Extending Intellectual Humility

Mara Neijzen

This is an unpublished conference paper for the 9th Annual Jubilee Centre for Character and Virtues conference at Oriel College, Oxford University, Friday 10th – Sunday 12th September 2021.

Jubilee Centre for Character and Virtues

University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom

T: +44 (0) 121 414 3602 F: +44 (0) 121 414 4875

E: jubileecentre@contacts.bham.ac.uk W: www.jubileecentre.ac.uk

Abstract

Responsibilist epistemic virtues – acquired traits such as intellectual humility and diligence – are becoming increasingly relevant for online knowledge acquisition and social interaction, due to the prevalence of misinformation and heated debates on the web. This paper proposes a method to enhance responsibilist virtue by applying the extended cognition framework to epistemic virtues. This results in an example of extended intellectual humility, described as the coupled dynamical system of a hypothetical agent and a hypothetical app, where the design of the latter is based on empirically supported methods to increase awareness of one's cognitive limitations.

I. Introduction

Responsibilist virtue epistemology (e.g., Baehr 2011; Code 1984; Zagzebski 1996) is a growing field which focusses on acquired intellectual traits and attitudes, such as integrity, open-mindedness, and intellectual humility, as their epistemic virtues. This is in contrast to responsibilism's older sibling, reliabilist virtue epistemology (e.g., Greco 1999; Sosa 2017), which describes the virtues as faculties and competences (e.g., perception and memory) which reliably lead to true beliefs. One of the main differences between responsibilism and reliabilism, is that the former postulates virtues as necessarily acquired personal qualities. Nevertheless, there is no robust account of responsibilist virtueacquisition, apart from mentor-based proposals (most notably Baehr 2013, who presents practical guidelines for educators to helps students acquire intellectual virtue). Moreover, following e.g. Baehr and Zagzebski, responsibilist virtues are admirable traits that agents *ought* to acquire. So, an interdisciplinary account of virtue-acquisition can be a useful addition to responsibilist virtue epistemology.

An interdisciplinary project on digitally enhancing epistemic character virtues can give novel, empirically supported insights on acquiring and enhancing these valued traits. In this paper I propose one way to enhance intellectual humility (an epistemic virtue which entails an agent's understanding of their cognitive limitations), by making use of our existent integration with our phones, modelling the agent's intellectual virtue through extended cognition. As will be discussed in the next section, according to the extended cognition thesis, cognition is a dynamical process consisting of feedback loops which encompass brain, body, and occasionally external objects (Clark & Chalmers 1998; Clark 2007; 2011).

Extending intellectual virtue is not a novel idea. While this line of enquiry has thus far only been applied to faculty virtue, most notably memory, it has been argued that responsibilist virtues can also

extend beyond the brain and body. Battaly (2018) not only argues in favour of the possibility of extended cognition and responsibilist virtue coming together, she also presents an example of extended open-mindedness.

Open-mindedness is an intellectual virtue which entails the disposition to set one's own cognitive standpoint aside in order to take other, possibly opposing, views seriously. Battaly presents two hypothetical openminded agents, Ivan and Olga, who are both employed at the government to consider and suggest changes to current policies. Where Ivan generates and considers appropriate alternatives in his head, Olga does so with pen and paper: whenever Olga recognises that a situation requires open-mindedness, she grabs her notebook and a pen to employ the "circle of viewpoints" method she learned in school. Both Olga and Ivan get similar results regarding policy. Moreover, Battaly asks us to imagine that the example adheres to the *parity* and *coupling* characteristics – requirements for extended cognition which will be discussed in the next section. So, Battaly names this an example of extended epistemic virtue, opening the door for other analyses surrounding other intellectual virtues and extension.

To further this project, the following section of this paper outlines extended cognition and the elements required for calling a process extended. I then outline a worry with regards to cognitive extension. This is the demand for a *mark of the cognitive,* which entails that defining a cognitive process as extended must include a definition of, or requirements for, calling a process cognitive in the first place. While this worry has been discussed within philosophy of mind, where extension proponents provide possible counterarguments, not all of these mitigations are applicable to intellectual virtue. Therefore, in the third section I suggest a *mark of the virtuous*. This provides us with the required characteristics of extended intellectual virtue, which will be applied in the fourth section to describe an instance of extended intellectual humility. By coupling an agent to an app whose functioning is

based on empirical studies on how to improve intellectually humble thought, the agent's intellectual humility becomes extended through, and improved by, the app. While this is only one example of how intellectual virtue can be improved through methods gained from interdisciplinary research, it might open the door to further interactions between responsibilist epistemology and other fields, on top of furthering the debate within virtue epistemology.

