



Training Exercises E5 (Virtue Ethics)

with Example Solutions

With Solutions

Issue 1: *Gamification – revisited again again*

On previous exercise sheets, you saw the following case:

Tim is the manager of a supermarket. He ponders on introducing a gamification system in order to increase customer satisfaction. The system is supposed to work as follows: Every customer is asked to rate their cashier on a scale from 1 to 5 after checkout by pressing one of five buttons. The average of those ratings is calculated for every cashier and the employee with the highest rating gets a week of extra holiday as a bonus, while everybody else gets just their regular holidays. This results in them being a little less happy with their own number of days off than they otherwise would have been. Cashiers can always see their own average rating as well as the currently highest average, but not whom it belongs to.

1. What are plausible candidates for relevant virtues and vices in this scenario? Give reason for your choice (i.e., make plausible why virtue *X* or vice *Y* might be relevant for Tim's choice).
2. Is Tim morally allowed to introduce the system according the eudaimonian virtue theory sketched in the lecture? Choose some of the virtues or vices from 1. and answer this question under the assumption that the chosen virtues or vices are exemplified by Tim's choice.
3. Does your moral intuition that you had on the first exercise sheet have elements of a virtue theory? (Hint: Did you have the intuition that Tim's intention played a role? If yes, this might point into an affirmative answer to that question.) Which elements?

Sketch of a Solution 1:

The question here is quite vague and lacks some details that we would need. If you think that the case description has gaps which need to be filled, you can do so. If there are multiple plausible ways to fill a gap, then you should make a case analysis. In order to be a good ethicist, you have to be able to competently fill gaps in the cases you are given and do appropriate case analyses if needed. Exam questions, however, are always asked such that either there are no gaps or that it is apparent from the case that you have to make a case analysis. A full analysis of the case would usually be significantly

longer than this, because many more distinctions would be necessary in order to properly analyse the case.

1. Plausible candidates for virtues:

Justice If Tim cannot give everyone extra days off, but only some, and he believes in some kind of meritocratic justice, then this is a relevant virtue.

Supportiveness (This is not on the slides, but it is a plausible virtue.) If Tim believes that this system supports the workers best, then this is a relevant virtue.

Kindness Similar reasoning like for supportiveness: If Tim believes that introducing the system is an expression of kindness towards his workers, then this is a relevant virtue.

Prudence If Tim believes that it is a rational and clever idea to introduce the system, then this is a relevant virtue.

Plausible candidates for vices:

Greed If Tim introduces the system simply because he wants to earn more money, then this is a relevant vice.

Stinginess or Pettiness If Tim could easily give his subordinates more days off but he thinks he does not have to do so, then this is a relevant vice.

Ambition If Tim believes he can get a promotion due to an increased productivity after introducing the system and he only uses his subordinates as a stepping-stone in his career, then this is a relevant vice.

Spitefulness If Tim believes that his subordinates will have a worse relationship after introducing this system and he is eager to see this, then this is a relevant vice.

2. According to the eudaimonian virtue theory (as sketched in the lecture), the following holds: An option φ is right in context C (for agent A) if and only if a virtuous person (a virtuous version of A) would φ in C. Now, what does it mean for Tim to be a virtuous person? (To simplify things a little, we will limit ourselves here to (meritocratic) justice.) This means that a virtuous version of Tim rewards diligent workers, but he does so to an adequate amount. So, he does not reward them simply because of they are doing their normal work, but he also does not refrain from rewarding them if they are doing exceptionally good work. In this case, it is plausible that it is the right thing for a virtuous version of Tim to introduce the system because it enables him to do exactly that: rewarding the diligent while not rewarding the normal ones. So, Tim ought to introduce the system.

As a side note: we have seen in the lecture that the real reasons for which Tim does the action do not matter here. So, he still ought to introduce the system even if he is, de facto, motivated by all of the mentioned vices. It only matters that the virtuous version of him would introduce the system.

