Lyubomirsky, K.M. Sheldon, & D. Schkade "Pursuing happiness: The architecture of sustainable change" *Review of General Psychology* (2005) pp. 111-131

9. ^

George Eliot, *Scenes of Clerical Life* (1858)

Notes on Shame

This post examines the virtue of **shame**. It is meant mostly as an exploration of what other people have learned about this virtue, rather than as me expressing my own opinions about it, though I've been selective about what I found interesting or credible, according to my own inclinations. I wrote this not as an expert on the topic, but as someone who wants to learn more about it. I hope it will be helpful to people who want to know more about this virtue and how to nurture it.

"He who feels no shame of evil and does not hate it is no man. Shame and hate of evil are the beginning of virtue." —Mencius^[1]

"Where there is yet shame, there may in time be virtue." —Samuel Johnson^[2]

What is this virtue?

There's a terrible terminological muddle around shame. I'm going to use "shame" to mean an unpleasant sense that one has failed to live up to one's own standards in some way. To have a well-tuned virtue of shame (or sense-of-shame) is for this sense to reliably and usefully alert to appropriate things.

Arguably, this virtue might also include *responding* to this sense well: how you process shame, learn from it, dispose of it properly, and so forth.

Shame vs. guilt

Shame overlaps with guilt and sometimes "guilt" is overloaded to include shame among its meanings. I think "shame" is a better word for precisely describing the virtue. For one thing, you can be ashamed of something (e.g. not living up to your potential) without necessarily being guilty of some specific transgression.^[3]

Also, you can judge someone else to be guilty, or they may just objectively *be* guilty based on the facts of the matter—whereas shame is more of an introspective, subjective evaluation. It's true that you can try to *shame* someone, but for this to succeed it requires their cooperation: they must acknowledge and internalize the shame by becoming ashamed, or the attempt sputters out ineffectually. This is why you can say simply "you are guilty" but shaming takes a more complex construction: "you should be ashamed of yourself."

These linguistic categories are blurry, though: Ashamed people sometimes say they "feel guilty." Some people restrict "shame" to refer to an I-am-bad sort of shame, and use "guilt" for I-did-bad shame about specific discrete acts. Just to keep things confusing, I-am-bad shame is sometimes called "free-floating guilt," and at least one researcher insists that "shame" refers to feelings regarding violating social norms and "guilt" to those about violating personal standards.

In short, it's frustratingly easy for people who discuss shame and guilt to talk past each other, and it takes a lot of effort to decipher the literature about shame because people use the same words to mean many different, sometimes subtly different, things.

Shame vs. regret or remorse or embarrassment or humiliation

When people shame you, and you become ashamed in response, this feels similar to *embarrassment* or *humiliation*: the awful feeling of being held up to ridicule, scorn, rejection, disgust, and things like that.

One possible way to distinguish shame from embarrassment/humiliation is that the latter usually implies external judgement: you are embarrassed or humiliated because of how you appear to someone else. In contrast, you can feel shame over something nobody else is aware of.

But how do you distinguish shame from *remorse* or *regret* — also private negative feelings about something you have done?

While I was working on this note, I took a break to walk to the grocery store. Along the way I tried to recall an incident from my past that I'm ashamed of. It happened about twenty years ago, and although it was nothing tremendously awful, it did reveal a selfish, covetous, ungrateful side of me that even now makes me cringe a bit to think about. After getting home from the store, I inhaled a pint of ice cream. That was probably a mistake. I have reason to regret it. I would be embarrassed to mention it to my physician. But I'm not *ashamed* of it: I think of it as a foolishness but not a *shame*.

Maybe this is how shame and regret can be distinguished: When I think back on something I'm ashamed of, it seems to involve learning something unpleasant about my character — I thought I was a brave person, but I chickened out when it counted, or, I thought I was a generous person, but I was stingy at a time when generosity was called for. Mere regrets don't seem to have this component. I can regret a mistake I made, or a [decision](#) that seemed like a good idea at the time but turned out poorly, but when I regret, while I may notice that my wisdom was underdeveloped or that I could have been more on-the-ball about something, I don't notice some fundamental character flaw revealing itself. Things I'm ashamed of, on the other hand, hit deeper at the heart of me. There is more urgency to shame. When I think of something I'm ashamed of having done, this thought is accompanied by something more like "I must never do that again," or "how could I have done that?" or "I thought I was better than that". When I do something I merely regret, it's less emphatic: "jeez; I ate the whole carton of ice cream? that was dumb."

