# Zen and Rationality

# Zen and Rationality: Don't Know Mind

*This is post 1/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at the Zen notion of "Don't Know Mind" in rationalist terms.*

I'm a little unsure where "don't know mind" comes from. Sensei Google suggests it might be the Seon expression of the idea that in Zen is said "shoshin", often translated as "beginner's mind" but also carrying notions conveyed by translating it as "original mind", "naive mind", "novice mind", and "inexperienced mind" (noting that the character rendered "mind" is better translated as "heart-mind"). There's also a beloved koan often called "not knowing is most intimate" (Book of Equanimity, Case 20), and "don't know" is a good name to put to a particular insight you might have if you meditate enough. Regardless, "don't know mind" is a thing Zen practitioners sometimes say. What does it mean?

Depends on how you parse it.

The standard parsing is as "don't-know mind", as in the mind that doesn't know. This fits with the notion of soshin or beginner's mind, that is the mind that has not yet made itself up. In standard rationalist terms, this is the heart-mind that is curious, relinquishing, light, even, simple, humble, and nameless. Saying "don't know" is tricky, though, because there's the looming trap of the "don't know" that stops curiosity. Instead, this is the "don't know" that extends an open invitation of learn more.

You can also parse it as a command: "do not know" (since Zen falls within the Buddhist tradition that claims all you know is mind, "mind" is redundant here). This isn't an imperative to never know anything. Rather it's an encouragement to gaze beyond form into emptiness since our minds are often caught up in form (the map) and fail to leave space for emptiness (the territory[1]). More specifically, to know is to make distinctions, to construct an abstraction, to prune, to have a pior over your observations, to give something a name, to say this not that. Naturally, this means knowing is quite useful; it's literally what enables us to do things[2]. Yet it necessarily means leaving something out. "Don't know mind" is then advice to simultaneously let all in even as you keep some out.

Thus both interpretations converge on this idea that we can open ourselves to "knowing" more if we can hold fast to the realization that we don't already know it all.

---

1. If I'm being more careful, the duals of form and emptiness and map and territory don't perfectly align, map and territory being more akin to the phenomena/noumena or ontological/ontic split in Western philosophy. Nevertheless I think this is a good enough comparison to get the idea given the broader ways we sometimes talk about map and territory on LW. ↩

2. By "do things" and specifically "things" I'm here talking about things-as-things, that is things as reifications of the territory in the map, since properly things only exist in the map, not the territory, since without the map there is no thinginess, only a soup of stuff. ↩

# Zen and Rationality: Trust in Mind

*This is post 2/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at "trust in mind" from a rationalist perspective.*

I really like *Trust in Mind*, a Zen teaching poem attributed to the Third Patriarch, because it offers a vivid description of Zen practice that I've kept coming back to for new guidance and insight. Today I only want to talk about its origin and its title, because to delve in much more would be to write a book on Zen.

Purportedly it was written in response to the rising popularity of Pure Land Buddhism in an attempt to encourage the people to practice the "true" Buddha way, and revolves around the idea of putting trust in mind rather than trust, as in the case of Pure Land Buddhism, in a deified Buddha that will lead people to the Pure Land if you chant his name. This seems a deeply humanist message, and from a certain perspective this is fair: Zen contains a strong humanist aspect in that it's all about doing the practice here and now, and who else is there to do the practice but you, a human? But this is where the cracks in the humanist interpretation start to show, because Zen will immediately turn around and ask "oh, so where's this 'you' that's doing the practice?" and "does a dog have Buddha-nature?". Zen's tricky like that, offering you something with one hand while pulling the rug out from under you with the other, repeatedly trying to get you to look up and see what you've been leaving out.

Turning to the title, the "mind" here is again the heart-mind or heart-body-mind. This is an idea that comes up a lot in Zen, so it's worth saying a bit about it. Partly it reflects the traditional idea that a person's emotions or soul reside in their heart rather than their head, and the use of shin/心 to mean heart and mind is common in classical Chinese writing, not just in Zen/Ch'an texts. But if we're charitable and ask why would someone without the benefit of modern medicine think an import part of themselves resided in the heart, I reckon it's deeply related to the way we talk about *feeling* our emotions. That is, most people report experiencing emotions and sometimes general thought somatically, i.e. in the body. Sometimes this ability to literally feel emotions is impaired, and even when it's not we can get better at it through focusing on the mind-body-"heart" connection, and Zen sees this as vital to its project because it's a key part of studying the self to forget the self.

