7-25-2023

The Vices of Virtues: Making Room for Moral Testimony in the Life of the Virtuous Person

Maria Waggoner
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Philosophy-Psychology-Neuroscience

Dissertation Examination Committee:
Allan Hazlett, Chair
Anne Baril
John M. Doris
Ethan Kross
Ron Mallon

The Vices of Virtues: Making Room for Moral Testimony in the Life of the Virtuous Person
by
Maria Waggoner

A dissertation presented to the Office of Graduate Studies
at Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

June 2023

St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgments

encouragement and support, I would not be the philosopher that I am today. I have learned, first-hand, about being a good mentor and advisor from being on the receiving end; Allan has set the bar high when it comes to what a good advisor looks like – a bar that I can only one day hope to approximate.

I am thankful to have parents who have been supportive of my studies in philosophy since the beginning and have never questioned the value of doing philosophy or expressed concern about my job prospects (although I'm sure they had some at times!), but only showed encouragement and pride of my work.

Most importantly, it is not without my husband, Adam, that I could have crawled out various philosophical and emotional holes along this journey. His patient listening, confidence in my ideas, and persistent investment in me as a philosopher have been exactly what I've needed. His creativity and insight have surely shaped my project – all of the plausible ideas in what follows are Adam's; all of the implausible ones are mine! Adam is the best sounding board, the most supportive colleague, and an immensely loving spouse and friend. He, along with our wonderful son, Athanasius, provides me with a constant light of joy and always a larger (and better) perspective.

And lastly, praise be to God, from whom all blessings flow.

<div align="right">Maria Waggoner</div>

Washington University in St. Louis

June 2023

ABSTRACT OF THE DISSERTATION

The Vices of Virtues: Making Room for Moral Testimony in the Life of the Virtuous

Person

by

Maria Waggoner

Doctor of Philosophy in Philosophy-Neuroscience-Psychology

Washington University in St. Louis, 2023

Professor Allan Hazlett, Chair

This dissertation sheds light on new and unacknowledged difficulties that we face in striving to be (more) virtuous. By making use of empirical literature from moral, affective, and perceptual learning, I explore the potential cognitive and psychological relationships between having a virtue in one context and the tendency to exhibit vices in another. I do this by showing how morally good behavioral habits can also lead to morally inappropriate actions, when a virtuous moral perceptual system can give rise to moral illusions, and when our basic evaluative affective responses differ in their degree of sensitivity, leading to having some virtues while lacking others. I then connect this to recent debates on the nature of moral understanding and its relationship to virtue. I argue that virtue, rather than merely a lack of it, can lead to a deficiency in one's moral understanding. I use this to make room for when a virtuous person can and ought to defer to another's moral testimony.

Chapter 1: The Focus of Virtue: Attention Broadening in Empirically Informed Accounts of Virtue Cultivation

1.1: Introduction

Recently, empirical standards of virtue-based theories have been raised,[1] with many now focusing on giving empirically informed accounts of how we might go about developing virtue. Some have suggested we change the situations we find ourselves in, so that the situational cues we encounter trigger different cognitive and behavioral effects.[2] Others have argued we directly modify our cognition by adopting particular intentions, resulting in better behaviors.[3] And still others have suggested we indirectly modify our cognition, through routine practices like meditation.[4]

In this paper, I first look at empirically informed proposals of virtue cultivation that rely on direct modifications of cognition – in particular, those that use varying techniques of goal pursuit – in order to change one's behavior. I argue that these techniques of goal pursuit effectively change behavior due to the attention narrowing they bring about, and further show that such attention narrowing can threaten the appropriate exercise of phronetic-related capacities. When these phronetic-related capacities are threatened, two derivative problems arise: (1) One can end up acting in

---

[1] Largely thanks to philosophical situationists (Doris, 1998, 2002; Harman, 1999, 2000, 2003) who cast doubt on the empirical adequacy of virtue ethics.

[2] See, for instance, (Doris, 2002; Kamtekar, 2004).

[3] See, for instance, (Snow, 2006, 2010, 2016; Kamtekar, 2004; Besser-Jones, 2008; Railton, 2011; Webber, 2016).

[4] See Upton (2017); Other sorts of indirect modifications might be daily exercise or getting sufficient sleep.

morally inappropriate ways, and (2) Even in cases where one performs the morally appropriate action, one nonetheless can fail to notice and appreciate features of moral value. I argue that in light of these concerns, such techniques of goal pursuit – at least, by themselves – serve the most good for those who are merely trying to avoid vice, rather than cultivate virtue. And so, I suggest that such accounts of virtue cultivation are incomplete. I then go on to suggest that these undesirable effects of attention narrowing brought about by the use of these techniques during goal pursuit may be ameliorated by also engaging in certain indirect modifications of cognition, particularly those which broaden attention. My suggestion, then, is that these direct and indirect modifications of cognition might best facilitate virtue cultivation when employed together. While the particulars of how we might best go about this are currently unclear, I end this paper by exploring one option that might be further investigated.

1.2: An overview of goal pursuit in proposals of virtue cultivation

Various empirically grounded proposals of virtue cultivation have invoked the use of goal pursuit. While different proposals emphasize different aspects, I take 'goal pursuit' to involve the following three features: (1) Adopting a goal (2) Selecting a course of action to achieve that goal (3) Executing the course of action A common empirical model of goal pursuit, The Rubicon Model of Action (Heckhausen & Gollwitzer, 1987) lays out four stages of goal pursuit, where the first three map on to the features I

mentioned above: One first begins with deliberation about what goal to adopt. In adopting a goal, one commits to it by forming an intention, which moves one from having a mere wish and into the volitional stages, which consist first of planning and then of acting. The planning phase involves selecting certain subordinate goals or plans that detail a specific course of action which is to be taken for the sake of a larger goal.[5] During this planning phase, "[one] should address questions of when and where to start acting, how to act, and how long to act" (Gollwitzer, 1990, p. 57). After formulating a plan of action, one then moves on to initiating that action. The Rubicon Model also posits a fourth step: After carrying out the action, one evaluates what one has done. When the goal is not reached, one might adjust the goal or return for another try. These phases of the Rubicon Model can be seen in Figure 1.1[6]

---

[5] Detailing a specific course of action in the planning phase will involve making subordinate, or more specific and short-term goals. See *Section IV*, where I suggest that planning a course of action is a way of setting a subordinate goal. The Rubicon Model of Action – and in particular the movement from adopting a superordinate goal to forming plans and setting further subordinate goals – is a plausible model of how we naturally go about pursuing goals. Gollwitzer and Brandstätter (1997), for instance, found that two-thirds of their subjects said that had naturally formulated set plans of action for their goals, without being instructed to do so. Likewise, Carver & Scheier explain that superordinate goals often given way to subordinate goals just below them (Carver & Scheier, 2003, p. 189).

[6] Figure 1 was taken, with permission, from Achtziger & Gollwitzer's *Figure 12.1* (Achtziger & Gollwitzer, 2018, p. 487)
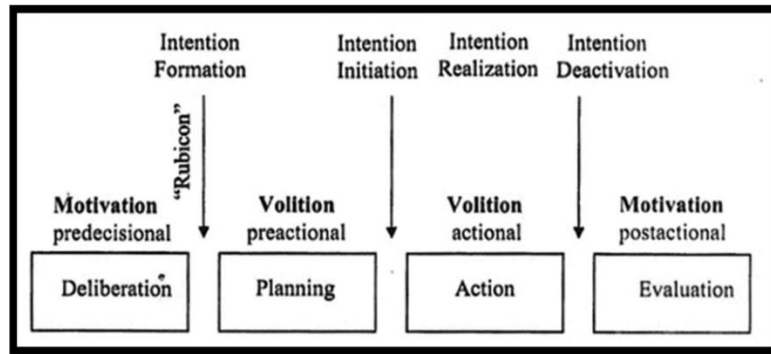
Figure 1.1. *Taken from Gollwitzer & Achtziger (2018), Figure 1 illustrates the four phases of the Rubicon Model of Action*

invokes the Rubicon Model of goal pursuit. Stichter suggests that goal pursuit plays an important role in self-regulation, which is necessary for the development of amoral or moral (e.g. virtues) skills.[7] Stichter explains that setting a goal, and committing to it, "is part of this process of forethought, [which] . . . motivates the next phase of forethought in planning what steps to take to achieve that goal" (p. 18). Once having completed the planning phase, you "can implement your plan and take action to achieve your goal" (p. 20). By committing oneself to a goal, and developing a plan to achieve that goal, one begins down the path of changing one's behavior. Such behavioral changes will require self-regulation in order to follow through with one's plan and to avoid alternatives that could prevent goal satisfaction. For Stichter, goal pursuit is an essential part of cultivating virtue, since Stichter holds that virtues are skills, and developing skills requires self-regulation. Goal pursuit is an important kind of self-regulation, as it helps us stay focused on carrying out particular actions without getting distracted or giving up.

---

[7] I am indebted to an anonymous referee at *Philosophical Psychology* for bringing Stichter's work to my attention.

The Rubicon Model of Action doesn't only explain how goal pursuit is carried out in a conscious and explicit manner, but also can account for a more automatic and preconscious method. Achtziger and Gollwitzer (2018) explain that in cases where the goals are pursued in an automatic way, the planned action is consistently performed in particular goal-relevant situations and the behavior eventually becomes paired with these situations, forming habits. In such cases of habitual goal pursuit, goals are activated and initiate behavior, skipping past the deliberative and planning phases because "all that remains to be done is to wait for the critical situation to arise . . . as soon as the critical situation is encountered, the respective goal-directed behavior is initiated" (Achtziger & Gollwitzer, 2018, p. 492). Implementation intentions, a specific technique used in goal pursuit, is a way of transitioning from explicit and conscious goal pursuit to a mode that is more habitual and automatized. When one is in the planning phase and explicitly forms plans of action to carry out, this typically consists of specifying the 'when', 'where', and 'how' (Gollwitzer, 1990, p. 57).

Implementation intentions specify if-then plans of action whereby one links specific situational cues ("if I am offered a beer . . . ") to a specific behavioral response (e.g. " . . . then I will ask for a sparkling water instead."). By formulating the 'if' portion of the plan, certain situational cues become more accessible, making it easier to pick up on such cues. By formulating the 'then' portion of the plan, it is thought that we will be better able to execute the behavioral response. When these specific plans are continually

rehearsed and carried out, they become habitual and automatic. Indeed, implementation intentions have proven effective in this realm, and have helped with behavioral modifications (Gollwitzer & Brandstätter, 1997; Gollwitzer & Sheeran, 2009; Webb & Sheeran, 2007). Several goal-pursuit accounts of virtue cultivation have made use of implementation intentions, including Besser-Jones (2008), Kamtekar (2004), Railton (2011), and Stichter (2018).[8] Stichter, for instance, suggests implementation intentions as a way to offload cognitively taxing self-regulation in the process of skill acquisition, including the moral skill of virtue.[9] Others, such as Railton (2011), focus on the ways that implementation intentions can be used to 'save' virtue from the situationist critique. While it might be true that particular situational factors trigger very specific behaviors, we not only change what situational factors become salient to us through the use of implementation intentions, but also the way we respond, due to the particular situation-behavior pairs or associations established.

Aside from the use of implementation intentions, others – such as Nancy Snow (2006, 2010, 2016) – have made use of the automatic activation of goals as a technique for developing virtue-related habits and carrying out virtue-related actions: Snow explains that our lives often involve pursing certain goals – such as being a good parent – which require the development of particular virtues, like patience, humility, and love. When the

---

[8] In particular, see (Kamtekar, 2004, p. 487–8; Besser-Jones, 2008, p. 328–9; Railton, 2011, p. 323).
[9] See Chapter 1 of Stichter (2018)

goal is held over a long period of time, is of high importance, and is continually pursued, the goal will eventually be activated in an automatic and preconscious way. In these cases, encountering certain goal-relevant stimuli can produce goal-relevant behaviors. Goal-relevant stimuli and goal-relevant behaviors will be more salient and more accessible when compared to stimuli and behaviors that are not associated with the particular goal one has set. In her case of being a good parent, Snow explains that such a person, "places great importance on the goal of caring for her child . . . When she encounters situational features that activate or trigger the representation of a goal, other things being equal, she will respond by acting in ways that promote goal attainment" (Snow, 2016, pp. 139–40). Snow supports her case by citing empirical research from those like Aaars and Dijkersterhuis (2000), Bargh (1990), and Bargh and Gollitzer (1994), which suggests that goals can be activated non-consciously when one encounters goal-relevant cues or is in goal-related circumstances. This body of empirical literature indicates that setting goals and forming goal-related habits can often result in activation of the goal unconsciously. And when this happens, "the activation of the goal to act automatically elicits habitual behavior" (Snow, 2010, p. 44).[10]   When goals that we set are activated, this influences what stimuli are made salient to us, and thereby how we respond to the situation at hand. Thus, Snow suggests that by setting certain goals (and not others), our actions might become more closely aligned to virtue.

---

[10] Here, Snow is referencing findings from Aaars and Dijkersterhuis (2000)

It is worth noting that Snow makes use of empirical literature of goal and behavioral priming, which has recently come under scrutiny. For instance, Snow invokes John Bargh's work, and several of the findings reported in Bargh's studies – such as Bargh et al. (1996), Bargh et al. (2001), and Williams and Bargh (2008), – have since failed to replicate, as shown by Doyen et al. (2012), Harris et al. (2013), and Shanks et al. (2013). If these recent replication failures indicate that there are no effects of goal priming on behavior, Snow's account might fail to give us empirically supported methods of cultivating virtue through behavioral changes.

While this body of literature is complicated and the effects are far from straightforward, a complete dismissal of goal priming effects would be too quick, as these replication failures do not necessarily show that there are no such goal priming effects. Dijksterhuis et al. (2014), for instance, argue, that "we should not . . . believe that the initial findings [of goal priming effects] were false positives . . . [as the priming effects] have been obtained in many different published experiments" (p. 208). A recent meta-analysis carried out by Weingarten et al. (2016) looked at 352 effect sizes, from 133 published and unpublished studies, revealed small but robust, significant goal priming effects on behavior (p. 490).[11]

---

[11] For further research that has found such priming effects, see: (Bry et al., 2008; Dijkersterhuis et al., 1998; Dijksterhuis & Van Knippenberg, 1998, 2000; Galinsky et al., 2008; Haddock et al., 2002; Hansen & Wanke, 2009; LeBoeuf & Estes, 2004; Lowery et al., 2007; Nussinson et al., 2010; Schubert & Hafner, 2003; Chen & Latham, 2014; Ganegoda et al., 2016; Papies, 2016; Payne et al., 2016; Van der Laan et al., 2017; Latham, 2018; Stajkovic et al., 2019).

Others – such as Klatzsky and Creswell (2014), Loersch and Payne (2011, 2014), Payne et al. (2016), and Schröder and Thagard (2013), – have accounted for the mixed results by arguing that these priming effects are more nuanced than initially thought. Previously overlooked moderators can help explain under what circumstances priming effects do and do not take place. Given the strong evidence for efficacious use of automatized implementation intentions, one potential moderator may involve the specificity of the goals adopted. While Snow seems to suggest that broader, superordinate, goals – such as being a good parent – can be automatically triggered by the situational cue, it is unclear if this is in fact what Bargh suggests. In his (1990) piece that Snow draws from, Bargh speculates it is the more specific goals – such as the particular plans of action – that are automatically activated when encountering particular situational cues:

> In general . . . it would seem that the more abstract and less concrete the goal – that is, the broader the array of behaviors that will satisfy it – the less likely it will be for that goal to become capable of direct activation by the environment. This is because the 'longer' a cognitive pathway is (i.e., the more links it contains), the less likely it is to become automated; the more abstract a representation, the greater the number of analytic steps both between it and the relevant environmental feature detectors on the one hand, and the action effector units on the other. (p. 117)

If the specificity of the goal adopted does in fact influence the effectiveness of goal-automaticity, then we might think that Snow's suggestion is correct, but only in a restricted sense – perhaps virtue could really only be cultivated with the use of more narrow goals, rather than merely vague, or superordinate, goals which she speaks of.

Even if it were true that Snow relied on shaky empirical findings, we can set these specifics aside, for there is a general point that still holds: Adopting and pursuing goals – whether done in an explicit and conscious or automatic and unconscious manner – will likely impact our cognition and influence our behavior. If it were to turn out that a particular technique of goal pursuit, such as goal priming, doesn't actually change behavior at all, then this method would be irrelevant for virtue cultivation, since the purpose of using goal pursuit as a way to cultivate virtue is so to *change our behavior* to better align with virtuous actions. And so, if a particular technique of goal pursuit in fact fails to bring about behavioral effects, then this technique becomes irrelevant for my argument, since such techniques are no longer goal pursuit accounts of *virtue cultivation*.

1.3: The role of phronesis-related skills in virtuous goal-directed behavior

In the previous section, I discussed recent accounts of virtue cultivation which invoke the various techniques of goal pursuit. While goal pursuit may involve adopting a larger, superordinate goal, it also usually involves formulating more specific plans or subordinate goals that are a means for achieving the larger goal.[12] But pursuing virtue-relevant goals will not necessarily amount to performing virtuous acts, for one will also need certain phronetic-related skills to discern which goal ought to be pursued at a given time, and what is the best course of action for doing so.

---

[12] Cf. note 5

Virtuous action requires at least two skills related to phronesis, which I call *perception* and *evaluation*. In exercising the skill of perception, one correctly identifies which goal(s) are relevant to pursue within a given context. The need for such a perceptive skill applies to both superordinate goals and the more specific plans for achieving that goal, for one must correctly identify which superordinate goals are relevant within a given context, as well as what course of action is currently a realistic means for achieving that superordinate goal. This skill of perception is largely attentional in nature: By directing one's attention to relevant goal-related stimuli, one correctly sees what goals are relevant within a given context. What I call perception is similar to what Dan Russell (2009) describes as the practical capacity of *comprehension*, which involves "the ability to 'read' a situation . . . so as to recognize what is salient" (p. 21). Darnell et al. (2019), likewise, detail a 'constitutive function' of phronesis, which "enables an agent to perceive what the salient features of a given situation are from an ethical perspective" (p. 118). Russell (2009) notes that comprehension is "not prescriptive but only discriminatory (p. 21), for 'reading' the situation tells us which routes of action are relevant, but it does not necessarily tell us which ones we ultimately ought to pursue. Thus, this skill of perception involves simply being aware of morally relevant features and routes of action within a given situation.