There's more hand-wavey vagueness in this definition than I'd like, though. Which foibles touch on "character" or "hit deeper at the heart of me" and how do I know? Does my sense of shame *tell me* which those are, or do I learn which those are some other way and my shame response learns this in parallel and adjusts accordingly? I feel as though I have developed a more refined and more exacting standard of character as I have matured, and that I have done this in a more-or-less deliberate, [rational](#), conscious way. When I am now ashamed of things that I did twenty years ago that I apparently was not sufficiently ashamed of doing when I did them, I feel as though my conscious shaping of my character has also shaped my more emotional sense of shame, not the other way around. But maybe I could tell a plausible story in which my shame evolves subconsciously, and I consciously concoct a description of my character to match its contours as they emerge.

Shame as a “quasi-virtue”

"To be so constituted as to feel disgraced if one does such an action, and for this reason to think oneself good, is absurd; for it is for voluntary actions that shame is felt, and the good man will never voluntarily do bad actions." —Aristotle^[4]

[Aristotle considered shame to be a "quasi-virtue."](#) For one thing, it seems closer to an involuntary emotion than a voluntary characteristic: for example, as with anger or fear, shame is often accompanied by involuntary physical symptoms (like blushing). But also, if you feel shame from a well-tuned sense of shame, it's probably because you've done something anti-virtuous, so if you *characteristically* or *habitually* exhibit shame — that is, if shame were like the other virtues in that regard — it would not be evidence that you are living a flourishing life of [eudaimonia](#) but that, on the contrary, you're screwing up an awful lot. But on the other hand, if you are incapable of feeling shame — if you're *shameless* — you are in need of a virtue that you lack.

Aristotle for this reason draws a parallel between shame and [self-control](#): If a person desires to do what is wrong, but then has the self-control to refrain from doing it, that's better than if they lacked self-control — but it would be better yet if that person were [temperate](#) enough not to have that bad desire in the first place. Similarly if a person does what is wrong, but then has the sense to be ashamed of it, that's better than being shameless — but it would be better yet if they hadn't done the wrong thing in the first place.

If we think of this virtue as a "sense of shame" rather than "shame" plain and simple, we can avoid at least some of Aristotle's awkwardness. If you habitually, characteristically, skillfully call yourself to account and weigh yourself in the balance of your high standards, then you can be said to have the virtue of a-sense-of-shame.

Essential shame

Sometimes shame seems to attach itself to you so strongly that it becomes part of who you are, rather than being something temporary that you work through and set aside. There are a number of perspectives on this sort of shame.

Often, this sort of shame is discussed as a sort of pathology or mistake: as "toxic" shame. For example, you are told that you are unworthy or terrible as a child, incorporate that into your self-image, and live your life under that awful shadow. Or powerful voices in culture decide that something about you (such as your sexuality or your race or your appearance) is an ineradicable flaw or crime against nature. In such cases you will be encouraged, in your favorite sources of pop-psych advice, to overcome this shame or heal it or free yourself from it by means of therapy or some ritual or other.

[Original sin](#) is a sort of essential shame that plays an important role in Christianity. Shame first appeared when Adam & Eve defied God to eat the fruit of the tree of knowledge, and thereby became aware of good and evil. The first symptom of their newly-won knowledge was shame. Sin — shame-worthy thoughts, desires, and actions — was henceforth baked into human nature. Acknowledgement of this sin (and perhaps formal confession), humble shame before God, repentance, and begging for divine mercy, are important parts of Christian practice.

Essential shame may result from processing ordinary shame poorly. Rather than using shame productively to retune your decision-making process and to make necessary amends, in such a way that the shame does its work and goes away, you convert the

shame into essential shame and dye it into the fabric of who you are. This may be because you never learned how to go through the process well, or it may be because you have a lot of essential shame already and so adding a little more shame to that pile is easier than trying to work through it.