II.Extension

Since its conception more than twenty years ago, the theory of extended cognition has sprouted several theories and transformations. They can be captured in the three (generally acknowledged) waves of extended cognition: the first wave which focusses on *functionalism*, the second wave which focusses on *complementarity*, and the third wave which conceptualises cognitive processes as *extended dynamic singularities* (Gallagher 2018; Wheeler 2019). As the topic of extended cognition has become more popular over the years, the three waves as described do not reflect the nuances of the views contained in the topic – most of which involving various formulations of the necessary conditions for extended cognition. Every wave, however, can be described as having certain requirements for extension to occur. To present a blanket account of extended virtue, then, all demands discussed will be taken into account when describing extended intellectual virtue in the next section.

The first wave is extended cognition as it was first introduced by Clark and Chalmers (1998). It relies on the parity principle, which states that if a process would be categorised as cognitive when it occurs inside the skull, then it should also be categorised as cognitive when it occurs (partly) outside the skull. This is often explained through *Otto's notebook*: Otto, a man suffering from memory problems, always carries around a notebook in which he writes every belief he forms. He always employs this notebook to remember his beliefs. When he wants to go to the Metropolitan Museum

of Art to see an exhibition, he simply refers to his notebook to find the address. This is in contrast to another agent, Inga, who has a well-functioning memory and thus organically remembers where the MET is before she sets off to see the exhibition. In this example, Otto's notebook and Inga's brain serve the same function, namely containing, or forming the basis of, dispositional beliefs, and both are systematically and reliably employed to serve this function, guiding both agents' behaviour. While Otto's notebook and Inga's brain do not store beliefs in the same way (as memories do not function as static sentences on a page – more on this later), the function of storing beliefs for future application is shared. As such, it is argued, Otto's memory itself extends beyond the brain to his notebook. This extension of mind only occurs when certain *trust and glue* conditions apply, i.e., "the resource be reliably available and typically invoked.", "any information thus retrieved [must] be more-or-less automatically endorsed" and when "information contained in the resource [is] easily accessible as and when required" (Clark 2010, 46). Accounts adhering to these requirements fall within the first wave of extended cognition.

These trust and glue conditions are important for this extended cognition account, as they prevent *cognitive bloat*. This would occur when the mind or cognitive processes are over-extended, for example when all instances where an agent employs a tool for a cognitive task are included. After all, even though we regularly employ search engines to come to certain beliefs, we do not automatically endorse the information we retrieve as we would if the beliefs were attained through perception or memory. Likewise, we might automatically endorse the information contained in a number of books on our bookshelves – books we have read but whose content has been mostly forgotten, yet the information in these books is not easily accessible whenever it is required. We do not take these books (e.g., an atlas, a dictionary, an autobiography, and numerous fiction books that convey an important lesson regarding the human condition) everywhere we go. Nor do we typically invoke them when we want to recall certain facts, as computers, smartphones, and knowledgeable friends

are also available. If these conditions were not required, cognition could extend to many sources of information (e.g. books and search engines) and useful cognitive tools (e.g., calculators) – an effect that should be avoided due to its unintuitive nature.

The second wave ascribes less importance to the parity principle, instead focussing on how cognition can be supplemented and enhanced by processes which are either complementary to (e.g., Sutton 2010) or integrated with (e.g., Menary 2007) existent cognitive abilities. For instance, when calculating a sum, an agent may use pen and paper to aid the process. Where first wave theorists would say that this could be a case of extension, as when the sum was calculated in the head we would call it a cognitive process, Menary (2007) and other second wave theorists suggest that the cognitive process is *enhanced* by the pen and paper. In other words, it is possible that the agent would not have been able to calculate this sum in their head, yet they manage to do so by extending the process to the pen and paper. So, extended processes do not need to resemble intracranial cognitive processes in either function or structure to count as examples of extended cognition. Note that the parity principle is not necessary for extension in this section wave – in fact, it is for a great part defined by its rejection of the necessity of parity.

However, that this rejection of focus on the parity principle might rest on a specific interpretation of this principle, i.e., that parity only occurs when the external mechanism functions *in the same way* as intracranial processes. Instead, the parity principle was merely meant to test intuitions on what we might define as a cognitive process when we do not necessarily localise it in the brain (Kiverstein and Clark, 2009). So, when we store beliefs in a notebook, calling this practise extended cognition does not entail that we also "store" beliefs in the head, as easily and objectively retrieved as when we read our diary. Inga, who has an average capacity for memorisation, can also use a notebook, in the same way Otto does, to reliably enhance her memory, storing facts she would have otherwise

forgotten. In other words, complementary and parity are compatible. So, our extended cognition account can (and ought to) meet the demands of both the second and first wave to propose the extended virtue example.