Given that perception only descriptively highlights important ethical features of a situation, another phronetic-related skill is required to identify what feature or option is

most important – I call this skill, *evaluation*. Evaluation involves judging which goal ought

to be pursued within a given circumstance. Given this, correct evaluation will depend on

correct perception, but it will also go beyond insofar as evaluation involves judging

evaluative importance. Additionally, proper evaluation will need to occur both at the

superordinate and subordinate level. Darnell et al. (2019) describe a sort of evaluative

weighing at the superordinate level when they explicate the 'integrative function' of

phronesis, which involves "integrating different components of a good life, especially in

dilemmatic situations where different ethically salient considerations or virtues appear

to be in conflict" (p. 118). On the other hand, Russell (2009) notes a component of

phronesis involving evaluative judgment-making at the level of planning, whereby one

"correctly adjusts one's grasp of what one must do in particular circumstances as regards

a general end, such as acting generously or as a good friend" (p. 22). Thus, whether this

judging occurs at the superordinate level of goals or the more particular level of forming

plans to achieve that goal, I call this weighing of relative importance the skill of

evaluation.

Failures in either perception or evaluation can lead to problems for cultivating

virtue. Thus, any goal pursuit account of virtue cultivation will need to make room for

the importance of phronetic-related skills of perception and evaluation.[13] Without such

---

[13] I take the skills of perception and evaluation to be necessary elements to goal-based accounts of virtue
cultivation, but not sufficient. There are likely other elements of phronesis that will also need to be had in

skills of discernment and judgment, one will likely fail to appropriately pursue the appropriate goal(s) in a given circumstance.

1.4: Goal pursuit and the role of narrowed attention

In *Section 1.2* I considered a few accounts of virtue cultivation that rely on techniques of goal pursuit, along with noting the different phases of goal pursuit, such as goal adoption, planning, and execution of said action. In this section, I return to the empirical literature to highlight the importance of attention narrowing in the planning and action phases of goal pursuit.

Recall The Rubicon Model of Action, which posits four stages of goal pursuit. After the initial phase of deliberation, one forms an intention or commits to a goal. This commitment results in focusing in on forming a plan of action and then executing that plan. Whether this is done consciously and deliberatively, or unconsciously and automatically, the planning and action execution involve a 'closed-minded' cognitive state, in which attention is focused on the course of action. This is often described as 'goal shielding', in which other concerns or goals are blocked out, so that one can narrow in on one's particular plan of action.[14] Furthermore, the extent to which one can narrow in on

---

order to act in a virtuous manner. However, discussion of all relevant and necessary elements is beyond the scope of this paper.

[14] For more on the phenomena of goal shielding, see (Bélanger et al., 2013; Kruglanski et al., 2002; Shah et al., 2002; Veling & Van Knippenberg, 2006).

one's plan and action is related to the likelihood of successfully carrying out that action, for, as Gollwitzer explains, "it matters [for goal satisfaction] whether one can shield an ongoing goal pursuit from distractions . . . [and] competing temptations" (Gollwitzer, 1990, p. 494). Failing to reach the desired end-state specified by the goal is often a result of insufficient attention narrowing in on one's prescribed plans and actions, or a lack of goal shielding.

One way in which goal shielding occurs is by changes in visual attention. Van der Laan et al. (2017), for instance, found that when primed with diet-related goals, participants visually attended more toward goal-relevant, healthy foods, as measured by eye-tracking movements. This, in turn, resulted in an increase in healthy food choices. Likewise, when implementation intentions are activated by the relevant situational cue, attention is similarly captured and focused: Janczyk et al. (2015) explain that a "heighted activation of the stimulus specified in the if-part of an implementation intention appears to enhance early attentional processes, such as attentional filtering" (p. 208), which accounts for why certain stimuli are narrowed in on while others ignored. Wieber and Sassenberg (2006) found that when a participant was visually exposed to the cue associated with the if portion of an implementation intention, visual attention was drawn toward the cue and away from other stimuli. Webb and Sheeran (2007) likewise found that subjects who adopted an implementation intention *if X, then I'll do Y*, were more likely to detect *X* than subjects who didn't adopt this implementation intention.

Achtziger, Bayer & Gollwitzer (2012) found similar results with auditory attention as well.

While attention initially narrows in the Planning Phase, a 'close-minded' state and

narrowed attention occurs during the Action Phase as well. When carrying out the goal-

directed action, Achtziger and Gollwitzer (2018) note the following attentional effects:

> Individuals . . . do not consider alternative strategies, neither do they form
> implementation intentions or action plans . . . they ignore any potentially
> disruptive aspects . . . The action mindset focuses attention on those aspects of the
> self and the environment that sustain the course of action. (p. 492)

Indeed, (at least part of) the success of implementation intentions and goal-priming in

changing behavior is due to such sustained, narrowed attention during the Action Phase.

It is because one's attention stays focused on the action at hand that one is more likely to

complete it, avoiding distractions or temptations.[15] That narrowed attention plays this

important role in effective goal pursuit is evident when we consider the fact that the

desired outcome specified in a goal is more likely to occur when using techniques that

---

[15] To clarify, the narrowed attention present in the Action Phase is plausibly the same sort that occurs with
the use of implementation intentions. As mentioned above in *Section II*, when one's goals are habitually
initiated, "all that remains to be done is to wait for the critical situation to arise. . .as soon as the critical
situation is encountered, the respective goal directed behavior is initiated." (Achtziger & Gollwitzer, 2018,
p. 492). With habitual goals, one skips the Planning Phase and, when the relevant situation arises, the habit
is activated, and the Action Phase proceeds. Achtziger and Gollwitzer (2018) note that this not only applies
to habits in general, but goals furnished with implementation intentions, as well: "If, upon crossing the
Rubicon, the goal was furnished with implementation intentions. . .specifying when, where, and how
actions are to be initiated, all that remains to be done is to wait for an appropriate opportunity, to arise (i.e.
the "when" and "where" specified in the implementation intention). . .If (the opportunity that arises is) a
match. . .goal-directed behavior is initiated immediately" (p. 492). Thus, whether one undergoes explicit
planning or already has established habits or implementation intentions, the Action Phase proceeds. Since
the "actional mindset is. . .hypothesized to be one of closed-mindedness", and "action mindset should
emerge whenever people move effectively toward goal attainment," (ibid) this close-mindedness or
narrowed attention is present when one is in Action Phase and effectively moving toward her goal,
regardless of whether the goal is or is not furnished with implementation intentions.

employ specified, as opposed to vague, goals.[16] Plausibly, this is at least one reason why the use of implementation intentions has proven rather effective at achieving the desired behavioral outcomes.[17] Gollwitzer (1990) explains that "implementation intentions are subordinate to goal intentions" (p. 494) and so we might think of implementation intentions as involving very specific, subordinate 'goals.'[18] Penningroth & Scott (2008) explain that we can think of "subordinate-level goals . . . [as] specific activities that can be executed in order to meet a high-level goal" (p. 74) whereby they explain that these activities can take the form of implementation intentions, but also other means such as prospective memories (ibid).[19] The more specified the goal or prescribed activity is, the more alternatives will be shielded, and so techniques of goal pursuit that involve greater specification of one's goal will be more effective in bringing about the goal-specified outcome.

---

[16] See (Locke & Latham, 2019, 2013, 2002).

[17] That implementation intentions are very specific is not to say that this is the only reason for why implementation intentions have proven effective.

[18] In the following passage, Gollwitzer actually argues that implementation intentions are distinct from goal intentions: "[With implementation intentions,] it is not a person's self that is linked to a desired end state (as with goal intentions); rather the person commits himself or herself to respond to a certain situation in a specific manner." (Gollwitzer 1990: 494). However, it is unclear that implementation intentions and goal intentions really do differ in this way, as one could adopt a goal intention for a very specific, subordinate goal, such as to "brush my teeth every morning as soon as I get out of bed." Setting this specific goal of brushing your teeth immediately upon rising seems to link a response (brushing teeth) to a certain situation (getting out of bed), yet it isn't clear how this means it ceases to be a goal intention. In any case, even if there is a conceptual difference between goal intentions and implementation intentions, this difference would seem to be something other than the similarities in specificity and narrowed attention, since goal intentions could be formed in a specified way that is on par with a given implementation intention, and as a result, both would likely have the same sorts of attention narrowing effects.

[19] Prospective memories are desired future actions that one commits to, such as "I will go jogging at 7am tomorrow."

1.5: The downfalls of goal pursuit accounts

In the last section, I highlighted how narrowed attention plays an important role in effectively carrying various methods of goal pursuit. In this section, I argue that, while virtuous goal pursuit needs to be guided by the phronetic-related skills of perception and evaluation, the most effective techniques used in goal pursuit can threaten the appropriate exercise of these skills. I then go on to show that two further derivative problems arise when perception and evaluation are threatened: (1) one may carry out morally inappropriate actions, and (2) even if a person avoids (1), she may still fail to notice or appreciate important moral features.

1.5.A: Narrowed attention threatens appropriate exercise of phronetic-related capacities

As Beckman and Heckhausen (2018) observe, once "crossing the Rubicon, people tend to either forget about the alternatives they have rejected or play them down" (p. 115).[20] While such forgetting or downplaying can be a good thing when this results in disregarding information that ought to be ignored, sometimes we overlook information that merely *seems* irrelevant.

---

[20] 'Crossing the Rubicon' is a reference to Caesar's decision to cross the Rubicon. This phrase is used to indicate a movement from deliberation to commitment to a course of action.

Empirical research suggests that appropriate exercise of perception can be threatened by goal pursuit, especially when using a particular technique of goal pursuit that involves highly specific subordinate goals or plans, as this usually narrows or focuses attention. For instance, subjects using implementation intentions stuck to their plan of identifying *x* cue and responding with *y* behavior, even when a more efficacious path was available for achieving the same goal (Belyavsky-Bayuk et al., 2010; Masicampo & Baumeister, 2012; Parks-Stamm et al., 2007). Likewise, implementation intentions make one *worse* (when compared to those who weren't using implementation intentions) at exhibiting goal-relevant behavior when the specific situational cue is absent (Bieleke et al., 2017). In addition to proving counterproductive to achieving the specific goal at hand, using implementation intentions has also shown to interfere with noticing goal-relevant cues for *other* goals (Wieber & Sassenberg, 2006). In their review of goal-directed behavior, Ordóñez et al. (2009) note that "goals focus attention," but, "unfortunately, goals can focus attention so narrowly that people overlook other important features of a task" (pp. 7–8).

Recall that in the previous section, findings from Van der Laan et al. (2017) and Janczyk et al. (2015) indicated that goals influence action through their effects on directing visual attention. Likewise, Büttner et al. (2014) found that when adopting an implemental mind-set – as occurs during the Planning phase, participants focused their visual attention on objects in the foreground (Büttner et al., 2014, p. 1248). If the mere targets of

visual tracking are impacted by the goal being pursued, it is no surprise that one's goal(s) influence what stimuli one perceives and responds to in one's actions. This can become problematic when pursuing a given goal prevents one from perceiving other important stimuli relevant to other (perhaps, more important) ethical concerns.

In addition to negatively influencing perception, there is some evidence that at least some methods of goal pursuit could also impact appropriate exercise of evaluation. Belyavsky-Bayuk et al. (2010) found that even when subjects did notice an alternative option, if this alternative fell outside of their adopted implementation intention, subjects misjudged the alternative's value, incorrectly deeming it not valuable for attaining for their superordinate goals.

It is important to note that, if goal pursuit reliably results in goal satisfaction, it will be (at least in part) due to the attention narrowing effects of the particular goal pursuit-technique used. This means that, perhaps not all means of pursuing goals – such as the employment of more vague, or superordinate ones – will strictly narrow attention, and so will likely not face these negative consequences.[21] But, because such means of goal pursuit typically do not involve attention narrowing mechanisms, we should suspect they will be less effective in bringing about the desired behavioral changes. Given that I am concerned with techniques of goal pursuit that bring about behavioral changes for

---

[21] At least, superordinate goals will likely not strictly narrowing attention and avoid such negative consequences when these superordinate goals are not further furnished with more specific subordinate goals.

the sake of virtue cultivation, it is worth clarifying that I am concerned with the attentional effects that typically occur in techniques of *effective* goal pursuit.

Thus, whether it be through failures of perception or evaluation, techniques of effective goal pursuit narrow attention, which can interfere with the exercise of important phronetic-related skills, such that one overlooks and/ or undervalues alternatives that ought to be considered.

1.5.B: Two derivative problems

Given that narrowed attention underlies effective techniques of goal pursuit and that this can lead one to overlook and undervalue alternatives, morally inappropriate action can result. Even if we adopt virtue-relevant goals, this can nonetheless lead to less-than-virtuous action when circumstances are morally complex. Consider Darley and Batson (1973) Good Samaritan study, whereby subjects were instructed to walk across campus in order to attend a talk on The Good Samaritan. Subjects were divided into three conditions, whereby the degree to which they needed to hurry to arrive at the talk on time varied. We might say that (at least some of) the subjects employed the specific goal of *getting to the talk on time*, perhaps as part of their broader goals of *respecting the speaker*, or *learning more about being a Good Samaritan*. Subjects who were in the high-hurry condition plausibly adopted highly specific, particular plans of action that they would

walk fast, and *not stop to sight-see, talk to strangers*, etc.[22] While adopting and implementing

these specified plans likely helped the high-hurry subjects arrive on time to the talk, the

study also showed that these subjects were less likely to stop to help a confederate who

was disguised as a stranger in physical distress. In discussing their results, the

researchers noted that "it would be inaccurate to say that [these subjects] realized the

victim's possible distress, then chose to ignore it; instead, because of the time pressures,

they did not perceive the scene in the alley as an occasion for an ethical decision" (Darley

& Batson, 1973, p. 108). By giving the subjects a particular goal (*to get to the talk within a*

small time frame), whereby they then adopted even further specific plans or subordinate

goals (walking fast, not getting distracted by sightseeing or strangers) – their attention

became narrowed. Despite the fact that subjects adopted virtue-relevant goals, given that

the particular means to pursue their goal involved highly specific plans, and resulted in

narrowed attention, and so they fell short of what morality required of them to do.[23]

Yet, even in cases where engaging in such means of goal pursuit does not result in

morally inappropriate action, another problem may loom: Virtue may require us to

appreciate certain moral features, even if we should not, all things considered, act on

them. The attention narrowing that occurs with particularly effective means of goal

---

[22] We might think subjects did, in fact, adopt such goals, since those in the high-hurry condition were give these precise sorts of instructions.

[23] Conversely, broadened attention has been shown to increase pro-social behavior because when one's attention is broadened, she is more likely to perceive helping occasions in her immediate environment. (Mukerjee et al., 2018).

pursuit can infringe on perception and evaluation, making such appreciation harder to come by.

First, consider some similarities between goal pursuit accounts of virtue development and a McDowellian view of virtue, under which a virtuous agent is unified in her moral judgment and lacks inner conflict. According to McDowell, a virtuous person 'silences' – or does not attend to – reasons that run contrary to the virtuous act. McDowell claims that the virtuous person is simply concerned with "keeping [her] attention firmly fixed on what Aristotle calls 'the noble'" (1978, p. 27), such that other considerations will not enter into her practical reasoning. Like the techniques of goal pursuit we are concerned with, silencing accounts of virtue seem to operate by way of narrowed attention, screening out all non-goal related or non-virtue relevant features. In her critique of silencing accounts, Karen Stohr (2003) notes that if one were to not silence other features and instead recognized the complexity of the situation, the right action would be "hard for [one] to perform, despite the fact that it [would be] the right thing to do" (2003, p. 343). That silencing other features makes the right action easier to do is, likewise, the aim of using highly specified goals or plans in order to cultivate virtue: by narrowing one's attention, distraction and temptation is prevented, making it easier to follow through with the goal-relevant action. Thus, at least conceptually, both these techniques of goal pursuit and silencing accounts of virtue seem to involve a narrowing in on the action to be pursued, ignoring other features.

Given these similarities, objections raised against McDowell's silencing view might likewise apply for goal-based accounts. Some have argued that silencing accounts face the following problem: Virtue can require loss and sacrifice. Failing to appreciate those things lost or forgone is a mark against virtue, not in favor of it. For instance, Susan Stark (2001), argues that, according to Aristotle, "death and injury are painful, and even to the virtuous person, and in fact, *even more* so to the virtuous person, because she has more to lose than the rest of us" (p. 448). She goes on to explain what this means for the silencing account: "McDowell's virtuous person seems to act flippantly with regard to her own life and health. She proceeds into battle, seeing the danger, yes, but not internalizing the danger, not seeing it as danger to her, potential loss to herself. But virtue is not this flippancy" (ibid). Ann Margaret Baxley (2007) makes a similar observation, claiming that "virtue can have a cost, and a mark of the wise person is that she recognizes it" (p. 419). A virtuous person will see the various conflicting objects of value in the situation before her; if one paid no attention to them, we might plausibly think such a person to be cold and calloused, rather than virtuous.

