Engineering Virtue: Constructionist Virtue Ethics

Jakob Ohlhorst, VU Amsterdam

Abstract: Virtue ethics is traditionally a conservative project. It analyses the virtues that humanity has been relying on since antiquity. This conservatism unduly limits the potential of virtue ethics to contribute to moral progress. Instead, we should pay more attention to constructionist virtue ethics with the help of conceptual engineering. I will argue that revising and ameliorating the virtue concepts which a community uses directly and indirectly leads to a change of the virtues that exist in this community. By revising and innovating virtue concepts, we can re-make and improve the virtues that we have.

Virtue theory; Concepts; Metaphilosophy; Conceptual Engineering

1. Introduction

Contemporary virtue ethics has a conservative streak. It frequently conceives of virtues in the same terms and with the same perspective as its ancient and mediaeval precursors as well as our common sense conception. Authors look to folk intuitions and religious and philosophical authorities from a broad range of traditions to learn about what the virtues are. An example for this is Linda Zagzebski's (2010) virtue exemplarism. She argues that our concepts of virtue derive from the characters of admirable moral exemplars like Saint Francis of Assisi, Confucius, or Jesus Christ. Consequently, finding out *what virtue is* requires describing the character of these admired past exemplars.

This conservative tendency makes virtue ethics an often analytical enterprise. Methodologically, it aims to analyse the concepts of virtue that we already use, clarifying and sharpening these traditional concepts. This analytical methodology is historically seen a part of the programme of modern virtue ethics, as originally proposed by G.E.M. Anscombe in "Modern moral philosophy" (1958). She raises the

different intuitions about flourishing and justice that different people might have and assigns the task of harmonising them to philosophy. (Anscombe, 1958, pp. 18–19)

I want to propose a different conception of doing virtue ethics. Instead of analysing what kind of virtue, say, courage is, we ought to construct a new concept of courage* as a virtue that serves our moral and theoretical needs. We should be *constructionists* (Floridi, 2011) about virtue and *make* the virtue concepts that will help us flourish and understand moral character. This can be done through the method of conceptual engineering. This method is already being used implicitly by some virtue ethicists; this paper simply makes it explicit and argues for its philosophical and moral benefits. We can engineer virtues in two ways: Either we revise and improve an extant virtue concept, or we create a new virtue concept to satisfy our practical and theoretical needs.

I distinguish between broad linguistic *concepts,* which we share and communicate with, and narrow cognitive *conceptions,* which an individual uses to classify objects, interact with them, etc. (Koch, 2021a) Concepts arguably supervene on a combination of the totality of conceptions in a community and environmental facts.

I will first explain what I understand by conservative analytical virtue ethics. Second, I introduce the method of conceptual engineering. Third, I will show how engineering virtue concepts can create new virtues and what constructionism's advantages are over conservative virtue ethics. Finally, I will examine some objections and further directions of research.

2. Conservative virtue ethics

The default philosophical method in virtue ethics is conceptual analysis. That is, usually, virtue ethicists try to discover the true nature of particular virtues by analysing the concepts that designate these virtues. This is the same project that Plato pursued in the *Republic*: Answering the question "what is justice?" (Plato, 2000, p. 331c *ff.*) This approach is based on our intuitions about what is and what is not designated by particular virtue concepts. Note that analysis does not mean that we do not aim at normative evaluations.

Frank Jackson (2000) has proposed an influential account of conceptual analysis. The method consists in two steps: The first step is to gather the relevant folk conception on an issue, that is the intuitions and tenets that most ordinary people share about the concept. In the field of virtue ethics this includes for instance the intuition "that we know that the cowardly are less likely to do what is right than the courageous". (Jackson, 2000, p. 136) This store of common moral intuitions and tenets constitutes our *folk morality*. (Jackson, 2000, p. 130) Folk morality determines the semantic role that different moral concepts play; however it is not without internal contradictions. Also past authorities can be a source of moral tenets. I take this first step to be essential for conceptual analysis.

The second step analyses which precise conception or definition actually plays the semantic role that folk morality ascribes to the concept. The goal is to accommodate as many intuitions as possible. However, it may turn out that some intuitions cannot be accommodated because they are in contradiction with our other intuitions, or our scientific knowledge. (Jackson, 2000, p. 132) For example, the Aristotelian tenet that experiencing fear is incompatible with courage (Zavaliy and Aristidou, 2014, p. 178) may be rejected due to theoretically weighty incompatible intuitions and our psychological background knowledge. This means that conceptual analysis does not require that a single counterexample will overthrow a whole theory – there are different ways of executing the second step.

On a strict reading, conceptual analysis is extremely conservative because it does not revise concepts in any way: It just gives necessary and sufficient conditions for how we actually use our concepts – it gives us a theory of our concepts, not concepts.[1] However, Jackson recognises that it would be too strict if the moral concepts that we analyse were set in stone. For that reason, he permits that our "folk morality is currently under negotiation". (Jackson, 2000, p. 132)

This methodology already can be found at the inception of virtue ethics. Aristotle also attempted to reconcile commonly accepted intuitions – our *endoxa* – with our empirical

---

[1] The question here is what kind of concepts folk morality gives rise to: An epistemicist would argue that unknown but precise concepts flow from folk morality. But you may also argue that the concepts of folk morality are vague or unstable due to the contradictory intuitions underwriting them. I would like to thank [SK] for pressing this issue.

knowledge and with each other. (Reeve, 2012) Conceptual analysis is a conservative method in the sense that it is only revisionary, i.e. rejects tenets and conceptions, when it cannot be helped due to internal contradictions of the common intuitions or with our scientific knowledge and the goal of such a revision is only to purify the concept to be analysed from these contradictions.