Sometimes, essential shame is thought to attach to a person if their transgression is sufficiently serious or irreparable. If you cannot possibly make amends, perhaps you therefore cannot resolve your guilt and you are doomed to carry the shame of it with you always. You now have shame, like an injured person might forever have a scar or a limp.

German philosopher [Karl Jaspers](#) thought, for example, that [Good Germans](#) who had permitted and enabled the Nazi atrocities had thereby acquired a sort of permanent guilt that must shape the rest of their lives into making amends: “Our life remains permitted only to be consumed by a task,” he wrote.^[5] [Theodor Adorno](#) seemed to go further, implying that such crimes had covered everyone with shame, when he wrote that “after Auschwitz, to write a poem is barbaric.”^[6]

What good could it do to feel shame over something you did not personally participate in? Isn’t “collective guilt” a pernicious myth? If you demoralize shame and think about it more functionally — what it *does*, what it’s *for* — this sort of shame makes more sense. Shame over a personal misdeed helps you notice the disconnect between the sort of person you thought you were and the sort of person you turned out to be, and to course-correct accordingly. Shame over humanity’s evils might be similarly helpful. Perhaps you once had illusions about the dignity of humanity and thought that there was some modest limit below which you could not possibly sink. When you learn the depths to which humanity can fall, a feeling akin to shame might help you to adjust and refortify your defenses against participating in atrocity.

Weaponized shame

As already mentioned, it is a common human pattern to try to cast shame on someone else (“shame on you” is the paradigmatic form of this spell). Rather than causing them pain directly, you try to provoke their sense of shame so that they cause themselves pain.

You can also do this indirectly and anonymously and remotely. If you can contribute to a shared cultural understanding that *X* is shameful, then you can disincentivize and/or morally subordinate people to whom the tag *X* can be stuck. So-called [“slut-shaming”](#) is one example.

[Online shaming](#) mobs are often repulsive and can be shameful in their own way, but they are arguably very effective in establishing and policing certain (sensible or not) boundaries of behavior. (Jon Ronson’s usual sharp wit and insight is on good display in his examination of this phenomenon in his book [So You’ve Been Publicly Shamed.](#))

Since shame can be an effective goad, advertisers sometimes use it to sell products. What kind of a mother would deny her child Producto™? Your psoriasis is not only treatable with Miraclon-X™, but you ought to consider yourself a pariah for not having already figured that out.

I Wept on my First Wedding Anniversary



*When Henry slipped lovely gift ring
on my finger I sobbed: "It's too pretty
for my red, dishwashy hands"*

Shame as entertainment

Other people's shame can be entertaining. This helps fuel the outrage mobs of online shaming. It also forms a lot of the appeal of television like [Judge Judy](#), [Jerry Springer](#), "reality TV," and the "scandals," "gaffes," and such that make up a disproportionate amount of political reporting. Much tabloid-style press takes the form: celebrity caught doing something shameful.

People who deliberately act or present themselves in what would ordinarily be a shameful way for entertainment purposes are usually categorized as clowns of one sort or another. [Sarah Silverman](#)'s hilarious comic persona, for example, guilelessly exposes her own shameful thoughts and impulses again and again ("I don't care if you think I'm racist. I just want you to think I'm thin.").

Why is shame entertaining? I don't know, but if I had to cough up a [just-so story](#), it'd go something like this: We like to watch other people's shame because it teaches us where the boundaries are without us having to cross them ourselves and meet the consequences. I bet it's a hoot and a holler to the other cows when one of their fellows runs into the electric fence.

Related virtues and vices

A sense of shame helps you hold yourself to high standards, and so is kin to virtues like [honor](#), propriety, rectitude, and upstandingness. Shame can be thought of as an inward-pointing version of [judgment or righteous anger](#), one that preserves those virtues from being hypocritical. Shame is related to [humility](#) in that in order to be ashamed of something you have to be willing to acknowledge that you screwed up: If you are unwilling to permit your self-image to be deflated, you will turn a blind eye to shame.