While narrowed attention can interfere with the appropriate exercise of phronetic-related skills, which can lead to morally inappropriate action, we see that even in such cases where one ends up doing the right thing, another problem might still persist: Sometimes this narrowing in on the right course of action can amount to ignoring the complexities of the situation – including the costs that morality requires of us. Even if we

act rightly, silencing or overlooking such features might be a defect of character, rather than an aspect of it.[24]

1.6: Searching for a way forward

Thus far, I have argued that accounts of virtue cultivation which employ particularly effective techniques of goal pursuit tend to be successful due to their attention narrowing mechanisms. While boasting positive effects, such proposals are also subject to problems. However, I nonetheless see these methods of goal pursuit as playing an important role in guiding moral action, and so worth maintaining in some form. In this section, I will first suggest that such techniques used in goal pursuit, which operate via narrowed attention, might be particularly helpful not in cultivating virtue, but in ridding oneself of vice. Thus, these goal pursuit accounts are incomplete when it comes to the cultivation of virtue. What is needed is (at minimum) to balance narrowed with broadened attention. I propose that, when cultivating virtue, if one engages in direct modifications of cognition through techniques of goal pursuit – particularly ones that

---

[24] Even if the silencing account is correct and such objections are mistaken, it is unclear that goal-based accounts remained unscathed. Some (Schuster, 2020; Stark, 2001; Vigani, 2019) have pointed out that the silencing account might only apply to practical reasoning, and so only normative and/or motivating reasons are to be silenced. But this does not mean that all emotional responses ought to be likewise silenced. But, if this is true, then in order for one to emotionally respond to an object of value (even when it ought not, and does not, enter into one's practical reasoning), this object must be attended to. But various goal-based accounts seem to be in tension with such attention. If non-goal related features and values are overlooked and only goal-related aspects are focused on, then perhaps this account fares even worse than other silencing accounts.

involve a highly specified plan or subordinate goal, and so typically narrows attention –

one should also undertake indirect modifications of cognition that would result in some

amount of persistent, broadened attention. While the mechanism by which we are to

employ in this attention broadening is currently unclear, I offer an intriguing suggestion

of how it might be done by turning to empirical research on Open Monitoring Meditation,

highlighting its promise in expanding our attention.

1.6.A: When is mere goal pursuit most beneficial?

In *Section 1.5*, I made a brief comparison between techniques of goal pursuit and

silencing accounts of virtue, noting that they both seem to involve attention narrowing.

These similarities likewise bear out in some potential criticisms that both accounts face.

But now, I wish to draw again from the literature on silencing to point out the benefits

that these sorts of goal pursuit techniques likely have.

Critics of silencing accounts are happy to acknowledge that silencing seems like a

good thing to do when the features that are silenced are those which run contrary to

virtue. For instance, despite rejecting the silencing view, Jeffrey Seidman (2005) notes

that there are cases where the virtuous person should silence at least some options: "[the

virtuous person] will exclude courses of action which are incompatible with virtue from

the range of possibilities . . . If he is decent, these morally unsound possibilities may not

even occur to him" (p. 72). Likewise, Baxley (2007) notes that the silencing thesis gets

something right when it is applied to a small range of cases, namely those where "competing options are and should be silenced by the requirements of virtue. These are cases in which competing options in conflict with virtue are actually immoral or vicious" (p. 412). If a course of action is vicious, then it is likely advisable – and even a sign of virtue – to not let this option cross your mind.

A parallel sort of application can be made to certain goal pursuit accounts: At least some techniques used in goal pursuit are likely well-suited when guiding one's attention away from vicious options. Pursuing goals in this way results in overlooking actions that should never be chosen. And, in such cases, the worries I have brought forth in Section 1.5 seem to no longer apply. For, if these particular techniques used in goal pursuit involve directing our attention away from vicious options, then adopting this vice-ignoring strategy will likely never lead us astray. Like the concerns raised with the silencing thesis, the attention narrowing that is operative in these methods of goal pursuit only become problematic when the overlooked options are not vicious.

Interestingly, implementation intentions are often used when people are trying to kick a bad habit. For instance, a recovering drunkard might seek to avoid drunkenness by never ordering a beer, and the way he does this is by directing his attention away from beer to something else, such as sparkling water. It seems that, for the recovering drunkard at least, adopting a plan of action to always stick to sparkling water would rarely lead him astray. Given one's current vices, some courses of action – such as

26

drinking a beer – might rightfully be silenced. And, so, avoiding vicious courses of action might be where these techniques of goal pursuit can do a lot of the work.

But, avoiding vice is not necessarily the same thing as cultivating virtue. Indeed, the temperate man doesn't fail to notice the value in food or alcohol, but appreciates the value in the right way, to the right degree. But, before a recovering drunkard can acquire the virtue of temperance, he might have to (temporarily) silence some options – namely, those which lead him straight to vicious acts. Thus, various goal pursuit accounts face certain problems, but these problems specifically arise during the cultivation of virtue, rather than the mere avoidance of vicious behaviors. What I seek to do next is examine how certain techniques of goal pursuit might work not just in cases of vice avoidance, but also virtue cultivation. I suggest that we hold on to the efficacious techniques commonly used in goal pursuit, but compliment their use with an intervention that broadens attention, and so may help remedy the problems I have brought forth.

1.6.B: The need for broadened attention

In order to explicate my suggestion of how we might go about broadening attention, recall The Rubicon Model of Action that was discussed in Section 1.2, and the fact that attention becomes narrowed when one is developing a plan of action, and then again when one carries out that action. However, before and after these phases, attention is typically broadened to allow for deliberation of what one's options are and evaluation

of what one has done. Broadened attention allows for the taking in of information at these phases, in order to then select and narrow in on the best one. The Rubicon Model is depicted, again, in Figure 1.2, but this time color-coded according to the kind of attention that typically occurs at each stage.[25]

Thus far, I have argued that especially effective techniques of goal pursuit involve attention narrowing. However, when engaging in virtue cultivation (rather than mere vice avoidance), we might benefit from a more broadened mind-set. Incorporating broadened attention when cultivating virtue might involve moving from the Rubicon Model depicted in Figure 1.2, to a modified version depicted in Figure 1.3.[26]
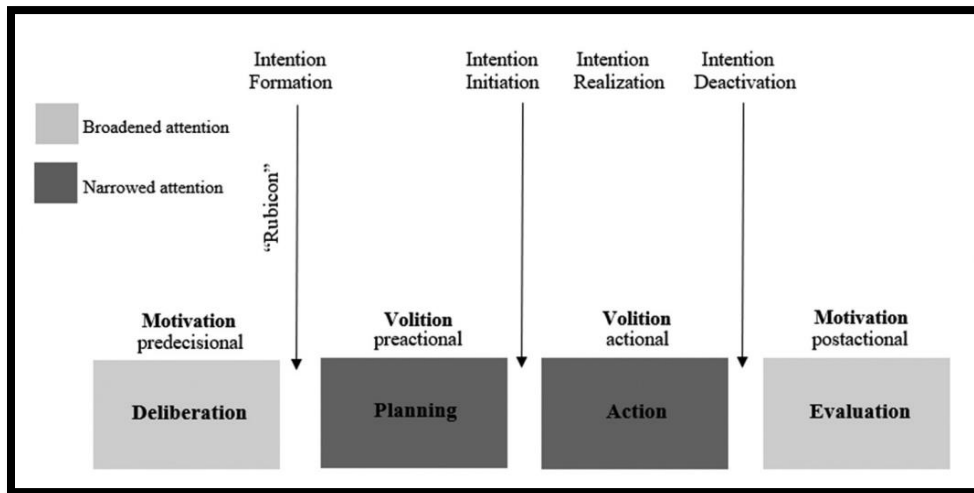
Figure 1.2. *Figure 1.2 depicts the Rubicon Model of Action, taken from Gollwitzer & Achtziger (2018), but additionally represents the attentional changes throughout the four phases. Indicated by light grey, broadened attention occurs in the Deliberation and Evaluation phases, while narrowed attention – indicated by dark grey – occurs in the Planning and Action phases.*

---

[25] Figure 1.2 was adapted from Figure 12.,1, Achtziger and Gollwitzer (2018, p. 487).
[26] Figure 1.3 was adapted from Figure 12.,1, Achtziger and Gollwitzer (2018, p. 487).
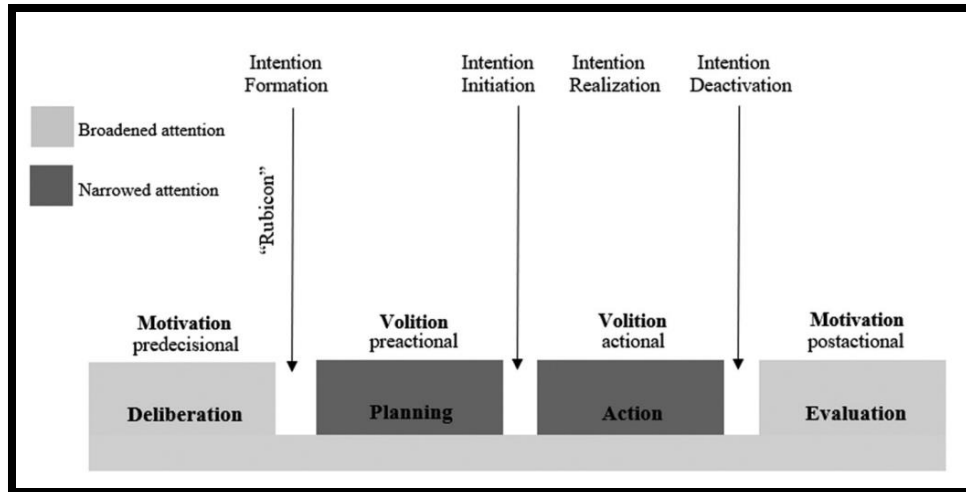
Figure 1.3. *Using the Rubicon Model of Action, Figure 1.3 shows what attentional patterns might look like during goal pursuit when one also has undergone practices to adopt a persistent, broadened mindset.*

This version suggests a certain degree of persisting, broadened attention which continues even throughout the more narrowed planning and action phases. Such increased broadened attention may be able to combat the problems that I have previously raised for goal pursuit. I now turn to empirical research on Open Monitoring Meditation, arguing that this may be one way by which one could acquire this broadened mind-set.

Open Monitoring Meditation (OMM) has received little attention in the literature on virtue cultivation, an exception being Candace Upton's (2017) paper in which she proposes (five different kinds of) mediation – including OMM – as an underexplored route for cultivating virtue.[27] Upton (2017) explains that OMM is a kind of meditation

---

[27] Upton groups all five different methods together and spells out their virtue-related benefits. Rather than advocating for meditation, in general, I rather propose we specifically look to the literature on OMM and the attention broadening effects that OMM can bring. In addition to focusing on OMM, rather than meditation more generally, my account also differs from Upton's, for I intend to incorporate the use of OMM with methods of goal pursuit.

where "the focus of the practitioner is to stay in the monitoring state, remaining attentive to any experience that might arise, without selecting, judging, or focusing on any particular object" (pp. 336–7). OMM has also been defined as when "the individual is open to perceive and observe any sensation or thought without focusing on a concept in the mind or a fixed item; therefore attention is flexible and unrestricted" (Colzato et al., 2012, p. 1). But, OMM is not simply letting one's mind wander – it uses executive or top-down capacities and, ironically, still involves directing one's attention: Garland et al. (2015) summarize the process as "involving a shift in attention", but where one's attention isn't intensely focused on a specific object in one's immediate awareness, but rather "monitor[s] the contents of consciousness while becoming aware of the quality of awareness itself . . . [taking on] a metacognitive state of awareness" (p. 301).

OMM is not only conceptually defined as broadened attention, as studies show that engaging in regular OMM produces this effect. For instance, Slagter et al. (2007) found that subjects who had 3 months of Vipassana meditation (a meditation similar to OMM) training showed greater attentional resource allocation when presented with two competing target stimuli. Those who did not have such training were less likely to notice the second of the two stimuli, as their attentional resources were consumed by the first target stimuli. Having been trained in OMM seems to enable people to expand their attentional resources, increasing awareness of the stimuli in their environment. Valentine and Sweet (1999) found that those trained in OMM were more likely to notice unexpected

stimuli, when compared to those trained in Focused Attention Meditation (FAM). Hodgins and Adair (2010) found that those trained in mindfulness meditation (of which OMM is a constituent part) were more efficient and accurate when it came to visual attention. Delgado-Pastor et al. (2013) report the beneficial effects of OMM-type meditation on attention via physiological indicators: expert-trained Vipassana meditators "showed greater P3b amplitudes to the [auditory] target tone after meditation" (p. 207), where greater P3b amplitudes typically occur when unexpected stimuli is perceived and attended to. These results "suggest that expert Vipassana meditators showed increased attentional engagement after meditation" (ibid). Note that Delgado-Pastor et al. (2013) reported that increased P3b amplitudes to auditory tones were observed *after* meditation. Likewise, attentional blink effects found by Slagter et al. (2007) were measured *after* completion of 3 months of intensive meditation. Perhaps, then, OMM doesn't just broaden attention *while* one meditates, but regular OMM may also result in a broadened mind-set that would persist – at least to some extent – throughout other activities, including the undertaking of various techniques of goal pursuit.

Lastly, not only does OMM result in increased awareness and so likely mitigate the negative effects of attention narrowing that come along with the particularly efficacious methods of goal pursuit, preliminary research suggests that there is in fact a relationship between broadened attention, ethical awareness, and ethical action: Ruedy and Schweitzer (2010) note that "individuals who are less mindful may fail to recognize

ethical challenges or to appreciate conflicts of interest" (p. 73). Hong (2019) found that mindfulness mediates unintentional unethical behavior that often arises in pursuing performance goals. While much more research is needed, there is reason to think that regular OMM leads to broadened attention, enabling one to notice a variety of important moral features.

1.6.C: Speculations on the benefits of a hybrid account of goal pursuit & OMM

Thus far, I have discussed how regular OMM might counteract problematic consequences that are often present when engaging in goal pursuit. I now turn to speculate about how pursuit of virtue-relevant goals might counteract detrimental effects of OMM. I conclude that when regular OMM is practiced alongside the pursuit of virtue-relevant goals, OMM doesn't just help ameliorate potential problems that plague various techniques employed in goal pursuit, but pursuing virtue-relevant goals may also help ameliorate problems for OMM. I end this discussion by describing the virtuous mean of attentional breadth, which could perhaps be attained through combining particularly efficacious techniques of goal pursuit with regular OMM.

Recall that one potential problem of using particularly efficacious techniques of goal pursuit is analogous to what silencing accounts face, namely a failure to adequately appreciate moral values or alternatives even when one ought not to pursue these alternatives. Note, though, that another problem exists on the opposite extreme: rather

than a particular route of action or feature capturing all of one's attention, it would also be problematic if one were to equally appreciate all options without appropriately varying moral import. One might worry that, since OMM involves attending to one's thoughts and external environment in a *non-judgmental* manner, this does not lend to virtuous *judgment*. If OMM inclines one to become non-judgmental about moral concerns, we don't end up with a virtuous agent, but one who fails to draw any moral conclusions at all.[28] In a similar vein, recent research has also found that OMM interferes with one's autobiographical memory (Fujino et al., 2018) which could mean decreased access to one's moral identity, personal narrative, and moral ideals.[29] The general worry is that the psychological distance which OMM creates might also distance one from the 'good' and 'virtuous' parts of one's self, leading to less morally appropriate behavior, rather than more.

While OMM comes with problems of its own, it is worth speculating how engaging in OMM in conjunction with goal pursuit might alleviate such concerns. Due to their broadness and generality, superordinate goals have been described as akin to values or ideals (Carver & Scheier, 2001; Schwartz et al., 2001). If one sets and strives after goals, the use of OMM might not just help one become more attentive in general, but also

---

[28] I thank Sabrina Little for highlighting this problem. For a personal account of how taking up a non-judgmental mind-set can interfere with making evaluative judgments which morality requires, see Ratnayake's (2019) essay.

[29] I am grateful to an anonymous referee for bringing this research to my attention.

more aware of these importantly held (virtuous) values and goals. Ruedy and Schweitzer (2010), for instance, found that "those high in mindfulness are more inclined to bring their attention to their current internal experience, to actively observe and reflect on their thoughts and feelings" (p. 76). Likewise, Carlson (2013) argues that engaging in a metacognitive monitoring of one's current experience can remove informational barriers to self-knowledge. Ruedy and Schweitzer (2010) suggest that if one values being virtuous or moral, and one becomes more aware of such values, such experiences result in an "increase [in] the self-importance of moral identity . . . the importance an individual places on protecting or enhancing her moral self-image" (p. 77). Thus, OMM might lead to problems for virtue cultivation – including a sort of amoral detachment – when done in isolation. But, when carried out in conjunction with the pursuit of virtue-relevant goals, perhaps this leads to both a greater awareness and focus on one's moral values, rather than detachment from them.[30]

---

[30] Research indicates that OMM does produce the attention-broadening effects that I have noted, but that OMM practitioners often begin by focusing their attention on a particular object, such as with FAM. For instance, Slagter et al. (2007) detail Vipassana meditation as beginning by first "focusing or stabilizing concentration on an object such as the breath. . . [and then] broaden[ing] one's focus, cultivating a non-reactive form of sensory awareness or 'bare' attention" (2007, p. 1228). Likewise, Lippelt et al. (2014) make note that "once practitioners become familiar with the FAM technique and can easily sustain their attentional focus on an object for a considerable amount of time, they often progress to OMM" (2014, p. 1). This critical role that preliminary attention focusing plays in OMM makes for an interesting observation with respect to how goal pursuit – and the narrowed attention that comes along with it – could work in conjunction with OMM. Further research should look at whether the narrowed attention that occurs in goal pursuit might take a similar sort of functional role that FAM has played in helping a practitioner carry out OMM. I am thankful to an anonymous referee for highlighting this connection between FAM and OMM.

Another way of construing the possible benefit of practicing OMM in conjunction with goal pursuit is seen when we consider an application of the silencing thesis that is most plausibly conducive to virtue: Consider, for instance, the soldier who realizes she must enter battle and the sacrifices this will entail. Yet, if she dwells *too much* on what she stands to lose, this may keep her from following through with what she ought to do. The silencing thesis has something to provide when it comes to virtue cultivation, but perhaps it isn't exactly 'silencing'; rather maybe something more like 'quieting.'