Agent-based virtue ethics, like Zagzebski's (2010) or Slote's (1995), offer prime illustrations of this conservative method. Slote appeals to intuitions about what constitutes an admirable trait. Meanwhile, Zagzebski very explicitly appeals to the toolkit of analytical philosophy of language to determine what the virtues are – notably a direct theory of reference. (Zagzebski, 2010, p. 50) The direct reference to the admirable agent's virtue is fixed when we admire moral exemplars' character traits. This admiration is based on our intuitions about what is admirable – it is based on our folk morality. However, our folk morality does not give us a definition, it only fixes the referent. (Zagzebski, 2010, p. 52)

Target-centred virtue ethics is committed to the method of conceptual analysis as well. Swanton (2003) takes our folk conceptions of virtues as a theoretical starting point. She aims at a reflective equilibrium between those folk conceptions and more broadly eudaimonistic considerations about human flourishing – be they naturalist or not. Finding a reflective equilibrium is one of the possible ways how to execute the second step in conceptual analysis.

Meanwhile, eudaimonistic virtue ethics are not necessarily wedded to conceptual analysis. These approaches define virtues as traits that support human flourishing. Nevertheless, there is also a conservative analytical approach to the eudaimonistic project where our folk morality determines what we would count as flourishing and thereby what counts as virtue. Annas (2011), for instance, does rely on folk intuitions about flourishing as a starting point, and she does aim to preserve this folk morality in her coherentist framework. She just rejects the narrower conception of conceptual analysis as refuting a definition on the basis of individual counterexamples.

3. Conceptual engineering

Conceptual engineering or *conceptual ethics* (Burgess and Plunkett, 2013) gives us the right framework to explain constructionist virtue ethics. Conceptual engineering does not aim to analyse the concepts that we *actually use.* Instead, it tries to find out which concepts we *should be using* and how to revise our existing concepts in that sense.

Conceptual engineers first assess conceptual deficiencies, second, they design revised or new concepts that we should be using instead, third they influence the linguistic and epistemic community to adopt these new or revised concepts. A *successful process of conceptual engineering* produces a concept that is useful for the epistemic community and that is indeed adopted by members of the community, thereby also changing their conceptions and usage of the related term. (Sawyer, 2021) Note, that conceptual engineering may also fail, either because the developed concept does not satisfy its theoretical and practical purposes, or because it is not adopted by the community.

Let me illustrate this with virtues. As argued above, conservative virtue ethics asks the analytical question: "What is courage?" or sometimes even more specifically "what is Aristotle's concept of courage?" (Zavaliy and Aristidou, 2014) The conceptual engineer instead asks the normative question: "What should our concept of courage be?" or "which concept of courage would be best?" First, she examines what practical (and theoretical) deficiencies our current concept of courage has; next, she tries to propose a new and revised concept of what "courage" should mean, usually through a revisionary definition. Third, she spreads the word about the engineered concept.

If the engineering process is successful, then the new concept of courage* is adopted by the epistemic community. This changes the cases that its members call "courageous*" as well as their conception of what it is to be courageous*. To illustrate how revisionary such a new concept can be, consider Sally Haslanger's (2000) engineering proposal the concept woman as a person who is systematically oppressed due to biological features that society ascribes to her. This engineered concept pays no heed to the folk intuitions or past authorities that play a role in conceptual analysis.

The specific method that I advocate for in this paper is *practical conceptual engineering.* (Nado, 2021, p. 3) Namely, when engineering a virtue, we should look at the purposes that the concept should satisfy. Concepts can have multiple and at times incompatible theoretical and practical purposes. (Fassio and McKenna, 2015) One goal of conceptual engineering is to harmonise or salvage these incompatible purposes. Sometimes this may mean splitting an undifferentiated concept into several more fine-grained concepts, each serving different purposes. (Nado, 2021) The purposes of a virtue concept like courage are also multiple: It serves to instil military morale, to praise or criticise risk behaviour in a broad range of contexts, and to give individuals psychological and behavioural models to engage with the world and deal with their fears. Note, that we can add or modify the purposes that a concept to be engineered should satisfy.

Amie Thomasson (2022) takes a very similar approach called *pragmatic conceptual engineering*: In her terminology purposes are "functions" of concepts, and she distinguishes two kinds, interpersonal and textual functions – these correspond roughly to what I call practical and theoretical purposes. The example of the purposes of courage that I gave were interpersonal functions that influence social behaviour, while textual functions concern how we can transmit information. I will focus mostly on interpersonal functions in this essay. Apart from purpose-driven practical conceptual engineering, there is also Carnapian (1962) explication; an engineering method that focuses on a concept's textual function of precision.