The third wave of extended cognition focusses on the way brain, body, and environment interact with one another, stating that extension occurs when the intracranial and external processes are non-linearly coupled so that neither process can be modelled without including the other. Cognition is conceptualised as a dynamical process, consisting of feedback loops which encompass brain, body, and occasionally external objects (Clark 2007, §2). For the cognitive process to extend to an external object, the brain, body and object must mutually interact with one another, i.e., the brain affects the body, which affects the object, which affects the brain, and so on (Chemero 2009). We no longer necessarily interpret extension primarily through function and reliability, but rather through explaining the dynamical interaction between different systems – if these systems are coupled non-linearly, then these different systems constitute one process.

The nonlinear coupling between these systems entails that neither system can be modelled apart from the other. For *Otto's notebook*, it means that when Otto forms a belief, he writes this down in his notebook, where he sees other beliefs which might relate to the new belief, contextualising his knowledge. He regularly updates his beliefs when new information presents itself, and the everchanging contents of his notebook shape his thoughts and actions, which in turn affect the contents of the notebook. As such, Otto's brain, body and notebook are coupled, all constituting the cognitive process of memorisation.

The purpose of this paper – to provide an account of extended intellectual virtue – should entail that the suggested account should, in principle, be acceptable to anyone who accepts the possibility of extended cognition. As such, the account of extended virtue will adhere to the elements discussed in

all three waves. This entails the parity principle, the trust and glue conditions, supplementation through complementarity, and non-linear coupling. However, before applying these conditions to our example of intellectual humility, one line of critique against extension accounts is discussed.

Mark of the Cognitive

While this paper employs extended cognition as a useful tool with which to model coupled processes, extended cognition remains a relatively controversial stance. Objections to the thesis often come down to a disagreement over the boundaries of cognition through differing initial definitions (e.g., when cognition is defined as an activity of the brain, extension is a priori excluded). Yet some critiques are of the applicability of extended cognition to commonly used examples, without such (conscious) reliance on preconceived definitions of the mind. Perhaps the most wellknown line of such critique entails the *coupling-constitution fallacy* and the want of a *mark of the cognitive* (Adams and Aizawa 2001; Wheeler 2019).[1]

The coupling-constitution fallacy entails the critique that proponents of extended cognition mistake causal relations between x and y for x partly constituting y. Concretely, Adams and Aizawa (2001) take Clark and other extension proponents to mistake an agent's cognitive process that occurs in the brain, being causally coupled to an object such as a notebook, for the agent's brain, body, and notebook all constituting cognition.

[1] As the former argument entails the latter, I will discuss them as one. Moreover, this will merely entail a short overview of the arguments and Clark's response – I stay agnostic on whether either the arguments or Clark's responses are victorious.

This is where the notorious jab directed at Clark comes in: "Why did the pencil think that 2 + 2 = 4? Because it was coupled to the mathematician."

To avoid this fallacy, Adams and Aizawa suggest the requirement of a mark of the cognitive – a definition or set of characteristics that allows us to recognise cognitive processes. Following Wheeler (2019), we can formulate Adams and Aizawa's arguments in two categories. First, they argue that a mark of the cognitive is needed in order to claim that a cognitive process is extended. Second, they require an argument for what such a mark of the cognitive may entail. The coupling-constitution fallacy falls in the first category of critique – a mark of the cognitive is needed in order to call an extended process cognitive, otherwise it could be a cognitive process causally coupled with some external object or process.

The first category of the critique is not accepted by Clark (2008) as a legitimate worry. He takes Adams and Aizawa to argue that some processes or bearers of processes are cognitive, while others are not. Recall the joke placing cognition in the pencil – this would imply that objects, e.g., the brain, neurons, pencils, or hands are cognitive. Yet, to cite Clark (2008: 87), it is of course "crazy to think that a V4 neuron thinks, and it is (just as Adams and Aizawa imply) crazy to think that a pencil might think. […] [T]his talk of an object's being or failing to be "cognitive" seems almost unintelligible when applied to some putative *part or aspect* of a cognitive agent or of a cognitive system. What would it mean for the pencil *or* the neuron to be, as it were, brute factively "cognitive"?". In other words, while certain things are uncontroversially constitutive of cognitive processes, these things need not be "cognitive" themselves.

However, the critique also applies to the kinds of things that can be *bearers* of cognitive processes. According to Adams and Aizawa (2001), only things with certain characteristics can be part of a cognitive process, as only those things have the capacity to bear cognitive processes. They argue that