If virtue involves that certain reasons or alternatives be 'quieted', but not necessarily 'silenced', broadening attention – through something like OMM – *in conjunction with* techniques that focus attention in the pursuit virtue-relevant goals might be able to do the trick. We might suspect that OMM enables one to notice more features or alternatives, without necessarily being swept up by them. Yet, the attention narrowing effects present in various techniques goal pursuit might also lead to subscribing these features or alternatives relative importance, enabling one to 'quiet' some reasons and pursue others.

One way to describe the potential beneficial outcome of this hybrid account of virtue cultivation is that such quieting of a variety of moral features results in the virtuous exercise, or *mean*, of the phronetic-related capacities of perception and evaluation. Such virtuous perception involves noticing and appreciating a spectrum of features and alternatives, rather than just the one option she ought (in the end) to pursue. Yet, virtuous

evaluation will also at least sometimes involve 'quieting down' certain features or alternatives, for even though the virtuous soldier ought not to overlook the severe loss (e.g. death) that could come in battle, the relative import she ascribes to this potential loss must be kept in check, as she ought to evaluate her moral duty to go into battle as even more weighty and pressing.

This virtuous mean of exercising the phronetic-related capacities of perception and evaluation can be seen in Figure 1.4, whereby this depicts a balancing between two extremes of excessive attention narrowing and excessive attention broadening.
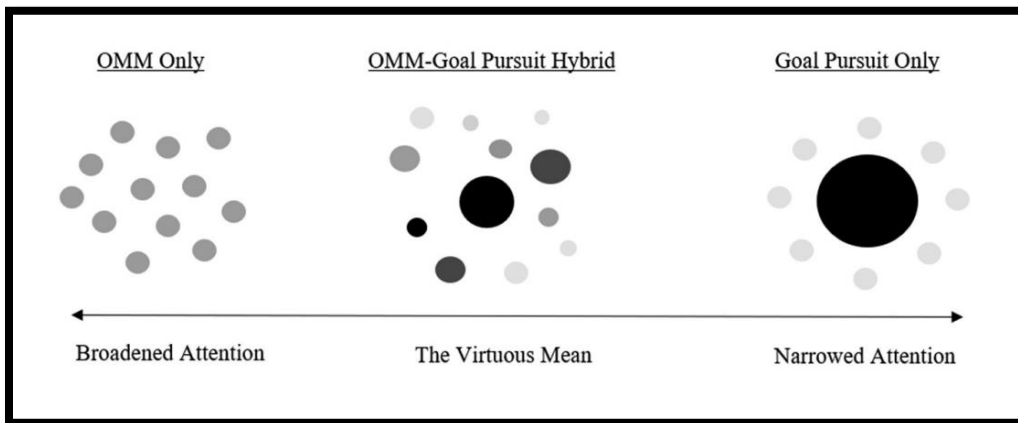
Figure 1.4. *Figure 1.4 shows where three proposals of modifying cognition and behavior (OMM-Only, OMM-Goal Pursuit Hybrid, and Goal Pursuit-Only) fall on the spectrum of attentional breadth and the evaluative importance given to various features.*

When attention is broadened too much, one is aware of a variety of objects or features, but remains non-judgmental, or fails to give any relative importance to such features. This is what we might expect with mere OMM. On the other hand, when attention is narrowed too much, one ignores all other features or options except for the sole item of her focused attention. This is what we might see when merely engaging in

the aforementioned techniques of goal pursuit – a strict sense of 'silencing'. Yet, somewhere in the middle between these two extremes is where we might find the virtuous balance: a 'quieting' of various features – to a greater or lesser degree. The virtuous person still notices a wide variety of features and appreciates their relative value. Yet, appropriate exercise of the phronetic-related capacity of evaluation will amount to an accurate weighting, where one judges these features or options as having different degrees of moral import in one's particular circumstance.[31]

1.7: Conclusion

In this paper, I have suggested that several goal pursuit accounts of virtue cultivation involve techniques that narrow attention. While attention narrowing often proves successful in modifying our behavior, problematic ethical consequences loom. By narrowing our attention, our phronetic-related skills can be threatened, leading to morally inappropriate actions. But, even where this does not occur, we might still fail to pick up on the moral intricacies of our situation that virtue requires of us. While I have argued this might not be problematic in cases where one is avoiding vice, cultivating

---

[31] One may wonder whether this hybrid proposal is threatened by what I suggest earlier in the paper, namely: If goal pursuit is most effective when attention is narrowed, then we should expect that introducing broadened attention will actually undermine goal pursuit. While it's possible that OMM has such detrimental effects, it's far from clear that it would. Earlier in the paper, I proposed that the lack of narrowed attention is what makes, at least in part, vague or broad goals rather ineffective. But why vague or broad goals don't often get satisfied is because one fails to maintain focus, and their attention gets swept by other distractions or temptations. Yet, the broadened attention that comes along with OMM is different; it isn't that one's attention 'bounces around', but rather it remains focused on a range of stimuli.

virtue requires a balancing of broadened and narrowed attention. I suggested that by using both goal pursuit and OMM, we might have a better chance in avoiding problems that plague either in isolation, and so lead to the cultivation of virtue.

Chapter 2: Virtuous Moral Illusions

2.1: Introduction

In an often-cited example, Gilbert Harman (1977) has us imagine ourselves "round[ing] a corner [to] see a group of young hoodlums pour gasoline on a cat and ignite it." In such a case, he suggests that "you [do] not need to conclude that what they are doing is wrong…you can see that it is wrong" (p. 7). The idea that we have a capacity for moral perception long predates Harman, as Aristotle not only describes something like moral perception, but also relates the *accuracy* of such perception to virtue: phronesis is the 'eye of the soul' (*NE* 1144a29–30); wise people 'see correctly.'[32] Similarly, Rosalind Hursthouse (2006) explains that the virtuous person "gets things right" (p. 103): The virtuous person not only acts rightly but also morally perceives accurately. In the past few decades, (Audi, 2013, 2018; Blum, 1994; Cowan, 2015; Cullison, 2010; Dancy, 2010; Goldie, 2007; many Jacobson, 2005; McBrayer, 2010; McDowell, 1998; McGrath, 2004, 2018; Roberts, 2003, 2013; Tappolet, 2016; Werner, 2016, 2020) have argued that we have moral perceptions or perceive moral properties, demands, or actions. But little attention has been given to one kind of *inaccurate* moral perception – namely *moral illusions* – and

---

[32] Aristotle holds that phronesis is a type of perception, for perception deals with particulars (*De anima*, 417b21–23, 417b27). Rather than following the deduction of universal ethical rules or principles (via logos or science), we ethically perceive our situation in a particular and unique manner (*NE*, 1142a23-30). For helpful discussion on this, see Rabinoff's Perception in Aristotle's Ethics (2018), specifically Ch. 4.

what this would mean for virtue.[33]   Just as we experience sensory illusions in sensory perception, we should wonder if there are moral illusions and whether we should expect the virtuous person to have them.

This paper investigates this underexplored phenomenon of moral illusions and how they relate to virtue and vice. While it is plausible that a virtuous person can have inaccurate moral perceptions from time to time, one might think it would be odd if one's virtue *manifests* itself in inaccurate moral perceptions.[34]   Yet, in sensory perception, according to a Bayesian computationalist framework, illusions *are* manifestations of a well-functioning and finely tuned sensory system.   So, plausibly, one also might experience moral illusions because of their well-functioning and finely tuned moral perceptual system.   And – for the virtuous – this finely tuned moral perceptual system is finely tuned to moral reality, having engaged in virtuous moral learning. For the virtuous person, moral illusions might very well be manifestations of their virtue.

I proceed as follows: In *Section 2.2*, I give a Bayesian, computationalist account for sensory perception and illusions. In *Section 2.3*, I look at how a Bayesian computationalist account could be employed for moral perception, predicting that we would also experience moral illusions. In *Section 2.4*, I explore what this means for virtue and vice.

---

[33] While the notion of 'moral illusion' has been largely neglected in the literature moral perception, the idea has been picked up on by Bruers (2016) and Moss (2009). Tappolet (2016) also mentions that recalcitrant emotions might be thought to be a kind of evaluative illusion.
[34] Just as the virtuous person can act out of character, the virtuous person plausibly can perceive out of character, or have an inaccurate moral perception. But when the virtuous person acts or perceives out of character, it is precisely *not* their character that is being manifested.

I look at how the virtuous and vicious person plausibly engage in moral perceptual

learning, such that the former would manifest in accurate moral perceptions and the latter

manifesting in inaccurate. I also examine how one's own agency would be involved in

forming the kind of moral perceptual system reflective of one's character. In *Section 2.5,*

I close by considering how the virtuous person should practically proceed when they are

under a moral illusion. I propose that we shouldn't necessarily think of the virtuous

person as always being autonomous. Rather, the virtuous person will sometimes rely on

another's moral testimony to help them navigate cases when they are under a moral

illusion.

2.2: A Bayesian explanation for sensory illusions

Consider the dots in Figure 2.1.a: The reader will likely see some circles as convex

and others as concave. But if we flip this image 180 degrees, as is depicted in Figure 2.1.b,

the circles that were perceived as convex are now seen as concave, and vice versa.[35] This

image is actually consistent with any of the circles being seen as either convex or concave.

One reason for why you view a particular circle as convex (or concave) is that your visual

system makes assumptions about where the light source is, and so in what direction

shadows are being cast from. It is thought that we have a light-from-above prior, or that

---

[35] Figure 2.1 is taken from Adams, et al. (2004).

our visual system assumes that light shines down from above (Adams et al., 2004). Because we live in a world where light (sun, lights on ceiling, etc.) generally does shine down from above, our visual system picks up on this information, stores it, and makes use of it in future perceptions.
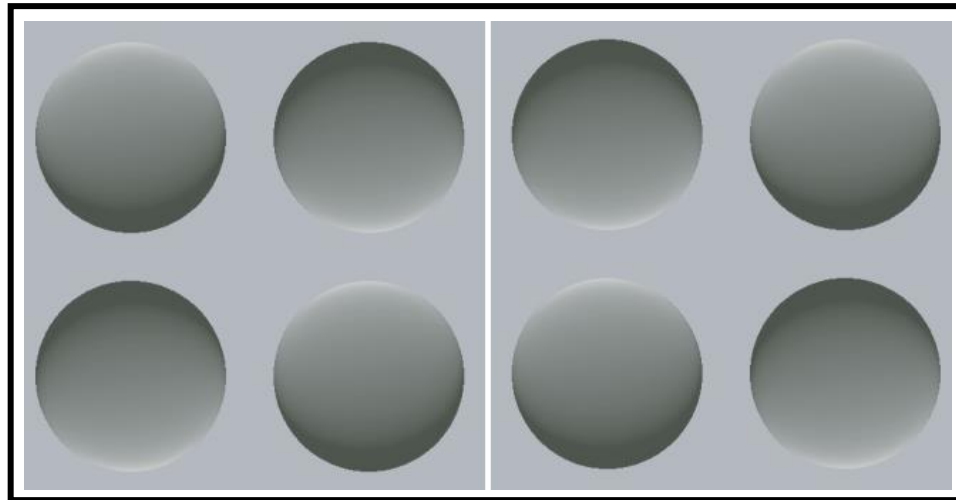
Figure 2.1. *Figure 2.1 illustrates our light-from-above priors. Figure 2.1.b (right) is merely just Figure 2.1.a (right) flipped 180 degrees. The circles that appeared concave in Figure 2.1.a now appear convex in Figure 2.1.b, and vice versa.*

This example of seeing circles as convex or concave illustrates how our perceptual system makes use of stored information – or priors – that are often learned over past experiences. According to a computationalist Bayesian framework, there are two basic components of perception: Incoming sensory information and assumptions about that incoming information.[36] Oftentimes, incoming sensory information is vague, uncertain,

---

[36] The reader may wonder why a Bayesian computationalist framework of perception is being privileged over alternatives. One reason is that alternatives – such as direct realist accounts of perception, including the Gibsonian ecological approach – face difficulties when accounting for genuine perceptual illusions (Smith, 2002, 2010; Millar, 2015; McLaughlin, 2010). Gregory (1997) explains that "[t]o maintain that perception is direct, without need of inference or knowledge, Gibson generally denied the phenomena of illusion" (p. 1122). While Gibson did discuss illusions, his explanation appealed to 'inadequate'

or ambiguous, and so our perceptual system uses stored information gathered from past experiences about what is the most likely percept. Put simply, the combination of these two components results in the percepts that we see, like convex or concave circles.

But oftentimes, the incoming stimuli consists of multiple features that are bound together, resulting in rich content. In such cases, stored information is used to make assumptions or predictions about how the individual features relate to each other. During perception, we use an "internal representation that includes single-element and co-occurrence statistics, as well as information about the *predictivity between elements*" (Avarguès-Weber et al. 2020, p. 25923, italics mine). As Fiser (2011) explains, our perceptual system forms the "*most likely …grouping* of its previous experience into independent representational units" (p. 141). Our perceptual system learns the statistics of our environment to make predictions about what elements or simple content tends to hang together, resulting in the rich perceptual content that we experience.

One example of this grouping occurs in multimodal perception: While it might seem that we have separate senses (sight, audition, taste, touch, etc.), they rarely operate in an isolated manner. Rather, incoming sensory information is taken in from various senses and integrated, giving rise to the percept that we experience. For example, when

---

information (Gibson, 1966, p. 288) and that we can rid ourselves of such illusions by employing "very special kind of selective attention" (ibid, p. 313). Since I wish to give an account of illusions – including moral illusions – I need to employ a model of perception that can adequately account for such phenomena, and this is why I make use of a Bayesian computationalist framework in this paper.

in conversation with another, the syllables we hear are the result of sensory information coming in from both audition and vision. However, how this information across various senses gets integrated is not always equal: In a review article, Landy and colleagues (2011) explain that "weights are assigned to the incoming sense data, in proportion to the reliability [of those sense modalities]" (p. 18). Priors – gathered from past experiences – are employed in assigning these weights in virtue of reliability of each sensory modality. Furthermore, the reliability of the different modalities, and so the weight assigned, is domain specific. O'Callaghan (2019) explains that "[v]ision carries more reliable information about spatial location than does audition…[While] in the temporal [realm]…auditory cues outweigh and dominate visual cues" (p. 39). So while our perceptual system weights visual information more heavily in the domain of speech perception, for instance, this is not always the case; visual information will be given less weight when it is generally less reliable, such as in the temporal domain.

Learning from past experiences and storing this information so to make use of it later is thought to be *optimal*, meaning that it is the best possible strategy for increasing reliability or maximizing accuracy over the long run.[37] Shams & Kim (2010) explain that "In carrying out basic perceptual tasks, the human perceptual system performs causal

---

[37] Optimality often assumes certain constraints, such as speed or time for instance. Reliability or accuracy over the long run could be further enhanced if one's perceptual system had an infinite amount of time to process information. But *given* that perceptual judgments or inferences need to be made quickly, making use of stored statistics that are learned from previous experiences with one's environment is the best possible strategy, maximizing accuracy in the long run.

inference and multisensory integration, and it does so in a fashion highly consistent with the Bayesian observer. This strategy is statistically optimal as it leads to minimizing…error" (p. 280).[38] In a similar vein, Seriès & Seitz (2013) explain that when one's priors reflect one's past experiences with their world, they are said to be optimal. Seriès & Seitz review various studies, showing how these priors can be updated across a lifetime, so as to match the statistics of one's environment, helping us navigate our physical world, maximizing accuracy in the long run.

Optimality can be defined as follows: A perceptual system is optimal when it maximizes reliability or minimizes error in the long run. This will require effectively learning the statistics of one's environment, so that the percepts approximate the physical world around them. In addition to a perceptual system operating optimally, the percepts that proceed from such a system are also said to be optimal: A percept is optimal insofar as it is produced via an optimal perceptual system.[39]

While an optimal perceptual system minimizes error in the long run, this does not necessarily mean it is infallible. Rather, it is precisely because it learns environmental

---

[38] The term *Bayesian observer* refers to a model where the 'observer' optimally makes use of priors to arrive at the most likely posterior perceptual judgment about the stimulus.

[39] For clarity, it might also help to consider cases of a perceptual system which is *not* optimal: Empirical research has found that those who experience hallucinations – such as those with schizophrenia – often have a systematic difference in the weight placed on their perceptual priors (Powers et al., 2017; Cassidy et al., 2018; Alderson-Day et al., 2017). When too much weight is placed on perceptual priors, perception relies less on incoming sensory data and more on stored information, giving rise to hallucinations or perceiving things that aren't actually there. This overweighting of priors is *not* optimal and is an overall less reliable perceptual system than those of neurotypical individuals.

statistics and makes use of this information that a certain kind of *inaccurate* perception –

namely perceptual illusion – can arise. For instance, because our perceptual system

maximizes reliability by giving more weight to visual over auditory information during

speech perception, illusions like the McGurk Effect arise. In the McGurk Effect, what

phoneme we hear someone speaking is impacted by the speaker's lip that we visually

perceive.   Despite the same phoneme actually being made, the difference in visual

information in Figure 2.2.a ('fa') vs. Figure 2.2.b ('ba') results in a different sound being

heard.[40] Yet, such "crossmodal illusions are not mere abstractions or quirks of processing.

They do not stem from accidental interference" (O'Callaghan, 2019, p. 34). Rather,

"[crossmodal illusions] are intelligible responses to extraordinary circumstances…[yet

i]n unusual conditions, it leads to an illusion" (ibid).