The practical conceptual engineer needs to pay no heed to the *principles* that govern the current use of a concept, i.e. the tenets of folk morality and past authorities. (Fassio and McKenna, 2015) As mentioned, Aristotle took for example the principle that experiencing fear is incompatible with courage to be true. (Zavaliy and Aristidou, 2014, p. 178) A conceptual engineer can simply throw this principle overboard and ignore it, if it does not align with the purposes that the concept of courage should satisfy.

By contrast, on a conservative approach through conceptual analysis, the principle will be kept onboard by default. To get rid of it, it has to be weighed up with other incompatible principles. The fearlessness principle is thrown overboard by most contemporary conservative virtue ethicists, but they need to justify this move by

arguing that it is incompatible with other tenets of folk morality. When we engineer virtue concept V, we do not need to pay attention to the objection that "this is not what V means", but the conservative virtue ethicist does.

4. Virtue Constructionism

The conservatism of virtue ethics as conceptual analysis limits its potential. If we only analyse what courage *is,* then we remain confined to options that are in the vicinity of the concept of courage that we already possess – we remain tied to our folk morality. Consequently, moral progress that occurs as a product of analysing our concept of a virtue will necessarily be incremental. For example, we can only *clarify* that courage does not require fearlessness. I do think that folk morality can be improved upon; I will simply take for granted that there are indeed deficient moral concepts in our epistemic communities. We will be bound to the virtue concepts that we already possess and about which we have intuitions, never going beyond them. We will always be thinking about courage, patience, and whichever further virtues that folk morality has transmitted to us.

This stands in contrast to virtue constructionism. With conceptual engineering, we can create new virtues concepts from scratch, and we can radically revise the virtue concepts that we possess already without the constraints of our folk morality. Note that there are two ways of engineering concepts: We can either introduce a new concept, or revise an extant concept. (Chalmers, 2020)

Whether we introduce a new virtue concept or revise a pre-existing one depends on the core purpose we want to satisfy with the concept. If there already is a folk concept that has the same, or closely related, purposes as the ones we want to satisfy, then we should re-engineer that concept to ameliorate it to better fulfil these purposes. If there is no extant virtue concept that serves these purposes in any way – for example because the purposes are bound to a novel context like social media – then we should create an entirely new virtue concept.

Virtues are a particularly fruitful object of conceptual engineering. Namely, I will argue that, if we successfully engineer a virtue concept, i.e. design a concept that is useful and ends up being actually used in an epistemic community, then we also *create* the

corresponding virtue. Note that successful engineering means that the concept is indeed broadly used in relevant contexts in the ways that it was intended to. After a concept has been engineered, the corresponding engineered virtue is created through three mechanisms. These mechanisms transform the new virtue concepts into real new virtues because successfully engineered virtue concepts influence the epistemic community on three levels.

The first mechanism that makes the engineered concept into an engineered virtue is the *mind-to-world* determination of virtue concepts. Mind-to-world determination means that a community's use and conception of a term determine the concept that it expresses. (Sawyer, 2021, pp. 10–11) Take for example the concept of democracy: Our current use and conception of democracy would not include Attic "democracy" or 18th Century French "democracy" which both restricted political rights to propertied males. The concept of democracy entirely depends on a community's use and conception of it, and ancient Greece's concept of democracy is profoundly different from ours.

This stands in contrast to concepts that are determined *world-to-mind* which fundamentally depend on factors independent of the community using the concept. Typical examples are natural kind concepts like water or gold whose extension is determined by the natural kind to which they refer. Their extension arguably has never changed. (Putnam, 1973) I am confident that virtues as practiced in society are no natural kinds, and that virtue concepts are determined completely by the epistemic community's use and conception of the relevant terms.

You might object that there could be natural kind concepts of virtue, where they simply designate dispositions that lead to an organism's thriving within some or all environments. However, first, I fear that such a pure naturalist conception of virtue may have difficulties separating morally good success from amoral success – especially given morally corrupt environments which require a different moral profile than, e.g., a functioning civil society or a nomadic community. In a totalitarian dictatorship, it is difficult to flourish in a naturalist sense without morally compromising oneself.

Jakob Ohlhorst

Second, virtues as natural kinds would be hard to individuate because every successful agent will have a different behavioural profile, which leads to a proliferation of virtue concepts – this is a version of epistemology's generality problem. (Conee and Feldman, 1998; Alfano, 2013)[2] Third, such natural "virtues" would fail to fulfil the full social function of virtues because they would not be praised and emulated as socially recognised virtues. I would argue that this social aspect of praise, admiration, and emulation is essential for being a virtue in a robust sense. Note that if virtues were (hidden) natural kinds an agent might still be admired for their character because it also satisfies the criteria for a distinct folk virtue. I consequently agree with MacIntyre (1981) that virtues supervene on social practices.