cognitive processes must be able to contain or process non-derived representational content. Cognitive representations are non-derived (i.e., their content is directly shaped and informed by experience, without the influences of man-made symbols, words, or other forms of representation), in contrast to linguistic concepts which get their content from social application.[2] When we apply this to Otto's notebook, then, this notebook cannot be a bearer of cognitive processes, as Otto fills the diary with linguistic representations with socially derived content, rather than non-derived representations. Clark responds with clarifying his functionalist argument: it can be argued that Otto's writings have the same function as cognitive representations of what the written words and sentences entail. In other words, while Inga uses cognitive representations to recall information,[3] Otto uses the representations on paper in the same way. If Adams and Aizawa argue against this possibility, they do not remain neutral with regards to the limits of cognition before defining its mark. In other words, their argument is question begging. When we accept this argument, then, the mark of the cognitive need not include specifics on what can be the bearers of cognitive processes. Secondly, Adams and Aizawa argue that a mark of the cognitive should also describe the characteristics of cognitive processes: cognitive process must adhere to certain characteristics found by empirical studies on how the mind, or in this case memory, works. These characteristics include chunking (where novel information is broken down and grouped into larger units of previously acquired information or concepts), priming (where previous stimuli influence successive stimuli), etc. Clark rejects this proposal as well, as the mentioned characteristics are anthropocentric and, again, question begging. First, it is questionable that all the characteristics Adams and Aizawa propose as necessary for cognition are applicable to the capacities of other earthly organisms – which would

[2] In this paper I will not discuss the problem of representations – what they entail and whether they are necessary for cognition.

[3] Or whichever purely organic explanation of memorising is preferred by cognitive scientists and other researchers.

certainly be desirable seeing the growing field of animal and even plant cognition. Moreover, their suggestion cannot account for the hypothetical event where we come across an alien species which, while intelligent, does not exhibit cognitive processes which function similarly to ours. It would be odd to deny that these processes are cognitive, even if they do not mirror our own. Besides this implied anthropocentric view, Adams and Aizawa's characterisation of cognition also begs the question, as the required qualities seem to be derived from how *our brain* functions. As extended cognition proponents aim to remain neutral towards the boundaries of cognition, basing these boundaries on what occurs in the human brain should not be the starting point.

This has merely been a very short overview of the extension discussion, without any stance on how cognition functions and whether a mark of the cognitive is needed in the first place. However, if we want to extend intellectual virtues, we may run into similar criticisms. For this reason, the following section will discuss whether a mark of the intellectually virtuous ($virtuous_i$) is needed and, if so, what this would entail.

III. Mark of the Virtuous

If a mark of the cognitive is thought to be needed for extended cognition, something similar would be the case for extended virtue. After all, the intellectual virtues are complex traits and attitudes which require many processes usually deemed cognitive (perceiving, attending, reasoning, etc.). While we have seen Clark's arguments against the need for a mark of the cognitive, these arguments may not all apply to virtue, especially when the argument refers to avoiding anthropocentricity. After all, when we follow Aristotle (NE; EE), both moral and intellectual virtues are species-specific excellences, making anthropomorphism more like a requirement than a problem. When it comes to

neo-Aristotelian moral virtue, the foundation for accepting certain virtues is often similarly species specific, describing the virtues as qualities that make one a good example of a human being (Foot 2002). Moreover, responsibilist epistemic virtues are also conceptualised as how we, specifically, as limited, social creatures, can be epistemically responsible (Code 1984). So, we are merely speaking of intellectual virtues and what they entail *for us*. If this excludes Clark's counterarguments from applying to extended virtue, then the suggestion of extended intellectual humility may find itself victim to the coupling-constitution fallacy, where the tool and agent merely causally affect one another, unless we can present a *mark of the virtuous*.

We can surely apply Clark's idea that something need not be inherently cognitive to be part of a cognitive process, by stating that something need not be inherently virtuous to be part of a virtuous process. After all, the prefrontal cortex tends to be a constitutive part of an intellectually virtuous process, as, for example, it is involved in processes that require executive function – the competence partly constituting activities such as planning and decision-making, yet it seems odd to call this cortex, or the individual neurons constituting it, virtuous.

However, recall that this critique also applies to whether certain things can be bearers of cognitive, or rather, virtuous processes. So, recall Clark's argument against requiring a specification of what can be a bearer of cognitive processes. His argument relies on extending the mind or cognitive processes by function. When something, such as Otto's writings, performs the same function as otherwise intercranial processes, the characteristics of the brain which would normally perform this function should not matter. This also seems applicable to intellectual virtue. After all, when we apply this to Battaly's example, Olga can use her notebook and writings in the same way, as it has the same function, as Ivan who uses his internal reasoning. In other words, if we accept the parity principle, we can also accept that the bearers of cognitive or virtuous processes need not possess specific

qualities, as long as the relevant "external" bearers are functionally equivalent to the relevant organic bearers.

Moreover, even when we do not agree with Clark's response or the parity principle in general, note that it is seemingly impossible to find a characteristic that is applicable to all elements that may constitute any intellectually virtuous process. For instance, while all virtuous acts may be described as agential, not all elements (including, e.g., neurons and organs) that may constitute a virtuous process can be said to carry or express agency. As such, defining what characteristics a bearer of a virtuous process must possess is seemingly impossible. Of course, we may think that every virtuous process is also cognitive, which means that Adams and Aizawa's suggestion (i.e., able to carry or process non-derived representational content) should also be applied to the mark of the virtuous$_i$. Yet note that not all intellectually virtuous processes are *purely* cognitive, as they also entail behaviour. For instance, for an epistemically generous agent, communicating their knowledge to others is part of the virtuous process, yet their vocal cords surely are not able to carry or process non-derived representational content.