There is some obvious tension between the virtue of shame and the virtue of pride. For example, one way to defy *being shamed* is to assert your pride (e.g. "gay pride"). On the other hand, it is *because* you have pride — because you value your character — that self-inflicted damage to your character is painful: that you feel shame.

A well-tuned sense of shame coordinates well with the ability to accept fault gracefully and to apologize, to repent, and/or to atone for one's misdeeds (those things might be considered virtues of their own, and I hope eventually to do a write-up on them). Shame helps you accomplish the "[fearless and searching moral inventory](#)" that is important in 12-Step programs, Christian [reconciliation](#) with God, and other such processes of renewal.

Sometimes the role of shame is assigned to a faculty called "conscience" — less as a virtue and more as a sort of sixth sense. I was struck by [what I read in the Dalai Lama's book Beyond Religion: Ethics for a Whole World](#): He says that "conscience" — a mysterious mental faculty that acts as a sort of ethical lodestone — is not part of his philosophical/psychological heritage. In its place is a conscientiousness motivated by self-respect and by consideration of others' opinions: self-respect says "this misdeed is unworthy of me" and consideration-of-others says "and I'll be poorly thought-of for doing it." This seems to me to describe something very similar to how shame functions.

The vice of deficiency regarding shame is “shamelessness” or one of its synonyms. The vice of excess — people who are ashamed too easily, or in an affected way — might be “bashfulness.” People who are overly-ashamed are sometimes described as “debased” or “feeling worthless.” This can wrap around and become effectively as bad as shamelessness if the debased person thinks they are already so miserably awful that one more bad deed can’t possibly stain them any worse.

A chronic, “toxic shame” — for instance one learned as a child at the hands of emotionally abusive adults — is a miserable thing to carry around and can take some hard work to shed. It can become a sort of essential shame: a feeling not so much of *having done* wrong, but of *being essentially* wrong. An essential shame of this sort can prevent the development of a useful sense of shame. After all, if you think you *are* shameful, then if you *do* something shameful this does not conflict with your self-image and so you have no reason to feel more ashamed than usual.

What good is (a sense of) shame?

“We can remove most sins if we have a witness standing by as we are about to go wrong. The soul should have someone it can respect, by whose example it can make its inner sanctum more inviolable.” —Seneca^[7]

Shame seems to be like other useful though unpleasant feelings that cause aversion, like nausea or pain.* Shame notices that we’ve stepped in it, and acts through visceral negative feedback to discourage the sort of prior steps that led there.

Shame is unpleasant enough that the threat of it can be sufficient to deflect us from our path, so, if we take the time to deliberate about our future actions, shame can also act proactively to discourage and prevent misdeeds. (In this way, shame can help to improve temperance.)

A good sense of shame is evidence that you measure your real-world behavior against your values and ideals, and give yourself useful feedback by which you can make course corrections. Shame is one of the mechanisms by which we learn from our mistakes.

A well-tuned, reliable sense of shame can protect you from going along with immoral mob- or state-encouraged atrocities, which can often come packaged with excuses that can seduce the intellect. “It was not our minds that resisted but something inside our breasts,” [wrote Alexander Solzhenitsyn](#) of how shame protected him and people like him from the temptation of joining Stalin’s secret police. “People can shout at you from all sides: ‘You must!’ And your own head can be saying also: ‘You must!’ But inside your breast there is a sense of revulsion, repudiation: I don’t want to. *It makes me feel sick.* Do what you want without me; I want no part of it.”^[8]

Shame can be a sort of “alarm” that warns you that you might be called to account for something you have done (or are thinking of doing) and so you ought to get your story straight and prepare to defend yourself.* In this way, shame does not so much protect you from doing wrong as it helps you protect yourself from the consequences.

Shame can be a way of learning about values that you hold subconsciously. If you become ashamed of something that you more consciously or rationally do not believe

to be shameful, and you notice this disconnect, you can then try to [integrate yourself](#) better.

Shame encourages you to withdraw from society. If you have recently done something that exposes you to scorn, when people see you that will likely be the most prominent thing they will associate with you, which is probably not to your advantage. If you lay low for a while, you allow for fewer opportunities for people to be reminded of your misdeed, and to associate you with it, and in this way you can reduce the damage to your reputation.