---

[40] Figure 2.2 consists of still images take from a video, produced by Mark Mitton, Josh Aviner, and Susanna Mitton (Nov 7, 2011). The video can be found here:
https://www.youtube.com/watch?v=PWGeUztTkRA

Figure 2.2. *Figure 2.2.a (left) and Figure 2.2b (right) consists of two still images takes from a video that illustrates the McGurk Effect. The differing lip movements impacts the phonemes heard, despite identical sound stimuli.*

Another example of an optimal perceptual system giving rise to an inaccurate perception occurs in cases of motion illusions: We live in a world where objects are usually static or move very slowly. As a result, we acquire a slow-motion prior. While employing this prior helps minimize error in the long run, it can also give rise to illusions. When incoming sensory information is vague or uncertain, we rely more heavily on our priors. This means that in low-contrast visual environments, more weight will be given to the slow-motion prior. This maximizes reliability in the long run, since we live in a world where objects tend to move slow. Nonetheless, this can result in an illusion – such as the Stepping Foot Illusion, a still frame of which is depicted in Figure 2.3.a. In low contrast conditions, such as when the blue block moves over the black line or the yellow block moves over the white line, the block is perceived as moving slow, due to the heavy weighting of the slow-motion prior. In the high contrast condition, the block is perceived as moving fast(er) (Figure 2.3.b). This results in an illusory experience of the blocks moving at different speeds, and so appearing as a 'stepping' motion. However, when the

contrast conditions change – such as when the background goes grey, the motion is seen

as smooth, and the illusion dissipates (Figure 2.3.c).[41]
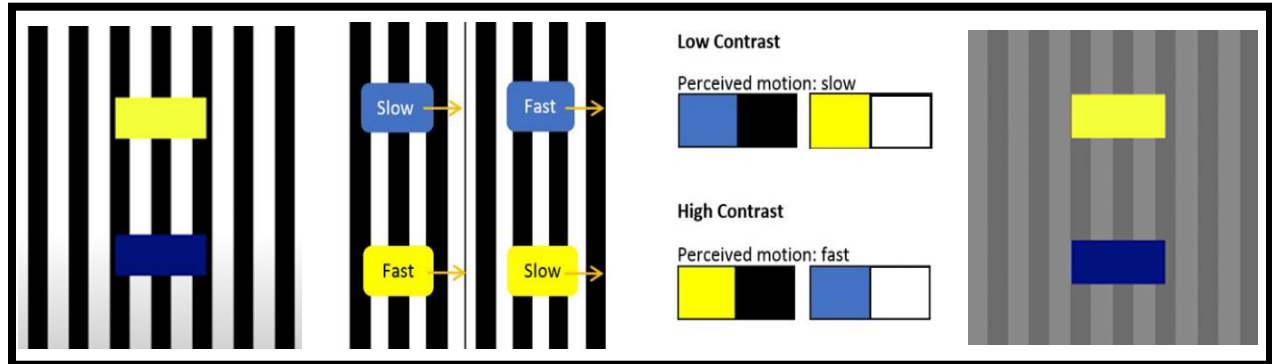
Figure 2.3. *Figure 2.3.a (left) and 2.3.c (right) consists of still images that illustrate the Slow Motion Ilusion. The illusion dissapates when the differences in contrasts disappear, as seen in Figure 2.3.c when the background turns to gray. Figure 2.3.b (middle) provides a visual explaining the the illusory stepping movements of the blocks.*

Weiss and colleagues (2002) explain that "many motion 'illusions' are not the result

of sloppy computation by various components in the visual system, but rather a result of

a coherent computational strategy that is optimal under reasonable assumptions" (p.

603). Stocker and Simoncelli (2006) give a similar explanation:

> [For an] observer who lives in a world in which slower motions are more likely to occur than faster ones and whose judgments are based on noisy measurements…the perceived speed and direction of a moving visual stimulus depends significantly on attributes other than its physical motion…[but that such] behavior can be seen as optimal (p. 578).

---

[41] Figure 2.3.a and 2.3.c consists of still images taken from a video produced by Hind Sight Grafyx, presented by the NJ Morris Museum (2015). Accessed December 19, 2022, at https://www.youtube.com/watch?v=wAMN1QW5ByM. Figure 7b was created by Rinueraeni (2018), Wikimedia Commons. Accessed December 19, 2022, at https://en.wikipedia.org/wiki/Stepping_feet_illusion

Thus, although the outcome might be an occasional illusion, this occurs precisely because the perceptual system is optimal, having learned and made use of probabilistic information that match environmental regularities, maximizing accuracy in the long run. In this section, I have summarized the Bayesian computationalist account of perception – how our perceptual system makes use of both incoming sensory data as well as stored statistical information which matches previous experience with one's environment. In doing so, we arrive at the most likely percept. When a perceptual system operates in this way, it can be said to be optimal, producing optimal percepts. But our perceptual system – precisely *because* it operates in this optimal manner – also gives rise to occasional illusions. In what follows, I'll turn to moral perception and examine what applying this Bayesian computational account would amount to for moral perception and the implications it has for virtue and vice.

2.3: Moral perception and moral illusions

In this section, I give two possible interpretations of moral perception – one that is *actually* perception, and another that is perception-like intuition – and argue that we should expect both of them to operate much like our sensory perceptual systems, producing optimal moral percepts or percept-like intuitions. But if this is so, then we should also expect moral illusions to arise.

Before beginning, I should clarify that I do not intend to commit myself to the phenomenon of moral perception, nor am I giving these arguments to imply that the phenomenon does indeed exist. Rather, what I wish to do is give a conditional argument: *If* moral perception works in the way suggested, *then* we should expect it to operate in an optimal manner and result in moral illusions.

2.3.A: Moral perception: The argument from rich content

One argument given for moral perception begins with the observation that we seem to be able to perceive complex content – like dogs, tables, and trees. As mentioned above, we do this by making use of stored statistical information about the relations between the components, binding them together into complex units. This stored statistical information involves the *predictivity between elements* (Avarguès-Weber et al. 2020, p. 25923), forming the "*most likely …grouping* of its previous experience into independent representational units" (Fiser, 2011, p. 141). Our perceptual system learns statistical information about what is the most probable complex percept, given the individual elements. Priors are made use of, resulting in rich content that we end up perceiving. As depicted in Figure 2.4, for instance, we do not just see black spots. Rather, our perceptual system makes use of stored statistical information about what

dogs look like. By making use of this prior, we perceive a dog rather than just black

spots.

Figure 2.4. *Figure 2.4 depicts the well-known image of the Dalmatian Dog, by R.C. James, showing how we can visually represent black spots as a Dalmatian.*

Some have argued that we might expect moral perception to work in a similar way:

If we can bind simple content into more complex content to perceive morally relevant

perceptual stimuli – like perceiving another's pain, perceiving intentionality in action,

and perceiving another's emotions – then why couldn't that content further get bound

together to create a more complex perception that has *moral content* – such as wrongness,

badness, cruelness, or kindness.   Robert Audi (2013) compares perceiving moral

properties to perceiving high-level properties in aesthetics (p. 35-37). Robert Cowan

(2015) suggests that "if perceivers can represent, e.g. natural kinds, in experiences, then

it is perhaps less incredible that agents can have experiences of wrongness" (p. 668). In a

similar vein, Sarah McGrath (2010) argues that if we can perceive morally relevant

content, like another's pain or intentional action - then why couldn't those morally

relevant contents be further unified together, forming even more complex moral content, such as moral wrongness? (p. 165).

Furthermore, there is empirical evidence supporting the claim that we do perceive things like another's pain or intentionality in action: Singer and colleagues (2004) found that when one watches another individual in a pain-inducing situation – such as when one's finger is being sliced by a knife – this activates brain regions in the observer, like the ACC and anterior insula, which are also activated one experiences painful stimuli for oneself. This perception of another's pain is sensory, although perhaps more somatosensory, rather than visual. Insofar as we perceive and sense our own pain, so too it seems that we can perceive and sense another's pain. Additionally, there is some empirical research which indicates that we can perceive intentionality in action. Scholl & Goa (2013) imaged subjects' brains while either viewing moving arrows pointed towards a moving disc (Wolfpack condition) or viewing a similar scene, except the arrows were not pointed towards, but rather perpendicular to, the moving disc (Control condition). Subjects in the Wolfpack condition – a condition where subjects often report the arrows to be 'chasing' the disc – had regions in their brain activated that are involved in visual perception. Given that the lower-level features, like the movements of the disc and arrows, were the same across both conditions, the researchers concluded that our visual system seems to visually perceive intentionality.

So, we might think that these things are the basic building blocks of at least some moral perceptual content. If that's right, then it doesn't seem so strange to think that there might actually be a thing such as moral perception, and that it is a result of binding simpler, morally relevant, content together with the aid of stored statistical information which is reflective of one's moral environment.

But, if this is how moral perception works – namely, by learning and using priors based on what is the most probable or predictive grouping of simple elements – then we should likewise expect it to operate in an optimal manner, and so sometimes give rise to moral illusions. I will go on to further illustrate how moral illusions are optimal percepts, resulting from an optimal moral perceptual system below in *Section II.C.* But before that, I'd like to turn to one other way that moral 'perception' might operate – namely via moral intuition.

2.3.B: Moral intuitions as moral 'perception'

Perhaps we should not think of moral 'perception' as actual sensory perception but interpret 'perception' in a more metaphorical way. Perhaps what is actually going on in moral 'perception' is something like moral intuition. Intuition has often been said to be akin to perception, insofar as both are impressions or presentations of the world (Bengson, 2015; Chudnoff, 2020). So, if moral 'perception' is actually just moral intuition, then we should look at what mechanisms are at work in moral intuition to see if it, too, is

optimal, operating according to a Bayesian computational framework, and so might also give rise to moral illusions.

I think there is good reason to think that moral intuition does operate in an optimal manner and so will also result in moral illusions: It's recently been suggested that intuitions are 'smart' and 'rational' – "the result of learning complex statistical relationships" (Railton 2017, p. 182), "guiding behavioral selection via the balancing of costs, benefits, and risk" (Railton, 2014, p. 833), in a way that "approximates Bayesian updating" (ibid, p. 835). Woodward and Allan (2007) suggest that intuitions, including moral intuitions, are the result of not just the current environmental input, but involve "a complex repeated game of some kind" (p. 185) and are "based on experience-dependent probabilistic models" (p. 186). Nichols and colleagues (2016) propose that moral intuitions involve general moral rules, learned through tracking environmental regularities and statistical updating that are optimal, or approximate Bayesian learning (p. 549). Kleiman-Weiner and colleagues (2017) put forth an account of moral learning where noisy and incomplete observations, along with innate priors, are used to build more complex moral models that are then used later in Bayesian inferencing. Innate priors are updated, and the stored information begins to match environmental regularities, taking into account observed behavior of others and feedback of one's own behaviors.

To illustrate how moral intuitions are the result of 'learning complex statistical relationships' that involve 'a complex repeated game of some kind', consider Peter Railton's interpretation of Haidt's case of Julie and Mark who engage in incest. Haidt gives his subjects a vignette describing two siblings who engage in incestual sex on one occasion, where it's stipulated that they use birth control, and no negative psychological consequences occur. In fact it "makes them feel even closer to each other" (2001, p. 814). After reading the vignette, subjects are then asked if it was okay for Mark and Julie to make love. Subjects often say 'no' but have trouble coming up with reasons that are applicable to the particular features of this case. They will often cite reasons like potential genetic defects to the offspring or emotional trauma of Julie and Mark. But these consequences are precisely ruled out by the stipulations of the case. When pressed, subjects are dumbfounded; Haidt concludes that our moral intuitions are not rational, and often had without reason.

Railton (2014), however, argues that our moral intuitions are smart and rational, especially once we look at how they get things right over the long run. Railton gives an analogous case to illustrate this, showing why the subjects' intuitions about the wrongness of Mark and Julie's actions aren't irrational or without reason. This case is of Jane and Matthew: Similar to Mark and Julie, Jane and Matthew are brother and sister and decide to engage in playing Russian Roulette on one particular occasion: "As it happens, the gun does not go off, and neither suffers any lasting trauma from the

experience. They both enjoyed the game but decide not to do it again. They keep that night as a special secret, which makes them feel even closer to each other." (Railton, 2014, p. 849)

Railton thinks that, in this case, it is obvious why it was wrong for Jane and Matthew to play Russian Roulette: "Jane and Matthew carelessly put all this at risk for the sake of a potentially 'interesting and fun' evening. [This is n]ot OK, despite the fortunate outcome" (p. 849). Even though Jane and Matthew come out unscathed in this particular instance, if they were to *keep playing* Russian Roulette, it would not end well. In entering into this social arrangement, both Jane and Matthew put each other at risk, and this *risk* they pose on one another is morally wrong.[42]

Our moral intuitions are based on running the probabilities and registering the risks, over the long run. In this way, our moral intuitions involve a 'repeated game', based on 'learning complex statistical relationships.' Our moral intuitions are optimal, getting things right *in the long run.* However, because our moral intuitions are formed by learning the statistics of our environment in an optimal manner (Nichols et al, 2016, p. 551; Railton 2014, p. 835; Kleiman-Weiner et al., 2017), we should also expect moral illusions to arise.

---

[42] Jane would not agree to play this game without Matthew and vice versa. Thus, by entering into this social agreement, they put each other's lives at great risk, which is morally wrong.

2.3.C: Moral illusions as optimal percepts

Now I'd like to give two examples of moral illusions that we could expect to occur within a Bayesian computational model of moral perception. In other words, these moral illusions would be ones that arise from optimal moral perceptual systems, arising from moral perceptual learning that maximizes reliability in the long run. While the first case – JESSICA AND MICHAEL – is a variant of Railton's case and quite toy-like, the aim of it is to help the reader understand how an inaccurate moral perception or intuition might be the result of an optimal moral perceptual system. The second case – LISA THE NURSE – is a more realistic case that will be used throughout the rest of the paper as further issues of virtue, vice and moral testimony are brought to bear.

> JESSICA AND MICHAEL: Jessica and Michael are hiking together in the middle of Alaska. One evening while hiking, they come across a gun with ~7 million chambers but has only one chamber loaded. They decide it would be interesting and fun if they tried playing Russian Roulette. As it happens, the gun does not go off, and neither suffers any lasting trauma from the experience. They both enjoyed the game but decide not to do it again. They keep that night as a special secret, which makes them feel even closer to each other.[43]

Note that in firing the gun at themselves, Jessica and Michael have approximately the same risk of dying as they do if they were to go on a ride in the car for fun. If your perception in this case is that it is still too risky for Jessica and Michael to fire the gun, then *you* are likely experiencing a moral illusion. But plausibly, this illusion is optimal. This illusion arises because the moral norms, rules, and statistics that we use to get

---

[43] I am grateful to Allan Hazlett for suggesting these modifications to Railton's original case.

around in our world are reliable and maximize accuracy in the long run. But, when we employ these learned norms, rules, and statistics in this hypothetical situation – which is quite unlike our everyday experiences in our own world – we are led astray, and an illusion arises.

Consider the following facts about our own actual world:

1) We live in a world where 7 million chamber guns do not yet exist. Our intuitions that have been formed around firing guns are not reflective of these kinds of guns.

2) We have been taught, again and again, to never point a gun at someone, even if it is unloaded.

3) We have also likely been taught, again and again, to never fire a gun unless you want to fire a bullet.

Given the great risk of harm in firing a gun, acquiring these sorts of intuitions about gun shooting seems quite optimal. But this nonetheless may also result in an illusory perception regarding the riskiness of playing Russian Roulette in the hypothetical case when one uses 7 million chamber gun.

Now consider a second case – one in which the reader doesn't experience a moral illusion, but the person in the vignette does.

LISA THE NURSE: Lisa has spent her career as a hospice nurse, caring 1-on-1 for a handful of patients at a time, providing personalized care and support. She gets to know her patients and their families and walks with them through the dying process. In March 2020, the Covid-19 pandemic hit Lisa's community and she was quickly asked to pivot to tend to patients at a make-shift field hospital. These patients were not the sickest ones, were all expected to live, and were usually discharged after a few days. Lisa's new work involved seeing hundreds of

patients over the course of a week, and often with not much consistency. The name of the game here was quantity over quality. Given this, Lisa felt like her new job was asking her to do the morally wrong thing – that she was being asked to treat her patients like numbers; to treat them as objects rather than subjects; not walking with them and hearing and being part of their story.

Given the nature of Lisa's previous work experience, she learned to encounter her patients in a particular way. But now that she is placed in a drastically different situation – one that requires her to shift her ratio of time and care per patient, the priors that were acquired over her time as a hospice nurse also are now what give rise to a moral illusion in this new environment. The way that Lisa is asked to carry out her jobs *feels* or *seems* wrong; she perceives herself as acting wrongly.

One way to describe moral illusions is that they involve an inaccurate sensing of the relative weight or importance among conflicting moral values or considerations.[44] We could construe Lisa's case to involve an inaccurate overweighting of the value of giving personalized care, and inaccurate underweighting of the value of maximizing care for all patients. Lisa mistakenly misperceives the weights of the moral features, and so has an inaccurate all-things-considered perception about what feature wins out.[45]

---

[44] It's possible, and even likely, that moral illusions can arise in various ways. This inaccurate weighting among conflicting moral considerations is just one way that we could think of a moral illusion taking place.
[45] To be clear, these moral illusions are not just moral dilemmas. In a moral dilemma, there are two moral requirements that cannot both be satisfied – there is no right way about how to proceed. While I want to grant that there are cases of true moral dilemmas, I also want to make room for cases where there *is* a correct way to morally perceive and proceed when it comes to relative weighting of moral values, and that a moral illusion could involve misperceiving these weightings.

Recall that in Section *2.2*, we saw that vision is more heavily weighted than audition within the domain of speech perception. While this is an optimal strategy which maximizes accuracy over the long run, it nonetheless also gives rise to perceptual illusions, such as the McGurk Effect. This might provide us with a nice analogy for what is going on in at least some cases of moral illusions, such as Lisa's: The moral illusion that Lisa experiences is a result from past moral learning, whereby she accurately picked up on the statistics of her environment. However, relying on these priors in this new environment means that she continues to overweight highly personalized care for an individual patient while underweighting maximizing care. In her current situation, this results in a moral illusion.