Issue 2: *Bad Things?*

We likely (hopefully!) share the intuition that certain things are wrong most of the time (or even all of the time), including

- (i) murder

- (ii) stealing
- (iii) cruelty towards animals

Answer the following questions for each of the above:

- (a) Use the agent-based virtue theory sketched in the lecture to explain that, for most situations, actions of this type are wrong. Can you do that for all of the above items? If no, is this a problem?
- (b) Can you come up with a (probably counterfactual) scenario, where the agent-based virtue theory sketched in the lecture would allow the action? If yes, present such a scenario. If no, why not?

Sketch of a Solution 2:

- (a) According to the agent-based virtue theory (as sketched in the lecture) the following holds: An option φ is right in context C only if φ expresses a virtue. (To simplify things a little, we will limit ourselves here to the four cardinal virtues of Thomas Aquinas.) It is easy to see that in most (if not all) situations, committing murder does not express a virtue. In other words: it will be difficult to find a situation where murder is supported by virtuous reasons. Murder, normally, is neither prudent, nor just, nor courageous, nor temperate, nor can we normally find any other virtue that provides a motive for murder. On the other hand, we can usually find at least one vice that provides a reason for murder, making the action directly wrong. Analogously, the same applies to the other two things: stealing and cruelty towards animals.

Alternatively, one could even argue that cruelty towards animals and murder can never express a virtue as they are vicious per definition. For an act of killing to be an act of murder the killing has to be, for instance, motivated by greed or by malice. This gets more obvious for cruelty towards animals: this act includes, literally, cruelty which is a vice.

- (b) Potentially, a knight in a shining armor is allowed to murder the bad guy and, thereby, save the beautiful princess (as long as it is the appropriate way to do so) because this act is expressing courage and, imaginably, no vice. Murdering a tyrant makes a similar case. For stealing, Robin Hood provides a good example. He could allowed to steal from the rich and give the spoils to the poor, since this action could express prudence, justice, and courage, and, imaginably, no vice. A stealing Robin Hood could, therefore, be an almost (completely) virtuous person. Cruelty against animals can, imaginably, be allowed in scenarios like that of the knight in the shining armor. If the bad guy is not a person, but a dragon (i.e., an animal), then warding it off the princes can be seen as courageous cruelty against an animal.

Maybe one has to construct an even worse scenario to make this case plausible, where the princess will be violated if the knight does not slay the dragon and the bad guy slays the dragon more brutally than the knight if the knight does not do so. Imaginably, this act of cruelty towards the dragon, then, expresses compassion (even towards the dragon).

On the other side (and as said in (a)), being cruel to someone or something seems to be a paradigm example of an action that nearly conceptually expresses a vice, whereby it (always) would be wrong according to the agent-based virtue theory (and the same holds for murder).

If this is the case, and the act of killing the dragon expresses a virtue and a vice at the same time, the agent-based virtue theory owes us an account of what to do when there are conflicting virtues/vices. Ultimately, it should always tell us, what we ought to do.

Issue 3: *Williams*

Recall the first example from Bernard Williams:

(1) George, who has just taken his Ph.D. in chemistry, finds it extremely difficult to get a job. [...] The results of all this, especially on [his] children, are damaging. An older chemist, who knows about this situation, says that he can get George a decently paid job in a certain laboratory, which pursues research into chemical and biological warfare. George says that he cannot accept this, since he is opposed to chemical and biological warfare. The older man replies that he is not too keen on it himself, come to that, but after all George's refusal is not going to make the job or the laboratory go away; what is more, he happens to know that if George refuses the job, it will certainly go to a contemporary of George's who is not inhibited by any such scruples and is likely if appointed to push along the research with greater zeal than George would. Indeed, it is not merely concern for George and his family, but (to speak frankly and in confidence) some alarm about this other man's excess of zeal, which has led the older man to offer to use his influence to get George the job... George's wife, to whom he is deeply attached, has views (the details of which need not concern us) from which it follows that at least there is nothing particularly wrong with research into [chemical and biological warfare]. What should he do?¹

- (a) What ought George to do according to the eudaimonian account of virtue theories as sketched in the lecture?
- (b) What ought George to do according to the agent-based account of virtue theories as sketched in the lecture?
- (c) If you come to different conclusions for (a) and (b): Why do you come to different conclusions?