Given the second and third points I just made, I am also committed to the claim that virtues are not *basic moral* kinds (Sayre-McCord, 1997) – granting that these are distinct from natural kinds. On my view, virtues are social kinds – or complex social moral kinds – whose structure is sensitive to facts about both basic moral and natural kinds because our social behaviour is sensitive to these facts. In that sense, virtue concepts are like concepts of mental illness, bound to social contexts. (Bentall, 2003) Social kinds nevertheless function very similarly to natural kinds, in that they ascribe inner essences, which may explain the risk of confusion. (Neufeld, 2022, p. 6)

If a community's conception and use of a concept like democracy entirely determine the *concept* that it refers to, then *what democracy is* for this community is also entirely determined by the epistemic community's use and conception of democracy. The changing use and conception of democracy throughout history means that democracy itself has also changed throughout history. Successful conceptual engineering changes an epistemic community's use and conception of a concept. Therefore, in the case of mind-to-world concepts, it changes *what the concept is about*. This is the "worldliness" of mind-to-world conceptual engineering: Changing a mind-to-world concept means changing the thing it refers to. (Cappelen, 2018, p. 138) Conceptual analysis, by definition, cannot do this because it is tied to our folk concepts.

---

[2] The situationist challenge (Doris, 2002; Alfano, 2013) also concerns virtue engineers, as a constraint on the kind of concepts that they can design. A successful concept will be one that is actually instantiated.

Consider a simplified example about virtue: A conceptual engineer has successfully re-engineered the folk concept of courage in a way that people who resist against oppression in whatever way now count as courageous*, while oppressing others is incompatible with courage*. Let us postulate that re-engineering the concept of courage in this way was reasonable because its purpose is closely related to that of courage*, and their purposes would be better served by concept courage*. Given that the engineering was successful, the epistemic community now conceives of and uses the concept courage* in a way that resistance against oppression is a sufficient condition for it and someone who oppresses others is not counted as courageous*. Courage* is a mind-to-world concept – it is entirely determined by the community's use and conception. Given that this use and conception has changed from the original use and conception of courage, the concept, and the things that it refers to, have changed. There are now new instances of courage* that weren't there before, namely cases of resistance against oppression that didn't satisfy the original concept of courage – obviously these cases already existed before the engineering, but they were not efficacious as instances of courage* because they were not recognised as such. Instead they were simply a disjunct set of "natural" dispositions to resist oppression. This first mechanism is a deflationary way how virtue constructionism creates new virtues by simply reframing the world in our linguistic practice.

I also take certain behavioural dispositions to belong to the use and conception of a concept. For a virtue concept to really play the role of a virtue concept, it must be accompanied by certain cognitive, emotional, and behavioural dispositions, especially praise, admiration, and emulation of the designated trait. This behaviour arises out of our understanding of the normative valences that are part of a concept: The concept courage* assigns a positive valence to resistance to oppression, and when you understand the concept you will be disposed to act accordingly, e.g. praising such instances. That is, virtue concepts have illocutionary force. Slurs illustrate the behavioural role of valenced concepts very well – using a slur is itself an illocutionary act causing harm, (Neufeld, 2019) and it may also licence further harmful behaviour. (Liu, 2021) Learning to use a slur means to learn certain behavioural dispositions. The same goes for virtues, but positively – if I call someone courageous*, I am *ipso facto* praising them, and by this use I am licensed to emulate that person's behaviour.

<div align="right">Jakob Ohlhorst</div>

The second mechanism contributes more substantively to realising the existence of the engineered virtue. As mentioned, if the virtue engineering process was successful, then the new virtue concept is a useful concept that will be used frequently. When people learn the new conception, they will be struck by the concept being valuable and useful for them. This will mean that people will pay (more) attention to this engineered virtue. On the one hand, they will be more likely to praise instances of it as a virtue which in turn positively reinforces the virtuous behaviour. On the other hand, they may try to emulate the virtue themselves, behaving in ways to realise it. Indeed, without the engineered virtue concept available, people may not even be aware that there is anything to be praised or emulated there.[3]

To take our example of courage* again, the instances of courage* may for the first time be perceived as virtuous – i.e. to be praised and emulated – because the correspondent concept is now available for the first time. Agents may now strive to be courageous* and those who succeed will receive praise and respect from their peers for it because of the engineered concept courage*'s widespread use. Meanwhile, before the introduction of courage* – or if the concept had not been engineered at all – resisting oppression would have been considerably harder and less rewarding. There was no norm or standard available to follow and emulate, and one's community was less likely to recognise it as a good. Consequently, the fact that there is a new concept of courage* is a necessary condition for praising and emulating it as a virtue, and its widespread use makes it more likely that people act courageously*, thus multiplying real instances of courage*.

This second mechanism also occurs with extant virtue concepts or conceptual analyses that make a more precise conception available. Conceptual analysis might indeed lead to more adoption and emulation because the analysed virtue concept will be less contradictory and easier to explain and apply, just like with engineered concepts. However, adoption and emulation are likely considerably weaker with analysed instead of engineered concepts because the latter will be more salient through their difference from the folk conception.

---

[3] Compare this to accounts of how the introduction of the term "sexual harassment" only ever made the phenomenon tractable (Fricker, 2007, pp. 150–151).

Further, analysed concepts will only marginally differ extensionally or definitionally from folk concepts, if at all. This is liable to lead to confusion between the folk concept and the analysed concept because they will function interchangeably in most cases. It will therefore often be unclear what exactly we praise or emulate in that case. Analysing concepts will also not bring about any change in *which* virtues we praise and emulate.