This just leaves the final aspect of Adams and Aizawa's argument for a mark of the cognitive. Namely, its content entailing certain characteristics of cognitive processes. Recall that this was also rejected by Clark as these characteristics were anthropocentric and question-begging. This response, however, does not seem to be applicable to intellectual virtue, due to the mentioned anthropocentric quality of intellectual virtues. Recall that the reason why Adams and Aizawa's characterisation of cognition is question begging according to Clark, is due to the characteristics being derived from how our brain functions. This method, however, is part of how responsibilists define and analyse the virtues – not necessarily by focussing on the brain, but certainly by focussing on what the virtues

entail given how we, as human beings, function. Clark's argument against the need for a set of characteristics to recognise certain processes as cognitive, is not applicable to responsibilist virtue.

Virtue responsibilists and cognition theorists alike, then, could refuse to accept the proposal of extended intellectual humility without a *mark of the virtuous*$_i$. If we accept that it is not necessary for a part of a virtuous process to be virtuous itself, then a mark of the virtuous need merely entail a set of characteristics of intellectually virtuous processes. The following section first describes the characteristics usually ascribed to intellectual virtue, after which a mark of the virtuous will be defined.

The characteristics

In order to formulate the characteristics of extended intellectually virtuous processes, I first summarise the commonly accepted characteristics of responsibilist virtue. While the characteristics of intellectual virtues are diverse, especially seeing the broad spectrum of traits and attitudes that are categorised as such (epistemic generosity; attention; curiosity; diligence), certain characteristics are relatively uncontroversially applicable to all relevant virtues. To keep in line with Battaly's suggestion of extended open-mindedness, I

summarise the characteristics she presents in this paper. However, some characteristics are irrelevant for defining the mark of the virtuous$_i$, while another ought to be specified further.

First, responsibilist virtues are *acquired*, in contrast to being innate. Whether they must be acquired actively (i.e., whether the agent needs to consciously acquire specific intellectual virtues through practise) or merely via agential actions (where the agent acquires or merely maintains the virtues by acting in line with them non-accidentally) is up for discussion (Baehr 2011). Second, they are

praiseworthy: when an agent possesses a responsibilist virtue or acts in line with it, this trait or act is admirable. This element, however, could either entail that the virtues are admirable traits as they make someone better as a person (Baehr 2011, 2016; Zagzebski 1996), or the quality could be reduced to the previous characteristic, i.e., that these virtues are acquired and therefore "up to the agent" to attain, thereby making attaining them praiseworthy (Baehr 2011). Third, "[r]esponsibilist epistemic virtues require dispositions of action" (Battaly 2018, 197), they are a *disposition to act* rather than mere intentions or motivations. While this characteristic could seem merely definitional, where responsibilist virtues are defined as attitudes or dispositions, this characteristic can also be interpreted as a stability requirement: a responsibilist virtue is not merely a disposition to act sometimes or in some situations. Rather, these virtues guide behaviour across temporal and situational conditions. In other words, once an agent has acquired an intellectual virtue, she will act on this virtue, whether it is while doing her job or when interacting with family.

Fourth, they are *personal*: "Responsibilists argue that the disposition to perform virtuous epistemic actions isn't sufficient for epistemic virtue. They think epistemic virtues and vices are personal qualities that tell us who we are as individual thinkers—an agent's virtues and vices express what she cares about and values" (idem). Responsibilist virtues, then, express an agent's values. These values may entail highly valuing truth or knowledge, or may refer to the agent's positive evaluation of the traits themselves. Moreover, these values motivate the disposition to act in line with the relevant virtue. As such, intellectual virtues require an epistemic motivation.[4]

The fact that the virtues are acquired and praiseworthy are not constitutive of the actual virtues; they are merely requirements to call a process or trait intellectually virtuous. The only two characteristics

[4] Battaly states that whether the virtues are *reliable* in attaining epistemic goods is still a topic of discussion. As this reliability would not affect the mark of the virtuous

we can include in our definition of the mark of the virtuous are that virtues are dispositions and personal traits that require epistemic motivation. So, we ought to focus on what extending one's epistemically motivated disposition entails. This requires a description of how dispositions work.