Shame as signalling

Scott Alexander considered how legible displays of shame can signal to others that you are a good cooperation partner:

The conspicuous experience of guilt is an evolutionarily advantageous way of assuring potential trading partners that you will be punished for defection. The behaviors associated with guilt are costly signals that help differentiate false claims of guilt from the real thing and add to public verifiability of the punishment involved.*

He also suggested that signalling lies behind the phenomenon of “survivors’ guilt” in which people who survive — e.g. a battle, an atrocity, a disaster — when others have died will feel guilty about it. Were a survivor to feel too self-satisfied about their survival, this might provoke suspicion that they had been clever about shifting the risks onto the victims and were congratulating themselves. By instead feeling conspicuously bad about their good fortune, they make this seem less likely, but also preemptively punish themselves so nobody else feels any urgent need to wonder whether such punishment would be actually justified.

When one has done nothing wrong, it can sometimes be advantageous to paradoxically display guilt in order to prove one’s lack of wrongdoing.*

Shame can be a useful signal to broadcast when you have been caught doing something that is sanctionable. For example, criminal court judges in the United States will often adjust the sentences they pass down according to the amount of remorse they feel has been demonstrated by the convict. A convict who accepts guilt and demonstrates remorse can reasonably expect this to pay off in the form of a more lenient sentence. In this way, self-induced shame can be a sort of culturally-approved proxy for other-inflicted punishment.

If you signal that you can successfully be shamed, this may broadcast that you are on board with the current cultural consensus about correct behavior: that you can be counted on to behave according to the current fashion, or to get back in line should you happen to deviate from it. This may make you more predictable to others and therefore make it easier for you to join in cooperative arrangements.

Conversely, if you publicly demonstrate scorn for community consensus about shame, this can bolster your credentials as a radical independent thinker with genuinely revolutionary values. See, for example, [Diogenes](#) living in a tub, eating in the marketplace, masturbating in public, and so forth.

How can I better tune my sense of shame?

"Be attentive to the appearance of evil. There is an inner voice in your soul which always tells you about approaching evil. You feel unpleasant, you feel ashamed. Believe in this voice; stop and seek to improve yourself, and then you will defeat evil." —Tolstoy^[9]

For shame to work well, it needs to be well-tuned so the things that evoke shame in you are the same things you want to be disincentivized to do. However, it is not uncommon for one's sense of shame to be miscalibrated. You may over-learn things that adults wanted you to be ashamed of as a child, and continue to be ashamed of them as an adult when this is no longer helpful. Or someone may have tried to shame you in order to take advantage of you or for sadistic reasons, before you were confident enough to shape your own sense of shame, and you ended up stuck with a sense of shame that's more harmful than helpful.* Sometimes shame is used as a way of enforcing social norms and taboos, such that there may seem to be a consensus that certain things are shameful ("boys don't cry," "well-done steak is an abomination") when really they're just unfashionable.

If shame is something like nausea, but meant to apply to our decisions and actions rather than to our digestion — what if we could harness shame the way an alcoholic harnesses nausea by taking a medicine like disulfiram that will make her sick if she drinks alcohol? What if we could shape our decisions and our actions by giving ourselves immediate visceral negative feedback for things we want to avoid? Aspiring rationalists: what if, every time you relied on a logical fallacy to make an argument, you flushed bright red, felt a knot in your stomach, and became so mortified that you wanted to disappear into a hole in the ground?

Unfortunately, in my research I didn't come up with much about how to go about deliberately crafting your own sense of shame. Intuitively, shame feels like it is more moldable than, say, pain or nausea. As I mentioned, I am not ashamed of the same set of things now as I used to be, and my sense of shame seems to me to have tracked my conscious decisions about what sort of person I want to be and what I value. But I don't have much in the way of confident insight into how this process works or how one might optimize it. I have heard persuasive arguments that what I'm calling "conscious decisions" are better considered as rationalizing afterthoughts, and if so my intuition here might not be helpful at all.