The third mechanism through which a successfully engineered virtue is likely to be realised and proliferate is intentional cultivation. This is the planful, collective counterpart to the second mechanism. Institutions, like individuals, have interests in particular virtues. An institution whose goals align with an engineered virtue concept's purposes would arguably have an interest in fostering the new virtue. These fostering mechanisms are manifold: For example, there can be education emphasising the engineered virtue, propaganda promoting it, or policies supporting and promoting individuals exhibiting it.

Children's films frequently show and praise virtues. What is presented as courageous in a children's film will differ, depending on whether it is a pre-courage* or a post-courage* film, changing the models that are given to its viewers. Or take the German *Bundeswehr* (armed forces): It conceives of its soldiers as "citizens in uniform" who should pursue their civic and moral duties in addition to their military duties. Arguably, the concept courage*'s purposes match this ideal better than the more mercenary original concept of courage that does not include any civic purposes. Consequently, the institution might take diverse active measures to promote courage* within its ranks. In that way, the successful implementation of a concept with purposes that align with an institution's goals may lead to this institution's actively fostering this virtue.

Naturally, analysed and folk virtues are already being cultivated. (Snow, 2014; Kristjánsson, 2015) In that respect the two methods overlap – virtue cultivation is indifferent to which virtues it does cultivate. Still, it may be more valuable to cultivate intentionally engineered virtues rather than analysed virtues: The only case where cultivating engineered virtues could in principle not do morally and epistemically better than fostering analysed virtues is if our folk morality, and the concepts of virtue grounded in it, were already near perfect, in which case engineering virtue concepts

would be useless. As said, I am sceptical that our folk morality and past authorities already use the perfect virtue concepts and that we now only need to discover their precise nature and iron out the kinks through conceptual analysis.

Note that not all virtue ethics is conservative. For example, the virtue of epistemic justice and its counterpart vice of epistemic injustice (Fricker, 2007) are both a product of conceptual engineering in all but name. Epistemic injustice is a revisionary differentiation of the more general concept of injustice. Deflated credibility and missing hermeneutical resources did not count as injustices pre-2007. Instead, they were simply considered to be epistemically bad states. Fricker (2007, p. 5) also urges that we not focus on our (folk) intuitions about injustice, but instead on the wrongs that someone may be inflicted – she replaces the concept's principles, i.e. intuitions, with its purposes, i.e. reducing wrongs. (Fassio and McKenna, 2015)

This example also illustrates the fruitfulness of engineering virtue concepts: The virtue of epistemic justice as a social kind has become a subject of considerable discussion. The introduction of epistemic justice as a concept has generated widespread use of the concept and an awareness of the virtue. Epistemic justice is now literally a thing in social practice, and there have already been proposals to cultivate it more systematically. (Newbigging and Ridley, 2018)

Virtue constructionism can make great contributions to moral progress: Successful conceptual engineering allows us to improve virtue concepts more quickly than conservative virtue ethics because the latter remains tied to folk morality. Granting the success of the engineering project, these better virtue concepts are adopted by the community, changing the use and conception of the term. Given the mechanisms that I argued for above, our successfully engineered virtue concepts will also be realised, emulated, and propagated into real virtues by the epistemic community.

To successfully engineer virtues and thereby enable moral progress, conceptual engineers need to be transparent about the theoretical and practical purposes which they aim to satisfy with their engineered virtue concepts. That is, there needs to be a broad social debate about the purposes which virtue concepts should satisfy.[4] This

---

[4] [EK] is currently developing an account of participative conceptual engineering.

Jakob Ohlhorst

deliberation about the purposes of engineered virtue concepts can also rely on traditional virtue ethics because it possesses a store of knowledge about how different virtues fulfil certain practical and theoretical purposes. Finally, the concepts that are engineered must be designed in a way to ensure concept uptake: They must be easy to understand and their theoretical and practical advantages must be readily apparent, otherwise nobody will use the concept.

There are also some plausible constraints that serve to guarantee that our engineered concepts stay on topic. If we engineered the concept of courage to an extent that it now designates a right to party, then we have stopped engineering a virtue concept. There are core purposes around which individual virtue concepts function: Among others, there is the purpose to designate stable character traits that explain excellent behaviour of a certain type. Indeed, concepts in general appear to be centred on purposes. Purposes therefore give a good criterion of continuity. (Rose and Nichols, 2020) For example, the concept of courage's core purpose will be to explain and encourage excellent behaviour in the face of threats. In the meanwhile, it can happen that we have reason to even modify a core purpose, as Koch (2021b) has shown. Nevertheless, the general constraint on virtue engineering will remain: Engineered virtue concepts must designate stable character traits that explain excellent behaviour.

5. Objections and Perspectives

One might object that engineering virtues puts the cart before the horse. Virtues are extremely complex normative character traits that have slowly evolved through cultural or even evolutionary history. Engineering and revising the concept of a virtue will inevitably get it wrong and have unforeseeable consequences. (Marques, 2020) For that reason, the best we can do is to conservatively start with the virtue concepts that we already have and try to refine and clarify them step by step – do conceptual analysis in other words.