2.4: Moral illusions and their relationship to virtue and vice

Thus far, I have suggested that insofar as we are thought to engage in moral perception, then it should also be expected that we should experience moral illusions Furthermore, I've pointed out that given a plausible computationalist Bayesian account of moral perception, such illusions arise precisely because the perceptual system operates in an optimal manner, maximizing accuracy in the long run. In this section, I will look at what this means for the kind of moral perception that we should expect a virtuous (and vicious) person to have. As noted early on in this paper, virtue has been thought to involve an accuracy condition – a virtuous person "gets things right" (Hursthouse, 2006, p. 103) and "see[s] correctly" (Aristotle, *NE* 1144a29–30). However, I will suggest that

we should expect the virtuous person to be subject to moral illusions, and so the virtuous person will sometimes perceive incorrectly. Furthermore, these inaccurate perceptions are not ones that are simply out of character but arise precisely because of the excellences and virtues of the virtuous person having the moral perceptual system that they do – namely learning from one's virtue-conducive environment and employing this stored information an in an optimal manner. However, plausibly, this optimality is also present in most kinds of moral perceptual learning – including moral perception had by vicious persons. This will lead me to flesh out the difference between vicious and virtuous optimal moral perception.

Insofar as our moral perceptual system is optimal, it will learn the statistics from the environment and make use of this information in future perceptions, maximizing accuracy in the long run. While this is plausibly an *excellence* of the system – for a perceptual system that does not update or learn in this way will be less reliable – it is not necessarily just an excellence of a moral perceptual system that a virtuous person would have. Rather, whether one is virtuous or vicious (or somewhere in between), one's moral perceptual system will likely learn from one's moral environment in this way, updating itself given the statistical information, and making use of this information in future occasions to arrive at the most probable percept.

So, in addition to having an optimal moral perceptional system, the virtuous person will also need to have the right virtue-conducive experiences to provide the

virtue-conducive statistical information to learn on. There might be some environments

and experiences – perhaps, situations of extreme violence or environments structured by

racism and sexism – that, when learning optimally from them, one is *unable* to acquire

reliable moral perceptual faculties, or morally perceive the world the way that the

virtuous person would. Having the moral perceptual faculties of a virtuous person likely

requires having certain previous experiences with particular environmental regularities

(and not others).

At this point though, the reader might have a few worries: First, if our moral

perceptual priors update based on the statistics of our moral environment and so

influences future moral perceptions, does this mean that our environment just completely

determines the content that we morally perceive? In sensory perception, if one is placed

in an environment where light shines up from below for a long enough time, one's priors

begin to change to reflect this experience (Adams et al., 2004). This unique environment

seems to completely determine the priors that one's visual system learns and makes use

of in future perception. If moral perception is like this, then this seems problematic for a

few reasons. First, moral perception seems largely sub-personal, operating not at a level

of agency and character but at a level of mere computations. Second, if two people occupy

the same environment for a long enough period of time, then we should expect the

perceptual content of moral perceptual systems will be more or less identical. *But* this

seems dubious: quite plausibly two different people, with two very different characters

could occupy the same environment for several years and continue to morally perceive differently. We expect our characters and the moral perceptual systems they inform to persist through various situations and environments, rather than completely conform to them. We expect that a vicious and virtuous person could occupy the same environments and *retain their* virtues and vices, rather than converging to be identical in character. Call this issue the ***Sub-Personal Problem.***

In response, it must be pointed out that even if it is true that one's environment completely determines, forms, and updates one's moral perceptual system, this need not exclude agency and character from the picture. This is because what environment one *chooses* to enter into may be where one's agency and character is exercised. One might know, for instance, what being in a toxic work environment can do *to* oneself, and so it is for this reason that one selects one's work environment quite carefully. Rodgers & Warmke (2015) argue that being a "good situation-chooser" (p. 18) "is a disposition to act for certain kinds of reason: a virtue" (p. 22). The life of the virtuous person will involve an awareness of how one's moral environment might passively shape one's moral perceptual faculties and so will place oneself in virtue-conducive environments while avoiding vice-conducive ones.

Aside from choosing which physical environments to place oneself in, our agency and character is also often manifested in how we *subjectively* construe our *objective* situations or environments. Social psychologists Ross and Nesbitt (1991) explain that

"[t]here is significant variability in a given person's construal of events, enough to lead us, just on the grounds of interpretive instability, to expect that there will be nontrivial variation in behavior across two objectively almost identical situations" (p. 68) Dan Russell (2014) advances an Aristotelian account of virtue which largely rests on construing situations appropriately: "to act consistently for a certain good goal and construe situations appropriately in terms of that goal, is to have a virtue" (p. 54). What we take to be reasons, and what motivates us to act, is in response to how we subjectively construe the situation we are in. To construe our situation in a morally appropriate or accurate way, to detect the right moral reasons, and to act for those reasons, is what virtue involves. Nancy Sherman (1991) similarly observes the relationship between virtue and subjective construal: "The agent will be responsible for how the situation appears as well as for omissions and distortions. Accordingly, much of the work of virtue will rest in knowing how to construe the case" (p. 29)

In short, a virtuous and vicious person may subjectively construe their situations quite differently even if the objective features of their environment are the same. The environmental statistics that one's moral perceptual system is trained on might best be thought of as the statistics of *the subjectively construed environment*. And so, the regularities that one experiences and learns will themselves be a result of one's character. Given this, one's *subjectively construed environment* might determine the priors that are learned,

stored, and made use of in one's moral perceptual system, but this subjectively construed environment is itself a result of one's virtue or vice.

In addition to the Sub-personal Problem, the reader might also be concerned about what I'll call the ***Moral Feedback Problem****:* In what way can one be said to learn the *moral* statistical information of one's *moral* environment? In sensory perception, our perceptual system updates and learns based on feedback. When our perceptual predictions or priors fail, we don't effectively get around the physical world: we misperceive depth and bump into tables and chairs; we fail to register shadows and fall into potholes. When don't perceive accurately, we get clear feedback about the statistics of our external environment, and our perceptual system updates on this. But there doesn't seem to be similar moral properties and objects that we necessarily "bump into" when we get things wrong. Moral feedback doesn't seem clear and obvious in the same way that physical feedback is. So in what way can our moral perceptual system be said to learn the moral statistics of our moral environment?

While moral feedback is almost always more implicit than physical feedback, consisting of things like subtle bodily gestures, a particular emoji sent over text, or the tone in a friend's voice, that does not mean it isn't picked up and updated on. What we morally "bump into" and what we update on are things like social norms, hurt feelings, or acknowledgements of gratitude. The moral information and feedback we receive might also come from reading novels, learning about historical events, or adopting

65

certain religious beliefs or frameworks. This information may change our perception of

the goodness or badness (and the relative degrees of these properties) of certain actions

or states of affairs (e.g. death, suffering, etc.).   And, as mentioned above, what moral

feedback our moral perceptual system updates on will likely be a result of how we

subjectively construe these environments and feedback. Moral praise given from one's

church pastor might positively reinforce the pious churchgoer's moral perceptual priors

but have a completely different effect on the apathetic teenager who has no respect for

their parent's religion.   So how one interprets the feedback and how this subjectively

construed information updates one's moral perceptual system will likely be influenced

by one's character, moral concerns, values, beliefs, etc.

Another way that moral feedback is different than physical feedback is that while

physical feedback provides us with information about what is actually out there, it is

plausible that one can update their moral perceptual systems even when the feedback or

information updates on does not necessarily track *actual* moral reality.   Just as one can

update on social norms, even when those norms do not track moral truths or moral

reality, one can likewise engage in moral updating, picking up on the moral statistics of

their environment. Yet, these "moral statistics" should be understood as environmental

regularities that put forth morally relevant features, even when this departs from actual

moral reality.   For example, one could pick up on the moral statistics of one's

environment which puts forth false moral information that people of a certain race are

without human dignity. Updating based on this 'information' can still be done in an optimal manner, even though it does not actually track moral reality. Moral feedback for this person will not necessarily lead to updates which make their moral perceptual system better at navigating *moral reality*. Rather, such feedback will just make one better at minimizing 'errors', relative to their moral environment. This means that vicious, as well as virtuous, people engage in moral updating, picking up on the moral statistics of their (subjectively construed) environment, whereby this is done in an optimal manner.

We might now have some rough sketch of what the moral perceptual system of the virtuous person – and perhaps the vicious person – looks like: Both the virtuous and vicious person will have a moral perceptual system that will update given moral feedback. Their moral perceptual systems will be optimal, meaning that they will give rise to moral perceptions which will maximize reliability in the long run.[46] However, *what* this 'reliability' is relative to will differ: For the vicious, the moral feedback they will receive will lead them to navigate their environment in a way that leads to minimal disruption or "bumps", but this environment and effective navigation does not actually match up with moral reality (e.g. white slaveowners in the antebellum south who continued to operate business as usual). In contrast, the virtuous person will have a moral

---

[46] It is also possible that some people could have non-optimal moral perceptual systems, analogous to those who have non-optimal perceptual systems, as mentioned in footnote 7. Some neuroatypical persons, such as those with schizophrenia, are thought to have non-optimal sensory perceptual systems whereby they systematically overweight their priors and so regularly experience hallucinations. Plausibly, some individuals may have non-optimal moral perceptual systems and systematically overweight (or underweight) their moral priors, negatively impacting the reliability of their moral perceptions.

perceptual system that will update and learn on moral feedback which *actually reflects what is morally true*, rather than merely one's (morally deprived) environment. This is because the virtuous person will both select and construe their environment so that it is one which promotes and sustains truth-tracking moral perceptual facilities.

However, whether one is virtuous or vicious, insofar as their moral perceptual systems are optimal and update on the statistics of their moral environment, we should expect occasional moral illusions to arise. This means that for the virtuous person, she will occasionally experience inaccurate moral perceptions *precisely because* of the excellences and virtues she has: Given her virtues and the (objective as well as subjectively construed) environments she places herself in to, along with her optimal moral perceptual system which updates based on this statistical information of her environment, she will be much more likely to morally perceive what is actually morally true. However, when she finds herself in an environment that is atypical, or unlike that which her moral perceptual system has been trained on – such as Lisa the Nurse – her priors will lead her astray, resulting in an illusion.

A further interesting implication of this framework is that the vicious person might actually have an *accurate* perception in the precise situation where the virtuous person experiences a moral illusion. In other words, in at least some cases, the vicious could morally perceive correctly *precisely because* of their vice.[47] One example of this

---

might be that in wartime: a vicious person might have the correct moral perception that killing enemy soldiers is good, while the virtuous person might incorrectly see it as bad. I will come back to this issue, to some extent, below, but it surely requires more attention than I can give it.

2.5: How Should a Virtuous Person Proceed?

If it is true that the virtuous person will at least sometimes morally perceive incorrectly *because* of their excellences and virtues, then this leaves us with a few pressing questions: 1) how should this impact the way we conceive of the virtuous person? 2) how should the virtuous person proceed when they are under a moral illusion?

If it is right that moral illusions can arise out of one's virtue, then I think this should prompt us to reconsider the accuracy condition of virtue. The virtuous person might not always "get things right" in their moral perceptions; rather, virtue might lead one to get things wrong. However, the reader might resist this by pointing out that the existence of moral illusions might not apply to *the virtuous person,* for the virtuous person is an ideal, not any one of us human beings (Annas, 2004, p. 67). While it might be true that human beings – with all of their Bayesian computational perceptual quirks – are subject to moral illusions, no human being is fully virtuous anyways. So the *ideal* virtuous person does, indeed, get things right.

Assuming that there is such a thing as moral perception and that we should also expect moral illusions to arise, I don't think this should only impact the way we that we

conceive of human persons *en route* to virtue. It will also have implications for the notion of the *ideal* virtuous person. In other words, even if the virtuous person is an ideal that can only be approximated (and never reached) by any one of us, the phenomena of moral illusions ought to impact the way we conceive of this ideal – specifically, it ought to prompt us to do away with the accuracy condition.

Before giving my argument for why this is so, I want to first clarify the notion of the ideal virtuous person that I am working with. There are various senses of an ideal and the way that this ideal enters into the way we think about and pursue virtue. For instance, an ideal could be understood as something that we appreciate as excellent or deem as good; call this the *exemplary* sense. Another way of illustrating the *exemplary* sense of the ideal virtuous person is that, if it is true that we humans are subject to moral illusions and would not be able to ever reach fully the ideal of virtue, this would be unfortunate. There is something to mourn about the fact that we can never embody the ideal that we see as excellent. This exemplary sense of the ideal is *not* one that I necessarily think we ought to modify. Rather, there is another sense that I am concerned with: the *emulatory* sense. The emulatory sense is when the ideal virtuous person provides us with a goalpost to strive towards and with a psychology that we strive to mold our own to be more like. This ideal of the virtuous person plays an important role as we try to approximate it and make our own characters – including our moral perceptual faculties – more similar to it. Insofar as our moral perceptual system is

70

optimal and subject to moral illusions, it is this sense – the *emulatory* sense – of the ideal that I suggest we ought to adjust.

To show why we should modify the emulatory sense of the ideal virtuous person, I want to begin by making use of the concept of an *illusive ideal*, a notion which I get from Siversten (2019). In short, an illusive ideal is one where progress within one facet of the ideal undermines progress made in another facet. For example, Siversten argues that the informed, impartial spectator is an illusive ideal: the more informed one is, the less impartial one is, and vice versa. For a less academic example, take the ideal of an all-around excellent runner – one who is both an excellent endurance runner and also an excellent sprinter. It turns out that given the way our human muscular system works, we cannot be both, and that pursuit in one area undermines progress made in another. When one engages in sprinting-related training, her intermediate muscle fibers turn into fast twitch fibers, and so are necessarily not slow twitch. If this runner were to then aim at being an excellent endurance runner and so undergo endurance-related training, these intermediate muscle fibers would then become slow twitch, and necessarily be no longer fast twitch. The training that helps her progress within the realm of endurance running undermines progress within the realm of sprinting, and vice versa. Thus, the ideal of an all-around excellent runner – at least insofar as it is an *exemplary* ideal – is an illusive one, for when we try to approximate it, our efforts are self-undermining.

Insofar as our moral perceptual systems (if we have them) are subject to moral illusions, trying to mold our psychologies and cognition to be like the ideal virtuous person – one who morally perceives accurately in all environments – is on par with attempting the physical training necessary to be an all-around excellent runner. Our efforts in both realms will often be self-undermining. Suppose that one slowly changes her moral perceptual system through learning of new priors, given a new moral environment. This will amount to *changing* the moral priors she previously had, or losing progress made in other moral environments. In visual perception, if one is placed in an environment where light is made to shine from below, one will learn the statistics of this environment, and adopt a light-from-below prior. But, as a result, one's light-from-*above* prior will be 'unlearned' or greatly weakened. When placed in a situation where light shines down from above, one's new learning (e.g. adopting light-from-*below* priors) will result in an inaccurate perception. In short, one cannot have *both* light-from-above *and* light-from-below priors. So, to have accurate perception in all environments is an illusive ideal. So, too then, with moral perception. Obviously, it would be great if one could accurately morally perceive across a wide variety of environments, instantaneously adjusting their priors. And, likewise, it would be great if one's intermediate muscle fibers could be trained to be both fast and slow twitch, whichever being relevantly activated at a moment's notice, given the race she is partaking in. But that's just not how the human muscular system – and plausibly, not the human moral perceptual system – works.

Thus, if we do not modify the ideal of the virtuous person and continue to strive to perceive accurately in all moral situations, we will engage in a self-undermining pursuit, making our moral lives into frustrating and counterproductive ones. One modification, then, is that we should not try to pursue accurate moral perception across all environments, but rather should make room for something else to take its place. I suggest a positive modification which includes a social component – that of relying on another's moral testimony – within the ideal virtuous person. This leads us to addressing the second question I raised at the beginning of this section: how should the virtuous person proceed, when under a moral illusion?

Recall the case of Lisa the Nurse, who is thrown from her work as a hospice nurse into a makeshift hospital to manage Covid-19 patients. Now imagine one more background piece to her story: Lisa has an old nursing colleague, who she admires and takes to be a wise and virtuous person. This old colleague is now her manager in the makeshift field hospital. Before the pandemic hit, Lisa's manager worked as a public health nurse and has cared for patients in previous epidemic scenarios. Lisa's manager is aware of Lisa's background and training as a hospice nurse. Upon entering this new environment, Lisa's manager pulls her aside and tells her that the way that she will be asked to care for her new patients will *seem* wrong to her. But this is because of her training as a hospice nurse. Her manager asks her to trust her – what she will be asked

to do *is* the morally right thing. She ought to follow through with it, even if it appears wrong.

As described earlier, Lisa is plausibly under a moral illusion when she inaccurately perceives that personalized care outweighs maximizing care *in this new situation.* And, this illusion arises precisely because of her previous training, which made her a good and virtuous nurse. What is Lisa to do in this new situation? I think Lisa ought to defer to her trustworthy a manger and friend. Furthermore, I think this is what virtue amounts to in this case. Virtue might not always amount to *accurate moral perceptions* since it might also give rise to moral illusions. But, when it does, I suggest we reach into our moral toolbox and make use of other moral tools – like that of relying on another's moral testimony.

When we think of what resources are in our moral toolbox to help us deal with moral illusions, we can find inspiration from how we handle navigating non-moral perceptual illusions. When we encounter the Müller-Lyer illusion, for instance, we experience the lines as two different lengths when they actually are not. One way to acquire knowledge about the correct length of the lines could involve deferring to other folks who are not susceptible to this illusion. Some empirical research indicates that those who live in "carpentered" environments – spaces structured by straight lines, right angles, and square corners, such as we often see in Western cities, for instance – are more susceptible to this illusion than those who live in "uncarpentered" environments, where

dome-like structures are more typical. The Carpentered-world hypothesis posits that that the difference in susceptibility to the Müller-Lyer illusion is explained by differences in living environment, and the information that our visual system is updated on (Segall et al., 1966; Gregory 1968, 2009; Stewart, 1973). Those who inhabit carpentered environments learned the statistics of their physical environment, resulting in particular perceptual priors. But employing these priors lead them astray in the Müller-Lyer case. However, such people could arrive at a true belief about the length of the lines via relying on another's testimony – like those who have occupied uncarpentered environments, and so have different perceptual priors.