Dispositions are inclinations or tendencies to act in a particular way, and when it comes to responsibilist virtues, these acts are epistemic in nature, where the specifics rely on the virtue in question and the particular situation. For many dispositions, the tendency might be general. For example, Santa has a jolly disposition, which means that he generally acts jolly. The intellectual virtues, on the other hand, are often described as a mean between two vices. In Battaly's example, for instance, being openminded is the mean between being dogmatic and naïve (idem). So, when an agent generally acts open-mindedly, she is not consistently interpreting all other points of view as if they possess merit, as this would refer to being naïve. Rather, she interprets any given epistemic situation, responding with open-mindedness when the situation calls for it. After all, when we recall Olga, we see that she made the circle of views when confronted with possible policies, rather than whether she should buy peanut butter or read an alt-right blog, i.e., trivial or possibly epistemically harmful (Cassam 2019) acts. So, when it comes to intellectual virtue, the disposition should be in response to particular situations, namely those that call for epistemically virtuous action. Simply interpreting certain situations as epistemically loaded, then, partly constitutes the intellectually virtuous disposition, on top of it disposing the agent to act in a way that fits with the epistemic situation.

Note that this process and the reasons behind it need not be conscious, or even epistemically accessible to the agent. When an openminded and a closeminded person both engage in a conversation with someone who has different views, the openminded agent will likely listen to these views and seriously consider them, while the closeminded agent is likely to ignore or automatically

reject those views contrary to their own. When both agents are asked whether they found the other's views plausible and why, they might answer with a general reasoning which does not include a mention of open-mindedness.

From this discussion, we can define the mark of the virtuous$_i$ as being constituted by the following characteristics: it involves epistemically motivation, is attention-guiding, and entails the disposition to respond with an epistemic action. These characteristics are necessary for recognising an intellectually virtuous process, but it is up for debate whether they are sufficient. However, while it could be possible that this description leads to examples of processes resulting from traits or attitudes which are not currently deemed intellectually virtuous, this need not be a problem. After all, the relevant trait instigating the process must still be acquired and praiseworthy. If an acquired, admirable trait, based on epistemic motivation, which disposes an agent to epistemic behaviour, is discovered, I posit that this is rather a discovery of another intellectual virtue, instead of an argument against my case.

Extending virtue

With these characteristics at hand, recall the requirements for extension: for an intellectual virtue or virtuous process to be extended, it must adhere to the parity principle, the trust and glue conditions, supplementation through complementarity, and non-linear coupling. As we can see from this characterisation, we should still differentiate two forms of extension: extended virtue, i.e., the trait or attitude itself, or extended virtuous action; the behaviour resulting from having the virtuous disposition.

A virtuous act must result from an epistemic motivation and a stable disposition, which guides the agent's attention to the relevant aspects of the situation, allowing her to adequately respond with the

virtuous act. Focussing on virtuous acts leads to a myriad of examples where a virtuous agent regularly employs certain reliable tools to act on their dispositions. For one, recall the example of extending the cognitive process of calculation, where an agent employs pen and paper to solve a sum. With this in mind, we can imagine a diligent agent who regularly has to calculate percentages or other mathematical sums. As this agent is diligent, she would regularly, if not always, employ pen and paper or a calculator to arrive at and check her conclusions. This agent's diligent actions, when they adhere to the trust and glue conditions, are extended.

This is similar to Battaly's example of the openminded Olga. After all, openmindedness is the disposition to consider other points of view without letting one's own cognitive stance overrule these other points of view. So, where the disposition to act open-mindedly for Igor entails reviewing possible policies in his head, for Olga it entails employing her notebook to write down the different points of view with regards to the policy in question. Both Ivan and Olga recognise situations where open-mindedness is relevant, as their virtue directs attention to the relevant aspects of the situation, and, seeing their epistemic motivations, they employ either their reasoning or a notebook to reach their desired outcome. As such, Olga employs an external object to adequately act on her virtue, making her virtuous act extended, rather than her virtue itself.

The other possibility, extending intellectual virtue, entails extending an epistemically motivated trait or attitude which directs attention to the relevant aspects of a situation, allowing the agent to interpret relevant situations as epistemically challenging or as requiring epistemic action, which disposes agents to act in line with intellectual virtue within the context of the relevant situation.[5] While this certainly seems more complicated than extending epistemically virtuous actions, it is also

[5] Note that the disposition aspect, including its attention-guiding quality, still need not be a conscious process when the virtue is extended.

more in line with the overall aim of this paper: extending virtues to acquire or improve upon them. It is this second version that will be the focus of the remainder of this paper.

In short, for an *intellectual virtue* to be extended, the stable disposition to act in line with the virtue in question in response to an epistemically challenging situation, entailing epistemic motivation and attention guidance, must be extended. Moreover, for an intellectual virtue to be *extended*, the process must[6] adhere to the parity principle, the trust and glue conditions, supplementation through complementarity, and non-linear coupling.