I agree that constructing revisionary virtues may prove difficult and often frustrating. However, we are not forced to first engineer a novel virtue concept from the armchair and then present it to the world. Just as in ordinary engineering, we may develop prototype concepts and virtues that we subject to practical testing. In this case, the testing of the prototype virtue would be to try to emulate it in order to see its

consequences as well as to study instantiations of the prototype concept found in the wild. This engineering process will essentially be an accelerated intentional version of the cultural evolution that underlies our traditional virtues. Note also that constructionism about virtue concepts does not mean that we reject everything that we know about our traditional virtue concepts.

In a related manner, you might wonder whether we actually do *construct* any new virtues, when engineering them. May we not simply discover traits that fit our practical and theoretical purposes in the wild, and recognise them as virtues?[5] Is this not what Zagzebski does with her virtue exemplarism? We admire an exemplar and we consequently designate and emulate their character traits as virtues. I considered a related, semantic, version of this worry about virtues as natural or basic moral kinds in section 3. I see three options here:

On the first option, we have to distinguish two cases. The case a) is that we rely on past exemplars. We take our virtues from the great minds of eras bygone – Confucius, Jesus, or Zara Yaqoub – this is reasonable because great virtue will be exceedingly rare. It is also what Zagzebski (2010) does. In this case, we do not actually discover their virtues and designate them directly. Rather, we rely on accounts of their deeds and characters – the tradition. This tradition clearly is filtered through the folk morality of centuries, and little of their original, actual character will have survived. I take this to be a simple case of conceptual analysis with historical, agent-based constraints.

Case b) is more interesting. What happens if we come across a contemporary live exemplar in virtue? A sceptic might doubt that this is possible – even saints have skeletons in their closets. But let us assume that, say, the pope is a living moral exemplar whom we all admire and we can recognise and emulate his character traits as model virtues. In this case, either our analysis is not revisionary at all, and we are just using the pope as an illustration of the folk virtue concepts that we had already, as manifested by our intuitions about the pope's admirability. Or alternatively, the pope's virtuous traits are in some way revisionary with respect to folk morality which brings us to our second option.

---

[5] I want to thank an anonymous referee for raising this point.

Jakob Ohlhorst

The second option is that Zagzebski proposes a scientific explication. We replace our old folk virtue concept with the exemplar's virtues. It is a bit like explicating the meter by designating a meter-prototype. Carnapian (1962) explication is counts as a precursor to contemporary conceptual engineering because it revises the folk concept instead of analysing it. (Kitsik, 2020) Explication takes inexact contradictory folk concepts and replaces them with precise scientific concepts. An example by Carnap is that the concept fish has been explicated to exclude whales, which is revisionary and not just analysis. (Carnap, 1962, p. 6)

In that case, we also need to engineer a precise and revisionary definition of these virtuous traits, otherwise the concept will deteriorate through tradition again because our prototype is not immortal. I believe that this is the method that Zagzebski technically seen intended. In that sense she would defend a revisionary method of virtue ethics – however, this reading clashes with her account of how we find exemplars, namely through intuitions about admirability.

As a third option, we might also take the exemplar to be someone whose character traits promote the practical and theoretical purposes that we desire of our constructed virtue concepts. In this case, the exemplarist method would be a heuristic that helps us to study how virtues that we aim to engineer behave in the wild. In this case, the exemplar is secondary to the concept's purposes. Nevertheless, there will be many exemplars whose character traits are potential candidates for a virtue to be engineered according to our purposes.

A further objection that speaks in favour of conservatism over constructionism is to say that virtue might be abusively engineered. Think a *1984* style scenario where a malevolent dictatorship distorts the virtues into obedience or hatefulness and spreads them through propaganda to maintain government control. (Shields, 2021) I do not think that the risk of abuse should deter us, especially because the traditional virtues are not innocent either. On the one hand they themselves often arise from and served to maintain oppressive societies. Aristotelian virtues for example are the product of a slaveholder society. On the other hand, also contemporary society can use traditional virtues for oppressive ends. Consider for example the Chinese government's (distorting) appeal to Confucian virtues in its propaganda.

This point relates to another objection that constructionist virtue ethics would open the gates wide for relativism. I do not think that constructionism opens this gate; it has already been open given conservative virtue ethics' own traditionalist and culture-relative roots. Meanwhile, virtue engineers are transparent about which purposes, i.e. values and constraints, they aim to realise. The theoretical and moral goals of conceptual engineering are more transparent and responsive to reason than the folk morality on which conceptual analysis relies because they are explicit and can be debated.

In an alternative avenue, Nussbaum (1993) responds to the generalised relativism-accusation that virtue is not bound to culture-specific practices but rather universal human experience. A good virtue theory explicates how the cogniser should deal and cope with these universal experiences. Virtue is first designated as whatever constitutes a good handling of such a situation, without sensitivity to folk morality, and then padded out into a full virtue concept. In that sense, Nussbaum's proposal can be read as a non-conservative eudaimonistic approach.