The other key word in the title that I've rendered here as "trust" can also be and frequently is translated as "faith". Now, I've buried that down here in the middle of the post for a couple reasons. One, rationalists and secular folks in general, including my past self, are often allergic to the word "faith" because of the way it's used by Christians, especially evangelical Christians, and I didn't want to put people off before they had a chance to read any of this. Two, I think the "trust" translation is more likely to resonate with rationalists; that was certainly the case for me. In fact, use of the word "faith" in spiritual jargon only started to really mean something to me after I came to develop some trust in mind! But just what is this trust-faith?

A good way start to see trust-faith is to look on the "small" end where reductionism bottoms out into epistemic circularity and makes us choose between consistency and completeness. Much to my surprise, while writing this I learned even Eliezer

acknowledges the need for faith of this type, though as of the time of writing that linked article he seems to have still held out positivist hopes. Regardless, it's the kind of trust you put in the axioms of a system; the unconditional acceptance you must grant to the smallest atoms of your ontology to which you can reduce all else but are themselves irreducible; the monumental importance you place on selecting the base types out of which you build your models. It's the leap across the epistemological chasm to the side of pragmatism and winning instead of infinite skepticism and solipsism.

Once on that other side, you can explore the "big" end of trust-faith. This is the kind of trust in knowing that you know nothing, in dealing with unknown unknowns, in being okay with the truth, and the kind of faith required to leave space for the "don't know" that I talked about last time. It's "knowing" the nameless virtue with your bones, and the utter acceptance of the Litany of Tarski and the Litany of Gendlin. For me, Alan Watts put it best when he wrote:

> Belief, as I use the word here, is the insistence that the truth is what one would “lief” or wish it to be. The believer will open his mind to the truth on the condition that it fits in with his preconceived ideas and wishes. Faith, on the other hand, is an unreserved opening of the mind to the truth, whatever it may turn out to be. Faith has no preconceptions; it is a plunge into the unknown. Belief clings, but faith lets go. In this sense of the word, faith is the essential virtue of science, and likewise of any religion that is not self-deception.

When I first heard that quote, it reframed "faith" for me in a way I could feel comfortable using it. I think of it as similar to the reframing Bayesianism gave me on "belief", a word I had previously avoided using at all cost because of the way Christians around me growing up used it to mean blind acceptance of specific facts because their religion told them to. So just as we need not abandon "belief" because some people with anti-rational epistemics use it as a stopping word, we need not abandon "faith" either, even if "trust" is perhaps a more serviceable synonym.

And at that I'll leave off, although we're not actually done with unpacking "trust in mind" because we've only looked at a sliver of what "mind" really means in Zen, but we had to get through trust-faith today, and I think it's a lot on its own. I'm sure, though, I'll find my way to saying more about it if I keep this series up.

# Zen and Rationality: Map and Territory

*This is post 3/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at form and emptiness from a rationalist perspective.*

Rationalists have a few key ideas or memes (in the memetic sense), and one of them is "the map is not the territory". Lots has been written about this idea on LessWrong, but it's an idea with a history that stretches back for thousands of years, so it's perhaps not surprising that it's also one of the ideas at the core of Zen.

But in Zen we don't use the words "map" and "territory", instead preferring numerous other metaphors to point at this distinction. Let's explore a few of them, because each elucidates a different aspect of the truth pointed at by these duals.

Before Zen was Zen, Nagarjuna formalized this idea that there's a duality between map and territory in the two truths doctrine. He called these two pairs form and emptiness, pointing at the way our minds put our experiences together into forms or objects that are fixed, at least in our minds, yet ultimately reality is empty of these forms or any other kind of inherent distinctions, essences, or ultimate and timeless truths. Everything we know is provisional, taking a skeptical epistemic stance similar to Pyrrhonism.

Form and emptiness have their place in Zen, but more common is to make a distinction between the relative and the absolute. The relative is that which changes, which exists in our minds, which comes and goes. The absolute is that which exists prior to our perception of it; it's the space in which the relative arises. But Zen doesn't stop there. Form is emptiness and emptiness is form, as the Heart Sutra says, and the relative and the absolute can be thought of as dancing reality into existence, simultaneously unified and separate. Dongshan (Japanese: Tozan) explored this in his poem on the Five Ranks, a subtle teaching that can take some effort to penetrate but is worth the effort.