Sketch of a Solution 3:

- (a) According to the eudaimonian virtue theory (as sketched in the lecture), the following holds: An option φ is right in context C (for agent A) if and only if a virtuous person (a virtuous version of A) would φ in C. Now, what does it mean for George to be a virtuous person? (To simplify things a little, we will limit ourselves here to the four cardinal virtues of Thomas Aquinas.) It means that George is a prudent, just, courageous, and temperate person. Additionally, George has these virtues or qualities in the right balance. According to this account of virtue theories, George should take the job. Arguably, the action is prudent because it lets George and his family overcome their difficult situation. Additionally, it is most likely just, because George prevents harm from happening. Furthermore, one could argue that taking the job expresses courage and temperance, because it is against George inclination. For the sake of others, George abstains from doing what he, originally, prefers to do.

¹quoted from *Utilitarianism: For and Against*, with J.J.C. Smart, Cambridge: Cambridge University Press, 1973, pp. 93ff.

- (b) According to the agent-based virtue theory (as sketched in the lecture), both taking and not taking the job could be right. We can imagine that the action of refusing the job expresses a virtue, e.g., integrity, while it expresses no vice. On the other hand we can argue, as done in a), that the action of taking the job expresses several virtues, e.g., prudence, justice, and courage, while it expresses no vices. Accordingly, in each case it can be argued that George ought to do the action. If George acts out of integrity, he ought to not take the job, but if he acts out of one of the other mentioned virtues, he ought to take it.

Furthermore, if George gets greedy for money and wants to take the job because of this or he just does not want to deliberate about what to do and wants to take the job out of sloth, doing so is wrong and he ought not to do so.

- (c) We came to slightly different results because the theories differ. For the agent-based theory only the actually expressed virtues count (which could lead to both results, depending on the respective virtues that motivate the agent), therefore, it is possible that two mutually exclusive actions are both right (although not simultaneously), since both could express virtues. According to the eudaimonian virtue theory, however, we must deliberate which action a fully virtuous version of George would choose and this is always the same one, regardless of the virtues that George actually motivate. Because of this, we were able to come to a clear result: it is plausible that a fully virtuous person would take the job, therefore, George ought to do so.

Issue 4: *Virtue Theories and Machine Ethics*

This is an issue meant for creative minds – and for lovers of science-fiction novels.

The field of Machine Ethics (that we will revisit later in this lecture) is concerned with the challenge of implementing morals into computer systems.

Brainstorm alone or in your discussion group: Is virtue ethics a promising approach to that challenge? Can there be virtuous robots or algorithms? How could a virtue ethics approach to machine ethics look like? (Hint: Think about the emphasis of moral education and experience in context of virtue theories.) What are possible shortcomings of such an approach and why?

Sketch of a Solution 4:

Virtue theories emphasize the importance of moral education and experience. This speaks in favor of learning-based approaches: machine learning could be used to train a computer system to behave morally. Machine learning, however, only works with a lot of data and (at least in this case) with supervision. Additionally, it is not clear where we can get this data from. Furthermore, if the learning is an interactive process: can we accept that the computer systems do bad things while in the process of becoming better? How can we see that the computer system learns to act because of the right reasons? Here explainability and transparency issues emerge (we will talk about these issues later in the lecture). Also, it is questionable whether the predicate "is virtuous" can be applied to computer systems at all. Finally, virtue theory is notoriously hard to apply even to humans, so it probably won't get any easier with machines.