IV. Extending intellectual humility

Intellectual humility is a disposition or trait which makes one understand and admit to one's intellectual limitations. As is the case with all virtues, intellectual humility also has a motivational component – when the agent comes to understand her intellectual shortcomings, she will be motivated to fix them. Additionally, the intellectually humble agent will respond with the appropriate affective states when confronted with her intellectual shortcomings (Whitcomb et al. 2015). For instance, she will not become defensive or angry when they are pointed out to her, while still experiencing an affective response to this confrontation because, as we already established, she is moved by it to fix her flaws. As such, intellectual humility is connected to emotional regulation, as well as inquisitiveness, objectivity, and other virtues which are employed to discover or fix one's intellectual shortcomings. Moreover, (intellectual) honesty can be needed when the agent has to own up to their intellectual shortcomings, on top of understanding them.

[6] At least in this paper. Theorists who define extension only in accordance with some of these requirements are free to apply only their requirements to extending intellectual virtue.

In contemporary psychology, intellectual humility has been connected to wisdom, a topic of growing interest in the last three decades (e.g., Grossmann 2012; Staudinger and Gluck 2011; Sternberg 1985): to improve intellectual humility, is to improve wise thought. So, one line of empirical research entails finding the conditions under which wisdom, including intellectual humility, blossoms.[7] Grossmann (2012; 2014; 2020) identifies a significant positive relation between construing one's situation from a distanced, rather than immersed, perspective and increased intellectual humility.

Participants increased their intellectual humility through distancing themselves from their personal situation by using third person language and by describing their situation from an outside perspective. In the earlier studies, participants verbally described their situation to the examiner from this distanced perspective. Their reasoning was more intellectually humble, with participants admitting that they did not have sufficient information, could not predict the future, and accurately suspending judgement more frequently compared to the control group. The most recent study (Grossmann et al. 2021) asked participants to write a daily reflection on a challenging situation they experienced that day from the third person perspective, or from the first-person perspective in the control group. In the test group, participants' intellectual humility increased. This effect was still present a week later, when participants came in for their last appointment.

Reconceptualising one's experiences from a distanced perspective appears to be a fruitful way to increase intellectually humble behaviour. However, experiencing or reflecting on traumatic events, as well as during heated arguments – in other words, during episodes of emotional turmoil – immersion is automatic, countering the distancing effects of prior third-person writing (Grossmann & Kross, 2010; McIsaac & Eich, 2004; Wegner & Giuliano, 1980). Intellectual humility is, like other

[7] While the referenced studies also research other possible aspects of wisdom, such as the willingness to compromise, this paper merely focusses on the means that enhance intellectual humility.

intellectual virtues, a relatively stable trait, even when the agent is confronted with a challenging situation. While it is relatively easy to be intellectually humble in response to a news item far removed from one's identity, it takes virtue to remain that way when confronted with shocking information about a family member or with regards to a lifelong project. So, to keep the increase in intellectually humble reasoning stable, supporting an enhancement of the virtue itself rather than shaping behaviour occasionally, I suggest a means to strengthen one's intellectual humility regularly and reliably through the empirically supported means of writing about one's situation from a third person perspective.

The app

Apps are a pervasive technology in the digital world, with some directed at aiding education or other learning endeavours. For example, Duolingo is a language-learning app which employs methods of gamification to engage and motivate users to commit to long-term daily practice. These gamification methods include presenting the user with rewards ("lingots") when a lesson is completed, achievement tokens, and a visible streak representing the number of days Duolingo was used in a row. Moreover, the app sends notifications to the user, nudging them to open the app to start their daily practise. While Duolingo might not be consistently effective, especially when an agent wishes to truly learn a language rather than learning a select number of words and sentences relevant for travel, its use, similar to apps such as MyFitnessPal and Forest, still motivates agents to acquire desirable habits and skills. This paper does not concern itself with the question whether these specific apps can be employed in examples of extension. However, the motivating qualities are useful to keep in mind for extending intellectual humility.

I propose the possibility of an app which induces the positive effect of thirdperson narration on intellectual humility. Every day, on a time set by the agent, the app sends a notification. This notification prompts the agent to answer questions, formerly posited by the researchers, now posited by the app. For example, "what is the biggest challenge [name] face(d) today?" and "Picture yourself in the event and ask yourself,

"Why is he/she [referring to yourself] feeling or behaving this way?" (Grossmann et al. 2021, Table 1). The agent answers the questions in the third person – either through writing or through a speech-to-text method. Once every question is answered, the app shows the user her answers, which she is nudged to read as she must tap the "submit" button for every individual question. After she has done so, she has finished her task for the day, prompting a text such as "Well done!" to appear on the screen.[8]

The app can be used in a similar way throughout the day, most importantly before making big decisions or during emotionally overwhelming situations. The temporary distancing from the situation via writing would allow for a moment of ego-decentralised reflection before any course of action is decided upon. While this function can sound intrusive and thereby not likely to be used, only a short amount of time is required for the ego-decentralisation to be improved.[9] So, for example, when an agent is feeling particularly anxious, she can use the app to write about the situation causing the anxiety, or even analyse her situation to see what this cause might be. For instance, if this agent experiences an upcoming event, such as a public speech, as anxiety inducing,

[8] Of course, with suggested applications, whatever they entail, questions of datamining and possibilities of hacking arise. For this reason, the agent can either use the app as a diary, saving her entries with the ability to read them later, or pressing the "submit" button simply deletes the text in question.