Talking about relative and absolute can get a bit abstract, as can talking about form and emptiness, so there's another pair that's been used extensively in Zen teaching that, alas, holds little currency for us Westerners: guest and host, or alternatively vassal and lord. I don't have much to say on these because they mostly make sense in the context of the pre-colonial Sinosphere, but I mention them in case the metaphor resonates with you.

For Westerners, I think our philosophical traditions offer some alternatives. Kant offers us phenomena and noumena, which sadly misses the mark a bit as often understood by assigning essential form to the territory/emptiness/absolute by suggesting there are things-in-themself than nonetheless have thingness. Better are Heidegger's ontological and ontic, which are just fancy Greek words for something like "words or ideas about what is" and "that which is", respectively. Although even "that which is" is a bit too much to describe the ontic; better to say the ontic is the "is" or "being" or "to be". Put another way, ontology is like the nouns, and the ontic is like the verbs just on their own, without even a distinction between one verb and another.

An analogy I like that I borrow from topology is to liken the map/form/ontology to closed sets and the territory/emptiness/ontic to open sets. This is by no means perfect and if you think about it too hard it falls apart, but using my intuitions about closed and open sets helped me make better sense of the two truths, so I share it with you in that spirit.

And at that I'll end this post. I've not said much about the actual relationship between the two truths of map and territory or how their dependence on one another creates reality as we experience it. I'll tantalizingly hint that ideas about embedded agency go a long way towards exploring how the two truths play together, but exploration of that I'll save for another time.

# Zen and Rationality: Just This Is It

*This is post 4/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at "just this is it" from a rationalist perspective.*

When Dongshan, the co-founder of what would become the Soto Zen school within which I practice, was preparing to leave his teacher Yunyan and go out in the world, he asked Yunyan how he might summarize his teaching. Yunyan replied, "just this [is it]". Because the more we say the more we move into the world of words and away from reality as it is on its own prior to conception, this is often shorted in various ways to "just this" or "just is" or "this is" or "it is" or, perhaps best of all short of saying nothing and letting reality stand on its own, "is".

This is arguably the core teaching of Soto Zen and maybe all of Buddhism, to perceive and accept reality just as it is. Yet I see it all over the place in the LessWrong corpus, too. I'll mention a few of these.

Egan's Law posits that "it all adds up to normality". In "A Technical Explanation of Technical Explanation", Eliezer phrased a similar sentiment as "since the beginning, not one unusual thing has ever happened". Both point at the way that reality is just as it is, and the only way we can be confused or surprised is because we had an idea about how reality is rather than simply looking and seeing how it is.

I think this is a hard thing to remember, because to the kind of people that are attracted to Less Wrong, better models of reality are very attractive. I know they are to me! Yet it's very easy to go from accepting reality as it is and trying to better predict it to getting lost in the model that does the predicting and confusing it for the real thing. Thus, while at the same time we look for models with better gears that more precisely carve reality at its joints, we also have to remember those boundaries are fuzzy and that all models are ultimately wrong even and especially when they are useful. It's perhaps the great koan of Less Wrong to build better models while simultaneously accepting that all models are somewhere wrong.

To help us deal with this koan, we have a poem to help us. You might think I mean the Litany of Tarski, but you would be wrong, because that poem is about having beliefs correspond to reality, but "just this is it" is all about getting under those beliefs and just seeing what's actually being perceived. For that, we turn to the Litany of Gendlin:

> What is true is already so.
> Owning up to it doesn’t make it worse.
> Not being open about it doesn’t make it go away.
> And because it’s true, it is what is there to be interacted with.
> Anything untrue isn’t there to be lived.
> People can stand what is true,
> for they are already enduring it.

This was said by Eugene Gendlin of Focusing fame, a technique for helping you reconnect to your perceptions just as they are without judgement or modeling. The method is simple, yet its impact can be profound for many to get out of their ideas about how things are and to get back to what evidence they are actually getting about the world. Zen asks us to over and over again come back to this fundamental point

that reality is just as we perceive it, not what we believe about it, and that belief is just a useful mechanism for helping us better live our lives, if only we don't get tripped up into mistaking the map for the territory.

Finally, to return to "A Technical Explanation of Technical Explanation", it contains one other phrase that neatly captures the spirit of "just this": "joy in the merely real". If we can take joy in what actually is, if that can be enough for us, then all else becomes the playground in which we live our lives.