Nussbaum's position marks an intermediary position between conservative virtue traditionalism and radical virtue constructionism. Just like the explicatory reading of Zagzebski's approach, it is also a Carnapian (1962) explication. Consequently, also Nussbaum's approach is revisionary and therefore a type of conceptual engineering.

Explication as a method differs most notably from radical constructionism in Carnap's *similarity*-requirement: Most uses of the explicated concept must overlap with most uses of the original concept, an explication revises a given concept to be precise. (Carnap, 1962, p. 5) In constructionist virtue engineering, this similarity requirement is completely shifted to core purposes.

Explication is the only method of conceptual engineering that would be applicable to virtues, if we take them to be basic moral kinds (Sayre-McCord, 1997) rather than social kinds that are sensitive to complex moral, social, and biological facts. Even in the case of explication, the methodology is more "constructive" than Zagzebski or Nussbaum might imagine: Take the amount of conceptual construction that went into the development of the natural kind concept *periodic element* which only overlaps in small parts with its precursor concepts.

Jakob Ohlhorst

Consequently, we have differentiated the methodology of virtue engineering: On the moderate end, there is explicatory virtue engineering which aims to replace the vague folk virtue concepts with precise and exact virtue concepts. Nussbaum (1993) and Zagzebski (2010) can be read as pursuing this avenue. On the radical end, there is constructionist virtue engineering, which posits theoretical and practical purposes that a virtue concept should satisfy and a revisionary concept is designed on this basis. I would argue that Fricker (2007) does the latter by developing the novel concepts of epistemic justice and injustice.

As a final consideration, I want to examine directions for expansion of the project. The first direction of expansion is into virtue epistemology. In some domains, virtue ethics and virtue epistemology share a considerable overlap as the example of epistemic justice shows. It is both a moral and an epistemic virtue. But there are also purely epistemic virtues[6] that are subject to their own dynamics. I suspect that especially social epistemic virtues will be a thankful object for virtue engineering.

A field where the engineering of moral and epistemic virtue is particularly called for is the context of social media. The structure and dynamics of interaction on social media is so far from the contexts from which our traditional virtues arose that they have become useless or even harmful in these contexts. A solution for this problem is to radically revise our traditional virtues for social media contexts or to develop and introduce radically new virtues.

Vice concepts can and should also be engineered. This would sharpen our focus on peculiar kinds of human weaknesses and help us to suggest remedies to address them. For example, there may very well be vices to be engineered that relate to systemic threats like global warming. (*cf.* Sandler, 2013)

6. Conclusion

In this paper, I argued that virtue ethics should make a methodological turn. Instead of conservatively analysing the exact nature of the extant virtues, virtue ethicists should become virtue engineers, revise old virtues, and develop new virtues. Note that this is not an argument *against* the traditional method of conceptual analysis which still

---

[6] For example, virtue reliabilism (Greco, 1999).

Jakob Ohlhorst

may be valuable, nor do I want to claim that all virtue ethics until now has been conceptual analysis. Rather, this is an argument *for* virtue engineering. I proposed an appropriate methodology for this, conceptual engineering, and argued that this methodology is particularly effective. I also distinguished two styles of engineering virtues: Explicating them and constructing them. My principal argument for this methodological turn is that constructionist virtue ethics would accelerate moral progress because we can expressly design concepts to satisfy the theoretical and practical purposes that our virtue concepts should fulfil.

References

Alfano, M. (2013) *Character as moral fiction*. Cambridge, England: Cambridge University Press.

Annas, J. (2011) *Intelligent Virtue*. Oxford University Press. doi: 10.1093/acprof:oso/9780199228782.001.0001.

Anscombe, G. E. M. (1958) 'Modern Moral Philosophy', *Philosophy*, 33(123), pp. 1–19.

Bentall, R. P. (2003) *Madness Explained: Psychosis and Human Nature*. London: Penguin Books.

Burgess, A. and Plunkett, D. (2013) 'Conceptual Ethics I', *Philosophy Compass*, 8(12), pp. 1091–1101. doi: 10.1111/phc3.12086.

Cappelen, H. (2018) *Fixing Language*. Oxford: Oxford University Press. doi: 10.1007/BF01696331.

Carnap, R. (1962) *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Chalmers, D. J. (2020) 'What is conceptual engineering and what should it be?', *Inquiry*, pp. 1–18. doi: 10.1080/0020174X.2020.1817141.

Conee, E. and Feldman, R. (1998) 'The Generality Problem for Reliabilism', *Philosophical Studies*, 89(1), pp. 1–29. doi: 10.1023/A:1004243308503.

Doris, J. M. (2002) *Lack of Character*. Cambridge, England: Cambridge University Press.

Fassio, D. and McKenna, R. (2015) 'Revisionary Epistemology', *Inquiry*, 58(7–8), pp. 755–779. doi: 10.1080/0020174X.2015.1083468.

Floridi, L. (2011) 'A Defence of Constructionism: Philosophy as Conceptual Engineering', *Metaphilosophy*, 42(3), pp. 282–304. doi: 10.1111/j.1467-9973.2011.01693.x.

Fricker, M. (2007) *Epistemic Injustice*. Oxford: Oxford University Press.