[9] In the study by Grossman (2021), subjects were not obligated to write more than a sentence.

she can write about this event from a third-person perspective, asking why she is anxious about this and similar events and whether the things she is afraid of are plausible to occur.

Note that the app extends intellectual humility rather than memory – the app does not merely remind the agent of the limits of her knowledge. After all, when the agent writes about her situation from an immersed perspective, intellectual humility might be increased, but only slightly when compared to taking the distanced perspective (Grossman 2021). As such, the app is not merely a memory crutch. Rather, it shapes one's experience of, and thereby one's attitude towards, the situation, i.e., one's interpretation of the situation, including her position in it, becomes more intellectually humble.

These characteristics of the workings of the app show the intimate integration of the agent with the app, where they are intimately, non-linearly coupled. While the agent's initial interaction with the app may be mostly causal, once the agent experiences the lasting effects of their daily exercise, on top of automatically reaching for their phone when they have to make a difficult decision or experience the possibility of an argument when interacting with others, then the app and agent influence one another consistently. The agent's intellectual humility, then, is constituted by coupled processes spanning both the phone and the agent.

Moreover, the trust and glue conditions apply when the agent's phone is available to the agent at any point during the day. This seems to be a likely condition, seeing the presence of smartphones in our daily life. Moreover, the app is used reliably and systematically, due to the daily notifications and the agent's consistent integration with the app.

The app also enhances the virtuous process, as it aids directing attention to the agent's uncertainty with regards to a situation or piece of information, which the agent may have not noticed otherwise. It ought to also regulate the agent's cognitive response, such as withholding judgement or acquiring

more information (after all, an agent would only use this app if she is motivated to acquire intellectual humility or become more aware of what she knows rather than implicitly believes.[10] So, the app's complementarity with the virtuous process leads to enhancing the intellectually humble trait. This fulfils the complementarity condition *and* the parity condition. Recall that parity refers to the object fulfilling the same function as it would in the head, yet Otto was precisely unable to memorise using just his brain. So, parity rather refers to whether the process has the same function if it were "located" in the brain. Similarly, intellectual humility extended to the app provides the app/agent interaction with the same function as intellectual humility would for a truly intellectually humble agent.

Lastly, the gamification quality of the app can make the virtuous process more pleasurable. While the positive reinforcement through the app's ability to have the agent experience a sense of accomplishment is mostly useful for increasing the agent's engagement with the app, this pleasurable experience is also tied to intellectually virtuous behaviour (Aristotle Eud. Eth, Nic. Eth.).

Due to these characteristics, with extended use, the app becomes part of one's intellectual humility, rather than merely improving the agent's attempts at intellectually humble behaviour. The daily use will also keep the agent aware of the possibility to use it when faced with a challenging situation, such as during verbal altercations (either online or in person). This allows the agent to be more reliably intellectually humble across situations.

Lastly, note that this app can also extend intellectually virtuous action, rather than only the virtue itself. After all, one important aspect of intellectual humility is recognising one's intellectual

[10] Another aspect of intellectual humility is to admit one's cognitive limitations to not just oneself, but others as well. While the subjects in the mentioned research projects did describe their relevant cognitive limitations to the researchers, admitting limitations to strangers is plausibly easier than admitting them to friends or coworkers. As such, the app might not enhance the agent's intellectual honesty within one's social circle.

limitations, which can include retroactive reflection on moments in one's day where this virtue may have been relevant. Such an app, then, can extend and aid intellectual humility in multiple ways, allowing the agent to increase this trait in line with her epistemic goals.

Conclusion

Responsibilist virtues are stable traits and attitudes which are motivated by positively valuing epistemic goods. As these virtues prescribe complex behaviours and are acquired over time, regular practise aids this acquisition process. However, this need not be the only way to improve intellectual virtue. In this paper I have suggested an app which can extend an agent's intellectual humility, promoting sensitivity to situations where intellectual humility is relevant and aiding the appropriate cognitive response (e.g., reflecting, gathering information, or suspending judgement). The app helps agents to attend to their limitations, it helps them to understand their limitations, and it helps them own their limitations, for instance by attributing mistakes they make to their limitations. Moreover, the distanced perspective taking also aids a more fitting emotional response when they are confronted with their limitations, for instance by noticing a mistake they made or a friend mentioning their limitation. Through the daily notification, the app motivates the agent to increase their intellectual humility. It does not, however, motivate the agent to act on other virtues, such as curiosity, courage, and integrity. For these virtues, perhaps other methods of extension or scaffolding can be devised, but that is a topic for future research.