So too with moral perception: Lisa and her manager have different moral experiences and have occupied different moral environments. We can expect them to have learned different moral priors, and so see the world differently. When they are in their usual environments, they morally perceive correctly. But a quick switch of environments – like what Lisa undergoes – could mean experiencing moral illusions. And so Lisa might benefit from deferring to her virtuous and trustworthy manager, who has occupied similar environments for some time and has priors that reflect the statistics of this kind of environment.

However, I suspect my suggestion will receive pushback. Recently, several philosophers have claimed that relying on another person's moral testimony is at odds with being virtuous: it is a "moral defect of character" (Crisp, 2014, p. 132) and "a way of

falling short" (Hills, 2009, p. 113). The virtuous person shouldn't need to rely on another person's moral say-so, because a virtuous person should be able to figure out what to do for herself. A virtuous person is her own moral compass (ibid, p. 112), having the ability to "see what is true in every set of circumstances, being like a carpenter's rule or measure" (Aristotle, NE, 1113a32–1113a34). Thus, many might resist my suggestion that Lisa is both virtuous *and* ought to rely on another's moral testimony.

But I think this ideal of the virtuous person – at least insofar as it is an emulatory one that guides our own pursuits - is just an illusive one. This is because, if humans have such a thing as moral perception, it is plausible that it will give rise to moral illusions; virtuous human beings will not always be able to see what is true, in every set of circumstances. Rather, the same disposition that causes one to perceive accurately in one kind of environment will also be at work in causing inaccurate perceptions in another. And so to hold the emulatory ideal of the virtuous person as one who is her own moral compass, across a variety of situations that call for a variety of moral perceptual priors, is one that will likely prove to be self-undermining, futile, and frustrating when one tries to pursue it.

2.6: Lingering questions and concluding thoughts

One thing that may complicate this proposal is the fact that the vicious person may *also* have an accurate moral perception in the precise situation where the virtuous person experiences a moral illusion. As noted at the end of *Section III*, the vicious person also

has occupied different moral environments and adopted different moral perceptual priors. For instance, both Lisa's virtuous manager and her vicious co-worker might morally perceive that maximizing care for the greatest number of persons is the most important moral value or feature to pursue in this scenario. However, perhaps a vicious co-worker might have this moral perception because she doesn't see personalized care as good at all. Maybe she sees maximizing overall care as the only morally relevant value because the quicker patients can be treated and sent home, the sooner she can get reassigned from the field hospital and get back to her normal life. Sick patients are a disruption to her life; the only option is to treat them as quickly as possible.

Thus, one undesirable outcome of my argument might be that Lisa – a virtuous person – has no more reason to defer to her virtuous manager than her vicious co-worker, since both have the correct moral perception. Or, relatedly, even if Lisa should defer to her virtuous manager rather than her vicious co-worker, how will she know this? The reader might be worried that since both seem to have the same moral perceptions, which differ from Lisa's own, there will be no way for Lisa to figure out who the *virtuous* (and not the vicious) person is that she should defer to.

However even if Lisa's virtuous manager and her vicious co-worker have the same moral perceptions in this scenario, it's plausible that there would be other ways for Lisa to figure out which one is virtuous and which one vicious. For one, their intentions will likely differ, and Lisa's virtuous manager's intentions will likely be similar to her own,

given that she is also virtuous. Secondly, the vicious co-worker will likely have inaccurate moral perceptions in other domains that Lisa's manager will not, and Lisa could plausibly pick up on this: Lisa's vicious co-worker might perceive the janitor's chatting with patients as being lazy; Lisa's manager might perceive his actions as caring. Lisa could plausibly pick up on their varying attitudes and reactions and get a sense of who is virtuous and who is vicious. Relatedly, Lisa might *do better* if she were to defer to her virtuous manager rather than the vicious co-worker, for these differences in intentions and attitudes are likely not isolated from one's other moral perceptions and the ways that these other moral perceptions are linked up to actions and habits.

Nonetheless, much more needs to be done to thoroughly examine the potential role that the vicious person might have in these sorts of situations. Given the unusual circumstances in which moral illusions plague the virtuous person, it may very well be that the vicious person – and their moral perception – has an important role to play. This would be unexpected. But plausibly so are moral illusions, or that inaccurate moral perceptions arise precisely out of one's virtue. And this in itself, is what I have tried to show in this paper: if there is such a thing as moral perception, then we should expect that one's virtue doesn't just manifest itself in accurate moral perceptions, but also might be at work in cases of moral illusions.

Chapter 3: The Tension of the Virtues

3.1: Introduction

The unity of the virtues thesis has long been questioned. A quick look at those around us will make us doubt that if you have one virtue, you've got the rest. As Susan Wolf observes: "Gandhi was a paragon of courage, justice, and integrity, but he was a cold and unsympathetic husband. Mother Theresa was an exemplar of disciplined altruism but a harsh and difficult person…There is no obvious connection between [the virtues]." (Wolf, 2007, p. 146).

Interestingly, a handful of philosophers have gone even further – they argue that having one virtue might rule out having others. In short, they posit that there is a relationship between some of the virtues – but that relation may be negative in nature. Phillipa Foot suggests that perhaps one "can only become good in one way by being bad in another, as if …a kind of dull rigidity were the price of refusing to do what he himself

79

wanted at whatever cost to others" (1978, p. 397). A. M. D. Walker proposes that "The virtues must have their roots in, and be sustained by, the personality as a whole and different virtues requiring different types of personality, certain pairs of virtues can no more flourish in the same personality than different species of tree can flourish in the same soil" (1989, p. 352). Put even more strongly, Owen Flanagan argues that "[t]he idea of any individual possessing all of [the virtues] is incoherent. This is because some of the qualities on the list are inconsistent with one another and would, so to speak, cancel one another out. For example, vivaciousness, forthrightness, and physical courage are virtues. But so are serenity, tactfulness, and pacificism…the notion of one human individual possessing all the virtues in both subsets is not merely undesirable, it is impossible." (1991, p. 33).

Thanks to the situationist critique of virtue ethics made by those like John Doris (1998, 2002) and Gilbert Harman (1999), the past two decades have seen a boom in empirically informed accounts of character, virtue, and virtue cultivation (Miller, 2003, 2010, 2013, 2014; Westra, 2018, 2020, 2022; Stitcher, 2015, 2018, 2020, 2021; Snow, 2006, 2010, 2016). However, despite the fact that whether the virtues are in opposition to each other seems to, at the very least, involve empirical facts about human psychology, there is currently a gap in the literature when it comes to investigating what I'll call, The Tension of the Virtues Thesis, through an empirically informed lens. One aim of this

paper is to bridge this gap by considering whether, and how, particular virtues might be in psychological tension with each other.

In addition to addressing whether virtues are in tension with each other, this paper also considers further implications about what the truth of the Tension of the Virtues thesis would mean for our conception of the ideal virtuous person. If it is true that there are certain features of our psychology which result in some virtues being in tension with others, then we should modify our conception of the ideal virtuous person. I'll argue that our conception should be one whereby the ideal virtuous person is *not* necessarily one who has all of the virtues. In the last section of this paper, I suggest one way that we might begin to modify our notion of the ideal virtuous person, such that it might better take into account features of human psychology.

This paper proceeds as follows: I will first further consider the intuitive explanation for The Tension of Virtues Thesis that A. M. D. Walker points to – namely that certain virtues are in tension with each other due to underlying personality types. Plausible as it may sound, I argue that this explanation lacks empirical support, due to there being little empirical basis for the existence of personality types, or groupings of some personality traits but not others. I then investigate an alternative account that I think holds more promise – one based in sensory processing differences. By making use of a Bayesian predictive processing framework, I argue that people vary in how sensitive they are to sensory and affective information. Furthermore, I argue that some virtues will

benefit from or involve *reduced* sensitivity in affective processing while other virtues will benefit from or involve *increased* sensitivity. A person who is 'strong' with the latter set of virtues will be 'weak' with the former, and vice versa. Lastly, I'll then go on to address what this means for revising our conception of the ideal virtuous person. [48]

3.2: The explanation from personality types

One plausible explanation for why certain virtues may be in tension with each other invokes the importance of one's personality. The *type* of person who is good at being just is simply not also the kind of person who is good at being merciful.

Walker (1989) is not the first person to point to the role of personality types as the basis for why one person might have some virtues but not others. This line of thinking could be seen as having support from Aristotle: Aristotle explains how different people have different natural temperaments, and that some of these temperaments may be conducive to certain virtues but not others (1144bl-17). Greek medicine posits the theory of the Four Humors – or four basic temperaments – Melancholic, Choleric, Sanguine, and Phlegmatic. These temperaments, or personality types, group certain traits together and people were generally thought to fall into one of these four categories. Such personality types were thought to predict one's general dispositions of thinking, feeling, action, and being in the world. And certain personality types may be more or less conducive to

---

certain virtues: The choleric person might be well suited for being courageous but poorly suited for being compassionate or a thoughtful and wise reasoner.

While the Four Humors don't get much traction today, other theories of personality types are often made use of, especially within vocational domains like business leadership and healthcare. Friedman and Rosenman's Theory of Personality posits four personality **types** (A, B, C and D), and this framework has been used in medical and public health campaigns: People with a Type A personality are perfectionists but also experience a high degree of stress, and thus were thought to be at higher risk for developing high blood pressure and coronary heart disease (Friedman & Rosenman, 1974). Myers-Briggs posits that there are 16 personality types, each with their own strengths and weaknesses. The Enneagram posits 9 personality types, such as "the peacemaker", "the helper" or "the achiever." For any given personality type, this type is composed of certain traits and not others. For instance, if one is a "helper", then one is a people-pleaser and tends to put others needs before their own. Contrast this type with the "individualist" who is prone to being self-absorbed and self-centered. If one is a helper, one will have certain traits and not others, and some of the traits that the helper has are precisely the ones that do *not* compose other personality types – such as the individualist. Insofar as personality traits are the psychological stuff that character traits are built upon, we can start to see how falling into a given personality type (but not another) can provide the basic explanation for how certain virtues may be

in tension with each other. If one is a helper (and so not an individualist), they will likely have virtues related to helping others and likely lack virtues related to personal growth and achievement.

Although explaining The Tension of the Virtues Thesis via personality types is plausible and intuitive, it does not fare well against empirical evidence. Intuitive plausibility may be all that there is: "personality types were created via the clinical intuition of gifted psychiatrists, and they are not based on scientific methods" (Crocq, 2013, p. 148). For one, personality type theories show poor test-retest reliability. This means that if one were to take a personality type test and fall into a particular type – say, an individualist – this would not necessarily mean that they would be likely to fall into that same type when retaking the personality type test at a later date. Howes & Carskadon (1979) found that over 50% of subjects who retook the Myers-Briggs test 5 weeks later fell into a different personality type the second time around. Pittenger (1993) nicely summarizes the problem: "Briggs and Myers conceived of personality as an invariant. It is expected that by adulthood the personality preferences…will be stabilized and that test-retest reliabilities should be high…the [identified] reliabilities suggest that types have the potential of changing at each testing. If each of the 16 types is to represent a very different personality trait, it is hard to reconcile a test that allows individuals to make radical shifts in their type." (p. 471-2).

Second, personality types aren't great at predicting real world behaviors, and any predictive power they do have is just reducible to the particular degree of individual character traits. In considering the Asendorpf–Robins–Caspi (ARC) personality types – Undercontrolled, Overcontrolled, and Resilient – which are at least empirically supported when it comes to reliability, researchers nonetheless found they were still not predictive of behavior: Van Leeuwen and colleagues (2004) explain that these "personality types do not predict adolescent problem behaviour beyond what is predicted by personality dimensions" (p. 210) and that "five-factor [trait] measures predicted adolescent problem behaviour better than types" (ibid, p. 219). Similarly, Costa and colleagues (2002) found that "[a]ssociations [of traits] were shown to be solely a function of the trait information summarized by the types. Continuous trait scores were substantially better as predictors of the same criteria: on average, type membership predicted only 40% of the variance that could be accounted for by the five factors. (p. S84)

Rather than appealing to personality types, personality psychology largely makes use of personality traits instead. Both The Big 5 and HEXACO are often taken to be the 'gold standards' when it comes to personality trait theories, as they show a high degree of reliability and external validity, meaning they are predictive of real-world behaviors. However, most importantly, the precise notion of a personality type – whereby the traits are clustered together, such that individuals fall (or do not fall) into a given cluster – is unsupported by empirical research surrounding The Big 5. In summarizing the Big 5

taxonomy, John and colleagues (2008) note that correlations – whether positive or negative - between different dimensions are all very low, right around +/- 0.25 (p. 133). The authors explain that "the size of these intercorrelations represents barely 10% shared variance," (p. 133) meaning that only 10% of the (degree of) a given trait or dimension is explained by another.   In other words, scoring high (or low) on one trait – say extraversion – is not predictive of one's scores on other traits. But this just means that while the gold standard of personality embraces the construct of personality traits, it simultaneously casts doubt on the hanging or clustering together of personality traits, as personality type theories posit.

What this means then is that, insofar as personality traits plausibly compose or are relevant to the virtues one has, no particular personality trait or virtue is predictive of having (or not having) another. Thus, it seems like the empirical support is wanting for thinking that any given virtue would be constituent of a personality type, such that this personality type would significantly influence what other virtues we would be able to cultivate. In order words, appealing to personality types as the basis for The Tension of the Virtues Thesis seems wanting.

3.3: Building an alternative case: The relationship of pain sensitivity and empathy

Despite the lack of empirical support for personality types, and so the explanation they could play in explaining a possible tension between certain virtues, I think there is at least one empirically plausible alternative worth investigating. In the next section, I'll

argue that this alternative is one that points to differences in sensory processing, making use of the Bayesian predictive processing framework. But before getting to that, I want to turn our attention to a particular phenomenon concerning the relationship between pain sensitivity and emotional empathy. This particular relationship is a nice illustration of the broader framework that I will look at in the next section.

Several studies have pointed to the fact that endurance athletes, like Tyson Apostol (Figure 3.1) do well in their sport because they have decreased pain sensitivity – or increased pain thresholds and increased pain tolerance. This means that the point at which endurance athletes *detect* pain is a higher level of objective pain stimulus than non-athletes, and that endurance athletes can endure such stimuli for a longer period of time. Petterson and colleagues (2020), for instance, had endurance athletes and non-athletes engage in a cold pressure task, where subjects placed their hand in a bucket of ice water. Subjects then reported at what point they began to feel pain (pain threshold) and were asked to keep their hand in the water as long as they could stand it (pain tolerance). Endurance athletes showed both higher in their reported pain thresholds and pain tolerances. Gieser and colleagues (2021) similarly looked at differences in pain sensitivity between athletes and non-athletes, making use of painful heat (rather than cold) stimulation. Additionally, rather than using a self-report method, these researchers measured pain sensitivity via brain imagining. They found that "brain activations of

athletes versus nonathletes during painful heat stimulation revealed reduced activation

in several brain regions that are typically activated by nociceptive stimulation" (p. 5927).

<div style="border: 2px solid black; padding: 10px;">

<p align="center"><u>From the Mouth of Tyson Apostol</u></p>

- *"Lying to everybody, especially [to other contestants] Brendan and Sierra, actually brings me pleasure."*
- *"I love seeing people cry. When you crush their dreams."*
- *"I like to see Sierra scramble and mope, uh, I think it's funny because I've never liked Sierra. To me, she's of no worth."*
- *"I think she'll be really really upset when she gets voted off, and I'd like to see, uh, a freak out…I think that would be fun."*

</div>

Figure 3.1. *Figure 3.1 features a handful of quotes from Tyson Apostol, said about his fellow contestants on the game show Survivor. Apostol was a fierce a physical competitor on the show, as well as a professional endurance athlete. He nicely illustrates the relationship between pain sensitivity and empathy.*

But not only do endurance athletes show differences in their pain sensitivities, but

these differences are also correlated with differences in their degree of emotional

empathy, or empathy for pain.[49] Freund and colleagues (2013) compared ultramarathon

runners to recreational runners and likewise found that ultramarathon runners to have

decreased pain sensitivities. But additionally, this was also correlated with decreases in

emotional empathy scores.[50] This relationship, though, isn't actually too surprising. The

researchers note that this relationship is supported by a larger theory – The Shared

Representations Theory, whereby "[being] high [in] empathy for pain is associated with

a stronger activation of the affective parts of the pain network similarly to reactions to

---

[49] Empathy for pain is a component of emotional empathy (Fitzgibbon et al., 2010), and emotional empathy, is feeling what another person feels. This kind of empathy is contrasted with cognitive empathy, which involves taking the perspective of another person.

[50][50] Empathy scores were measured by the IRI, a standard, validated self-report measure of empathy.

physical pain" (p. 529). This theory is supported by a body of research which shows that emotional empathy makes use of the same neurological underpinnings – particularly, regions of the ACC and bilateral insula – that are used to sense pain in one's own body (Jackson et al., 2005; Lamm et al., 2011; Singer et al., 2004; Liu, et al., 2019). Mischkowski and colleagues (2016) found that taking prescription levels of the painkiller, acetaminophen, resulted in reductions in one's emotional empathy for feeling another person's pain. Danziger and colleagues (2006) also found that those with congenital deficits involving impairment in physical pain detection in their own bodies also showed decreased empathy for pain when viewing videos depicting another person in physical pain. Given this, we should actually just expect that if one – such as is the case with endurance athletes – has decreased pain sensitivity to their own physical pain, so too, they would be less sensitive to another's pain, or have decreased emotional empathy.