How to react to shame in better or worse ways

If you feel shame, and this prompts you to course-correct, to make amends, and to adjust your decision-making process, then congratulations: your response to shame is helping you get through life well. But there are a lot of ways this mechanism can go awry. Here are several:

1. You might ignore the connection between the shame and the specific deed you did (or failed to do) and try to convert the shame into essential shame: "I'm such

a fuck-up." That may *feel* like appropriate penance, but it actually gets you off the hook for doing anything about what you feel shame over, so it's an unproductive cop-out.

2. Instead of trying to reassess the decisions that you are ashamed of, you might try to shift the blame from the decisions to the context in which you made those decisions: "I feel terrible for cheating on my boyfriend. That's what I get for trying to go steady." or "I shouldn't have driven drunk again! Why won't they put a bar closer to my home?"
3. There's always denial: You could rewrite history so that the shameful thing never happened, or decide that it wasn't really shameful for [reasons], or it wasn't really shameful for you the way it would have been for other people because you're special, or something like that. This "works," but only in the same way that an anti-emetic works to help you digest more poison.
4. You might try to fetishize your shame ("I'm a bad boy and ought to be punished."). Some people get so hung up on this sort of thing that they deliberately do shameful things (e.g. pointlessly shoplift random shit).
5. Similarly, you might take pride in your shame for how it demonstrates what good standards you have. For example, Rudolph Hoess, the commandant of the concentration camp at Auschwitz, used the disgust he felt about the atrocities he efficiently carried out as a way of reassuring himself that he was really a good person deep down inside.^[10]
6. You might minimize your shame by pointing to other people who did things even more shameful and thinking "well at least I'm not as terrible as they are."
7. You might use your shame to virtue-signal: making a big public show about how bad you feel, engaging in a histrionic public confession, going on a sackcloth-and-ashes kick.
8. You might simply change the subject: "Sure I did something shameful, but the *real* issue here is..."
9. You might wax speculative about determinism and free will, recall that "nobody's perfect," wonder if a life without regrets is perhaps an impoverished one, suspect that a proper übermensch would not be ashamed of anything he does, note that you can't change the past, and so forth, and basically try to cover the whole thing up with a pile of self-help clichés and sophomoric philosophy until it all blows over.
10. You might lower your standards so that you no longer feel ashamed about whatever it is.

Because shame is unpleasant, and the process of taking responsibility, atoning, and improving can be difficult, there is great temptation to resort to evasions like these. As a result, people often get less help from their sense of shame than they could. One possible antidote is to be aware of these various patterns of evasion, and when you find yourself adopting one of them, remember to add the caveat: "...or on the other hand, is there anything I can learn from this feeling of shame in order to improve?"

1. ^

Mencius, *King Hwuy of Leang*, book VII, chapters 6–7

2. ^

Samuel Johnson, *A Journey to the Western Islands of Scotland* (1775)

3. ^

Oliver Goldsmith, in [*The Friar of Wakefield*](#) (1766) has his protagonist draw the distinction in this way:

Guilt and shame, says the allegory, were at first companions, and in the beginning of their journey inseparably kept together. But their union was soon found to be disagreeable and inconvenient to both; guilt gave shame frequent uneasiness, and shame often betrayed the secret conspiracies of guilt. After long disagreement, therefore, they at length consented to part forever. Guilt boldly walked forward alone, to overtake fate, that went before in the shape of an executioner: but shame, being naturally timorous, returned back to keep company with virtue, which, in the beginning of their journey, they had left behind. Thus, my children, after men have traveled through a few stages of vice, shame forsakes them, and returns back to wait upon the few virtues they have still remaining.

4. ^

Aristotle, [*Nicomachean Ethics*](#), book IV, section 9

5. ^

Karl Jaspers, *The Question of German Guilt* (1946)

6. ^

Theodor Adorno “Cultural Criticism and Society” (1951)

7. ^

Lucius Annaeus Seneca (the Younger), *Moral Letters XI.9*

8. ^

Alexander Solzhenitsyn, *The Gulag Archipelago* (1974)

9. ^

Leo Tolstoy, *A Calendar of Wisdom* (~1910) April 11

10. ^

Fred E. Katz, [*Ordinary People and Extraordinary Evil: A Report on the Beguilings of Evil*](#) (1993)