# Zen and Rationality: Skillful Means

*This is post 5/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at skillful means from a rationalist perspective.*

As part of Zen practice, a teacher may use or encourage the use of many skillful means. Sometimes called expedient means or upaya, these may include things like encouraging a student to take up a particular meditation practice like breath counting or labeling, assigning a koan, or giving the student a job within the Zen center, like alter care or bell ringing. These are meant to aid the student in their practice of the way.

Importantly, the idea behind these means being "skillful" or "expedient" is that they are not necessarily practices a student will continue with forever, but rather things the student should do now that will help them. What may be useful for one student to do may not be useful for another, and what was once a useful practice for a student may later become a hindrance.

A good example is structured meditation practices, like breath counting. When a student begins meditating, they might find it very difficult to stay seated for 30 minutes, even during timed meditation periods sitting with other people in a zendo. Giving them something to do, like counting their breath, gives them something to focus on and distract the part of their mind that wants to get up and do anything else. Over hundreds of hours, they'll retrain themselves to be able to stay seated even when distracted by gathering evidence that they can do it, and slowly the breath counting will stop being a skillful means to help them practice and will instead become a hindrance and a distraction from just sitting, at which point they might move towards a different, less structured meditation practice or spend less of the time in meditation counting breaths. If they kept counting their breath for years even after it was no longer necessary to get them through a sitting period, it would no longer be a skillful use of the means, but as long as the alternative would be failing to keep sitting so they have the chance to develop skills that will let them meditate more full heartedly, it remains skillful.

Rationality has its own version of skillful means via the practice of instrumental rationality, or systematically achieving one's ends. It's the art of finding ways to help one become stronger. It recognizes that you can't go straight to the ideal and convert oneself into a perfect Bayesian reasoner with infinite memory and thinking capacity, but instead must work with your fallible, human self to find tools and techniques that help you where you are now to get you incrementally closer to where you want to be.

It also recognizes that sometimes the skillful next step actually makes you temporarily worse off as you climb down the hill of a local maxima to move to a different, higher local maxima.

If you want to practice these skillful means, the CFAR handbook is a decent starting point for learning some of them, and a CFAR workshop is maybe an even better option. Further, many of the posts tagged Practical(2031) contain useful techniques to aid you in achieving your goals, as do posts tagged Rationality under the "Applied Topics" and "Techniques" groups of tags.

There's even some impressive overlap between rationalist and Zen skillful means. For example, there's the general act of noticing, be it noticing confusion or anything else, that's essential for studying the self in enough detail to work with it (and in Zen, to eventually forget it). There's trying things (famously Zen teachers may tell their student not to bother understanding something, but just to do it and see what happens). And focusing is, from a Zen perspective, one more useful means to reintegrating the heart-mind-body.

Importantly, both rationality and Zen acknowledge some version of the typical mind fallacy , carrying the realization that what's best for one person now is not necessarily what's best for them later, and that what works for one person may not work for another. Lucky for us we have so many skillful means to choose from on our journeys!

# Zen and Rationality: Karma

*This is post 6/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at karma from a rationalist perspective.*

Karma is one of those ideas that is poorly understood both because it means different things in different contexts and because there's a simple but bad abstraction of the idea that's easily understood and shared but points to something other than what we mean by "karma" in Zen, thus leading to bucket errors.

So, first, what is karma not. The Zen notion of karma is not any of the related ideas also called "karma" present in other dharmic traditions. It's also not the New Age or Theosophy notion of karma. All of these you are probably familiar with and might be expressed in one of the following ways:

- what you give is what you get
- the world will punish you for your bad acts and reward you for your good acts
- you'll be rewarded/punished in proportion to your acts in this life via the quality of your rebirth into a next life
- karma is a "sin debt" you carry forward, possibly from past lives, that you must work off to have a better life

In Zen (and some other Buddhist traditions), karma is nothing so involved. Instead, you can just read it as "causality" and you'll be 80% of the way of understanding it, with the remaining 20% some subtlety around how causal relationships relate to the fundamental flowing of one moment into the next that come from understanding how map and territory relate.

As you might guess, the other notions of karma relate to this one as heuristics about how earlier moments affect later ones. For example, the thread of moments out of which a person can be identified contain within them all the causes of the person's present condition, and it's often true that if you behave prosocially towards others they'll behave prosocially towards you, so a reasonable compression of karma that skips all the finer points look a lot like "if you do good stuff, good stuff will happen to you; if you do bad stuff, bad stuff will happen to you". Mix in some beliefs about reincarnation and you have karma as it is commonly encountered, bucket errors and all.