Greco, J. (1999) 'Agent Reliabilism', *Philosophical Perspectives*, 13, pp. 273–296. doi: 10.1111/0029-4624.33.s13.13.

Haslanger, S. (2000) 'Gender and race: (What) are they? (What) do we want them to be?', *Nous*, 34(1), pp. 31–55. doi: 10.1111/0029-4624.00201.

Jackson, F. (2000) *From Metaphysics to Ethics*. Oxford, UK: Oxford University Press. doi: 10.1093/0198250614.001.0001.

Kitsik, E. (2020) 'Explication as a strategy for revisionary philosophy', *Synthese*, 197(3), pp. 1035–1056. doi: 10.1007/s11229-018-1774-z.

Koch, S. (2021a) 'Engineering what? On concepts in conceptual engineering', *Synthese*, 199(1–2), pp. 1955–1975. doi: 10.1007/s11229-020-02868-w.

Koch, S. (2021b) 'Why Conceptual Engineers Should Not Worry About Topics', *Erkenntnis*, (0123456789). doi: 10.1007/s10670-021-00446-1.

Kristjánsson, K. (2015) *Aristotelian Character Education*. New York City: Routledge.

Liu, C. (2021) 'Slurs as Illocutionary Force Indicators', *Philosophia*, 49, pp. 1051–1065. doi: 10.1007/s11406-020-00289-0.

MacIntyre, A. (1981) *After Virtue*. Notre Dame: University of Notre Dame Press. doi: 10.4324/9781912281954.

Marques, T. (2020) 'Amelioration vs Perversion', in Marques, T. and Wikforss, Å. (eds) *Shifting Concepts*. Oxford: Oxford University Press, pp. 260–284. doi: 10.1093/oso/9780198803331.003.0014.

Nado, J. (2021) 'Re-engineering knowledge: a case study in pluralist conceptual engineering', *Inquiry*, pp. 1–26. doi: 10.1080/0020174X.2021.1903987.

Neufeld, E. (2019) 'An Essentialist Theory of the Meaning of Slurs', *Philosophers' Imprint*, 19(35).

Neufeld, E. (2022) 'Psychological essentialism and the structure of concepts', pp. 1–20. doi: 10.1111/phc3.12823.

Newbigging, K. and Ridley, J. (2018) 'Epistemic struggles: The role of advocacy in promoting epistemic justice and rights in mental health', *Social Science and Medicine*, 219(October), pp. 36–44. doi: 10.1016/j.socscimed.2018.10.003.

Nussbaum, M. (1993) 'Non-Relative Virtues: An Aristotelian Approach', in Nussbaum, M. and Sen, A. (eds) *The Quality of Life*. Oxford: Oxford University Press, pp. 242–269. doi: 10.1093/0198287976.003.0019.

Plato (2000) *The Republic*. Edited by G. R. F. Ferrari and T. Griffith. Cambridge: Cambridge University Press.

Putnam, H. (1973) 'Meaning and Reference', *The Journal of Philosophy*, 70(19), pp. 699–711.

Reeve, C. D. C. (2012) 'Aristotle's Philosophical Method', in Shields, C. (ed.) *The Oxford Handbook of Aristotle*. Oxford: Oxford University Press, pp. 150–170. doi: 10.1093/oxfordhb/9780195187489.013.0007.

Rose, D. and Nichols, S. (2020) 'Teleological Essentialism: Generalized', *Cognitive Science*, 44(3). doi: 10.1111/cogs.12818.

Sandler, R. L. (2013) 'Environmental Virtue Ethics', *International Encyclopedia of Ethics*. Edited by H. Lafollette. Oxford: Blackwell Publishing Ltd. doi: 10.1002/9781444367072.wbiee090.

Sawyer, S. (2021) 'Concept pluralism in conceptual engineering', *Inquiry*, pp. 1–26. doi: 10.1080/0020174X.2021.1986424.

Sayre-McCord, G. (1997) '"Good" On Twin Earth', *Philosophical Issues*, 8, pp. 267–292. doi: 10.2307/1523011.

Shields, M. (2021) 'Conceptual domination', *Synthese*, 199(5–6), pp. 15043–15067. doi: 10.1007/s11229-021-03454-4.

Slote, M. (1995) 'Agent-Based Virtue Ethics', *Midwest Studies in Philosophy*, 20(1),

pp. 83–101. doi: 10.1111/j.1475-4975.1995.tb00306.x.

Snow, N. E. (ed.) (2014) *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology*. Oxford: Oxford University Press.

Swanton, C. (2003) *Virtue Ethics*. Oxford: Oxford University Press. doi: 10.1093/0199253889.001.0001.

Thomasson, A. L. (2022) 'How should we think about linguistic function?', *Inquiry*, pp. 1–32. doi: 10.1080/0020174X.2022.2074886.

Zagzebski, L. (2010) 'Exemplarist vitue theory', *Metaphilosophy*, 41(1–2), pp. 41–57. doi: 10.1111/j.1467-9973.2009.01627.x.

Zavaliy, A. G. and Aristidou, M. (2014) 'Courage: A Modern Look at an Ancient Virtue', *Journal of Military Ethics*, 13(2), pp. 174–189. doi: 10.1080/15027570.2014.943037.