# Zen and Rationality: Continuous Practice

*This is post 7/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at continuous practice from a rationalist perspective.*

In Zen there is a focus on continuous practice (*gyoji dokan* in Japanese). The idea has some subtlety lost in translation: it's referencing practice of the Buddha Way that is maintained, preserved, sustained, ceaseless, and persevered through (likened to the need to continually sharpen a frequently used blade against a grinding stone) in a ring or cycle of ongoing change (this is both talking directly about the reality of practice and an allusion to the wheel of life).

Two common ways this shows up in Zen is as both an emphasis on continuing to train the same thing over and over because it's through doing "the same thing" that we come to realize impermanence, and a focus on maintaining one's practice after various "attainments" like kensho or enlightenment. The big idea here is that one is never really done learning and developing, so continual training and practice is necessary to prevent sliding into greater delusion. As a translation of the Bodhisattva's Vow puts it: "dharma gates are boundless, I vow to enter them".

Although rationality seeks somewhat different purposes than Zen, I think we find a similar current of thought running through rationalist writing, urging rationalists to continually practice the art. Some examples:

- Scott Alexander wrote that rationality is a habit to be cultivated. As such, cultivation of that habit requires ongoing work, which he captured with the phrase " constant vigilance ".
- OODA loops are a technique for continuously reassessing a situation and updating one's response to it. They can be generalized to operate over different scales.
- Many rationalists refer to themselves as "aspiring rationalists" in order to make clear that they haven't achieved the goal of being rational, but are instead always working at it.

And I think there's something in the ethos of rationality itself that suggests continuous practice. The idea that rationality is systematized winning carries with it the idea that winning, whatever that means for you, requires ongoing effort to actually use the system—the moment you stop putting in the work to reckon things is the moment you stop winning, even if only marginally. It's only through continuous application of effort that one can succeed, at least for a moment, at winning.

# Zen and Rationality: Equanimity

*This is post 8/? about the intersection of my decades of LW-style rationality practice and my several years of Zen practice.*

*In today's installment, I look at equanimity from a rationalist perspective.*

In Zen in particular and various Buddhist lineages in general there's a lot of talk about equanimity. Sometimes it refers to a particular meditative state, other times to a more general virtue of meeting the world on equal footing with no particular preconceptions about what one will believe.

It has a so-called near enemy, though, which is indifference, and this is central to the straw stereotype of the advanced Buddhist practitioner you've likely encountered. Think of the unflappable monk who continues to meditate while the building burns down around them, or more likely the idea that through meditation you should become a person who can suppress their emotional responses and never smile, frown, laugh, or cry no matter what the world throws at you. I don't know much about stoicism, but indifference is the same kind of thing people seem to mean when they say that a person is stoic.

Equanimity is something quite different. From the outside it might produce behaviors you could interpret as indifference, especially in training environments like a meditation center or monastery, but in fact it's something more nuanced than that, and something rationalists are quite familiar with.

In everyday language we might say equanimity is about being open to the possibility that the world is just as it is rather than how you think it is, and that whatever thoughts and beliefs you form should not be formed on the basis of what you wish were true but instead on the basis of what you actually observe.

If we focus only on beliefs, equanimity is about being a good Bayesian reasoner, updating fluidly and proportionately in response to evidence. The only thing to watch out for here is that we often model Bayesian reasoning in toy environments where physical reality is constrained to match the model, and Bayesian updating out in the real world includes a lot more uncertainty about the correctness of the fundamental assumptions the model makes than you'll encounter in worked examples.

Thus it may not surprise you that we find equanimity in the Litany of Tarski:

> If the box contains a diamond,
> I desire to believe that the box contains a diamond;
> If the box does not contain a diamond,
> I desire to believe that the box does not contain a diamond;
> Let me not become attached to beliefs I may not want.

There's some tricky phrases in this litany that I don't quite endorse like "I desire to believe" and "beliefs I may not want", but it conveys the right sentiment and it's part of the rationalist cannon, so it's worth mentioning. The last line mentions attachment, and non-attachment is another thing you might hear a lot about in Buddhism since it's about how to have an equanimous mind. But giving up attachments is a topic for another day.