Given this relationship between pain sensitivity and emotional empathy, we might begin to see how certain virtues could be in tension with each other. Plausibly, the virtue of perseverance might involve or benefit from being *less* sensitive to one's own pain and uncomfortability, while the virtue of compassion might involve or benefit from *heightened* sensitivity to another's pain or having greater emotional empathy. I look at these claims further in the next section, after giving a more generalized empirical account of how certain virtues may be in tension with each other, given differences in affective sensitivity.

3.4: Generalizing the account via Bayesian Predictive Processing

While there is this interesting empirical relationship between physical pain sensitivity and emotional empathy, in this section, I generalize my account as to how some virtues may be in tension with each other. This generalized account goes beyond the particular phenomenon of pain sensitivity and empathy, appealing to personal differences in sensitivity to affective or emotional information. Additionally, I'll argue that some virtues benefit from or involve having heightened emotional sensitivity while other virtues benefit from or involve having decreased sensitivity. Given one's general degree of sensitivity to emotional states, one will likely be able to have some virtues and not others. Thus, virtues will be in tension with each other insofar as they involve or benefit from a certain amount of emotional sensitivity. Importantly, the tension between some virtues resides, at root, within the affective or emotional components of virtues. However, I'll argue that this will likely spill over into other components of virtues – such as epistemic capacities of figuring out what to do and volitional or motivational capacities of actually carrying out the right action.

3.4.A: Bayesian Predictive Processing and emotional sensitivities

Bayesian predictive processing has gained much traction in recent years, as it is thought to underly perception, decision making, and learning, including within the moral realm. According to this framework, we continually make predictions about what incoming information we are receiving, given hypotheses about what is most probable. Within perception, for instance, our visual system makes use of these hypotheses, or

priors, to infer what is the most likely percept, given fuzzy or uncertain incoming sensory data. The more uncertain or ambiguous the incoming sensory evidence is, the more that priors are relied upon. On the flip side, the more precise or less fuzzy the incoming sensory evidence is, the less that priors need to be invoked. Consider, for instance, the case of Phantom Phone Vibrations that predictive processing theorist, Andy Clarke (2019), gives:

> I imagine that most people have experienced phantom phone vibrations, where you suddenly feel your phone is vibrating in your pocket. It turns out that it may not even be in your pocket. Even if it is in your pocket, maybe it's not vibrating. If you constantly carry the phone, and perhaps you're in a slightly anxious state, a heightened interoceptive state, then ordinary bodily noise can be interpreted as signifying the presence of a ringing phone.

Because your phone is usually in your pocket and often vibrates, your brain learns this information and stores these probabilities, so that when you experience these interoceptive sensations, you tend to interpret them as phone vibrations (and usually you are correct!). But, even in cases when you are wrong, the incoming sensory information is made sense of by invoking generally, quite reliable, statistically optimal, priors. The more vague the incoming data – as in the phantom phone vibration case – the more heavily that priors are weighted or relied upon.

The first main point relevant for my argument is that there are two basic features at work: (1) incoming sensory evidence and (2) priors, or stored statistical information. The more precise the incoming sensory evidence is, the more weight that such evidence is given and the less weight that is given to priors. And, the less precise the incoming

sensory evidence is, the more weight is given to priors. The second main point to make

note of is that the relative weights of these two factors are in a direct tradeoff with each

other – the more weight that is given to (1), the less to (2), and vice versa (see Figure 3.2).
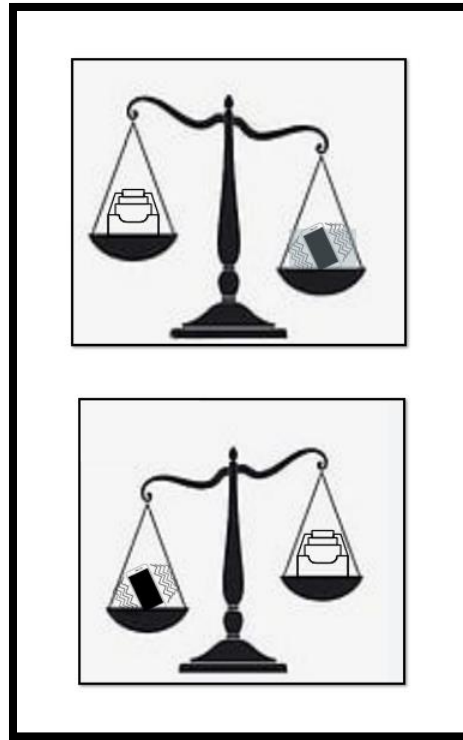
Figure 3.2. *Figure 3.2 depicts tradeoff between the weight assigned to priors (symbolized by the file drawer) vs. precision of incoming sensory data (symbolized by the vibrating phone).*

While one way that the relative weights are distributed is due to the precision of

the incoming environmental evidence, there is also neurological variation among

individuals. According to the Intense World Hypothesis for Autism (Markram &

Markram, 2010), those with autism, for instance, are thought to generally give less weight

to their priors or have 'hypo-priors' (van Boxtel & Lu, 2013). Incoming sensory evidence

is experienced as 'hyperqualia' and the world presents itself as 'too real' (Pellicano &

Burr, 2012), as is often corroborated by those with autism reporting a hypersensitivity to light, noise, and other sensory stimuli.

While those with autism place more weight on the precision of incoming sensory data, research indicates that there is a variation even within the neurotypical population. Ward (2019) explains that "simple sensory stimuli (e.g., noises, patterns) may reliably evoke intense and aversive reactions. This is common in certain clinical groups (e.g., autism) and varies greatly in the *neurotypical* population" (italics mine, p. 139; see also Favre et al., 2019). Thus, people's individual psychologies differ with respect to the general, default weight given to priors vs. the precision of incoming sensory data. Some people will be more sensitive to incoming sensory stimuli, some will be less.

Furthermore, incoming sensory information doesn't just include sounds or visual stimuli, but also interoceptive or somatosensory stimuli (Seth & Critchley, 2013). Given that somatosensory information is generally thought, at the very least, to be a component of emotion (Feldman Barret, 2016, 2017, James, 1884, Prinz, 2004, 2006), people will also vary in how sensitive they are to incoming somatosensory stimuli, and so vary in how strongly they feel their affective or emotional states. Even outside of invoking the Bayesian predictive processing framework, other research has also found similar personal variations of emotional sensitivity:

> This individual difference dimension [of affective intensity] is defined at one pole by persons who experience their emotions only mildly…and at the other pole by persons who experience their emotions quite strongly…Given the same level of

emotional stimulation, individuals high on the affect intensity dimension will exhibit stronger emotional responses, regardless of the specific emotion evoked" (Larsen and Diener, 1987, p. 1-2).

So, it seems like there is good empirical evidence to think that people vary with respect to their emotional sensitivities – some folks generally feel their emotional states more strongly, and others more weakly.[51]

3.3.B: Variation of emotional intensity for different virtues

In addition to the fact that people vary in their emotional sensitivities, there is a second important claim for my argument – namely that different virtues may involve or benefit from being more or less emotionally sensitive. At one end of the spectrum, there are virtues like perseverance, courage, or patience, which involve enduring difficult circumstances and uncomfortable conditions. One might be able to exhibit more patience when one is not overwhelmed by emotions and affective states of uncomfortability or distress. In the same vein, Lisa Tessman (2005) also notes that "[in] aiming at the mean of courage, they might instead develop…an inability to feel any emotions" (p. 126). While an inability to feel *any* emotions is likely a sign that one has overshot the mean and so

---

[51] Even if it is true that people vary in how intense they generally feel their emotions, reader may be concerned that these differences are minor, failing to contribute in any substantial or noticeable way to our emotional and moral lives. But these studies actually report decent effect sizes: Larsen and colleagues (1986) looked at being high on the Intensity Affective Measure (AIM) vs. low predicted how one would respond to various (objectively rated positive and negative) emotional stimuli and events. The effect sizes proved to be medium: "The Pearson correlation between the AIM and the averaged bad event rating was .32 (p < .01). The correlation between the AIM and the averaged good-event rating was .43 (p < .001)" (p. 805). And, they found that "this effect held no matter how objectively bad or good those events were" (p. 808), or across slightly, moderately or very bad/good stimuli.

hasn't hit virtue, this observation nonetheless indicates that to hit the mean when it comes

to courage involves *some* amount of emotional insensitivity, or a *reduced* ability to feel

emotions. From a more empirical approach, there is some evidence that having increased

self-control – a capacity which is plausibly part of virtues like perseverance, courage, and

patience – is related to reductions in how strongly one feels their emotions: "Individuals

high in self-control showed fewer emotional fluctuations and less intensity. This suggests

that emotional experience may be limited for those high in self-control." (Layton &

Muraven, 2014, p. 48).

We can contrast these virtues with another set – those that involve, or benefit from,

increased emotional sensitivity, or felt intensity. Robert Roberts (1989) carves up the

virtues into a few different groupings – including *virtues of willpower*, which include

virtues like perseverance, courage, and patience, which are virtues which "are powers for

managing emotions" (p. 294). In a similar vein, there are also *detachment virtues*, which

"require an *absence* of a certain range of emotions" (ibid) and include virtues like humility

and generosity.[52] But in contrast with this, Roberts also groups certain virtues into

passional virtues and *emotion-virtues*. Passional virtues include virtues like justice or

benevolence – "a person with the virtue of justice is a person with a passion towards a

just state of affairs" (ibid). Emotion-virtues are the "dispositions to have the emotion

---

[52] Roberts thinks that "generous person does not feel much regret at being detached from his goods" (p. 294), and so exhibits an absence of emotions in this way.

from which they get their name" (ibid), such as compassion and gratitude. Compassion, for instance, is typically thought to involve *feeling pained* at another's undeserved misfortune or suffering. And so, being able to be emotionally sensitive or feel the said pain seems to be an important part of the virtue of compassion.

Another way of motivating this point can be done by looking at two different ethical frameworks or approaches to virtue and emotion: On one end of the spectrum, we have Stoicism and on the other end, Sentimentalism. While the Stoics generally see emotions as being at odds with virtue and embrace virtues of self-reliance and self-management, Sentimentalism roots moral virtue in emotions, especially emotions of fellow-feeling or empathy, prizing virtues like caring, kindness, sympathy, and compassion (see Slote, 2017). The contrast between these two frameworks nicely illustrates the association between certain virtues (self-management and self-reliance virtues, such as self-control; altruistic virtues of compassion, sympathy, kindness and caring) and the importance of being moved by one's emotions.

While I have suggested that some virtues involve or benefit from general reduced emotional intensity, and others from general heighted emotional intensity, this is not to say that one couldn't also have excessively reduced or excessively heighted emotional intensity, indicative of vice. Just as being completely emotionally numb is not the appropriate emotional sensitivity for courage (as discussed above), it likewise is the case experiencing one's emotions *too strongly* may push one beyond the virtuous mean of

compassion or kindness and into the realm of vice. Aristotle, for instance, explains that

being pained too strongly at another's loss does not necessarily result in compassion, but

instead in horror, and "tends to drive compassion out" (*Rhetoric*, 1386a20–24). Empirical

work on compassion and felt distress at the suffering of another likewise observe this

relationship: Being *too* distressed at other's plight actually results in disengagement and

can undermine compassionate responses and actions (Batson, 2011). So, just because

certain virtues – like those of say, compassion or justice – plausibly involve or benefit

from

having

greater

a
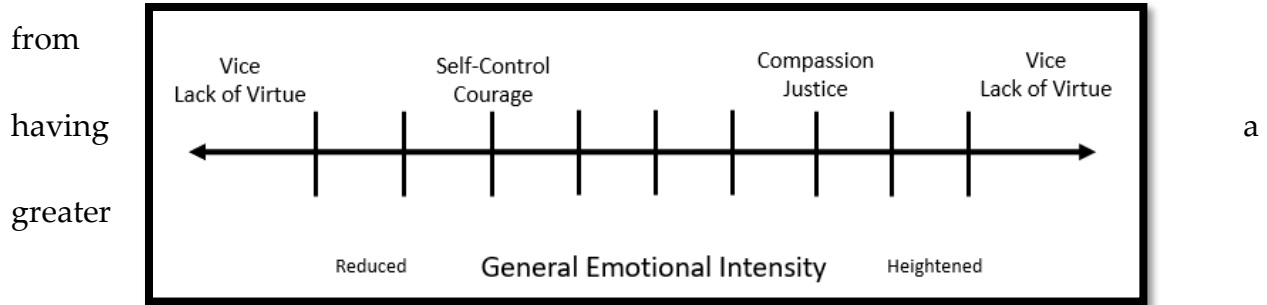
Figure 11. *Figure 11 depicts where certain virtues lie along the spectrum of general emotional intensity.*

heightened emotional intensity, this does not mean that one couldn't generally feel *so*

strongly that their emotions lead them astray (Figure 11).

Nonetheless, while being on one end or another of the spectrum with respect felt

emotional intensity may be constituent of vice, there is still room for a large amount of

variation of how strongly one generally feels one's emotions which might be involved in

having some virtues, but not others. Given that people seem to vary with respect to

strongly they feel incoming affective information, or emotional states, and that different

virtues plausibly benefit from or involve generally feeling emotions more (or less)

strongly, an empirical Tension of the Virtues Thesis begins to take shape. If a person generally is less sensitive to their emotions or feels them less strongly, they will likely be able to have the virtues of self-control and courage but will greatly struggle to have other virtues like compassion or justice. And, of course, vice versa. Thus, at least some virtues – like self-control or courage – are likely in tension with other virtues – such as compassion or justice – in virtue of these empirical features about our human psychology, namely differences in people's sensory processing systems.

3.4.B: The role of the emotion in virtue

Thus far, I have argued that some virtues are in tension with each other, due to differences between virtues when it comes to emotional intensity – some virtues will generally involve or benefit from reduced emotional intensity while other virtues will generally involve or benefit from heighted emotional intensity. It should not go unnoticed, though, that the tension lies specifically within the emotional components of virtues. In this section, I will flesh out a bit more for what this tension among the emotional components means for the tension of these virtues, as a whole. I will argue that while the tension is at root, within the emotional component or domain, emotions importantly influence other capacities or components of virtues – including epistemic and motivational components. Thus, there will be probable, downstream tensions that arise between other components of these virtues.

Aristotle explains that "[Emotions], and in general pleasure and pain can be experienced *too much or too little*, and in both ways not well" (1106b18). Experiencing the right emotions, to the right degree, is "the business of virtue" (1106b24) So, insofar as one fails to experience the appropriate emotional intensity called for, one lacks an important aspect of virtue. But I think we should further wonder whether other aspects of virtues will be negatively affected. For instance, if one does not have the appropriate emotional intensity that is called for by the situation, one will likely be less motivated to carry out the appropriate action. Generally speaking, emotions motivate action. Some (Frijda, 1986; Deonna and Teroni, 2015, Scarantino 2014, 2015) hold this motivational component to be the fundamental feature of emotion – emotions just are motivations to act, feelings of action readiness, or preparing the body for action. But even for those who don't hold that every emotional episode necessarily motivates particular actions, emotions are *generally* thought to be an important aspect of our motivations. We should expect that differences in emotional states and their intensity will, in general, impact differences in the actions carried out. Support for this point can actually be found in the above discussion of compassion, and how feeling *too* much distress or pain in another's plight often results in withdrawal and avoidance. And, experiencing not enough will likely keep one from being moved to act compassionately at all. So, the right emotion, of the right intensity, generally impacts one's motivations to act, and so one's actions.

Aside from impacting our motivations, we should also expect that failing to have the appropriate emotion to the appropriate degree or intensity will negatively impact our epistemic capacities. Tappolet (2016), for instance, holds that emotions are evaluative perceptions, or a way of perceiving value. And furthermore, just as visual perception immediately justifies (sans defeater) beliefs about the object being perceived, so too, emotions can immediately justify (sans defeater) our evaluative beliefs. But even those who are critical of this perceptual account of emotions nonetheless still point to an important epistemic role of emotions. Michael Brady (2013) suggests that emotions guide and capture our attention: "[emotions are] of central importance for making the value of objects and events salient for us, such that without emotion many important or significant objects would pass us by" (p. 23-24). And, by way of guiding and capturing our attention, emotions then also "motivate us to search for reasons or evidence that bear on the accuracy of our initial emotional responses" (p. 129). In searching for the underlying reasons which (in reality, do or do not) justify our emotions, Brady argues emotions play an important role in leading us to evaluative understanding: "emotions promote our understanding of our evaluative situation, since to discover the reasons why some object or event has some evaluative property just is to come to understand one's evaluative situation with respect to that object or event" (p. 147).

All this to say – whether emotions directly justify beliefs, promote evaluative understanding, or simply just guide our attention, they seem to play an important

epistemic role in our evaluative and moral lives. So, we should expect that an inaccuracy of emotional intensity will negatively impact these epistemic capacities. For instance, if one fails to have the correct degree or intensity of the fear in a particular situation, one's attention may fail to be sufficiently captured, one may arrive at an inaccurate belief, and/or one may fail to reach evaluative understanding for what is actually the dangerous-making features of the particular situation. Insofar as these epistemic goods are negatively impacted, we should likewise expect negative effects on the actions we carry out: arriving at a false belief about the degree to which something is dangerous or not will likely impact how we act or deal with this danger; failing to give sufficient attention to the dangerous stimuli may lead us to take it as less important than it actually is, impacting how we react to this stimuli; failing to understand why the stimuli is sufficiently dangerous may further impact how we navigate future similar situations, as we will have failed to grasp the underlying danger-making features, etc.

So, even if the heart of the tension of virtues is within the emotional components of these virtues, we should expect that this tension will not simply remain within this domain, but spill over into other aspects of virtues, such as motivational and epistemic components. Insofar as these components are impacted, we should also expect our general tendency to carry out virtuous action to be impacted. Given how many components of virtue are likely to be affected, we have reason to think that this tension will expand beyond merely the emotional components of the virtues, and truly be one

where the whole of one virtue (courage) is in tension with the whole of another (compassion).

3.5: Modifying the ideal of the virtuous person

Thus far, I have argued that certain virtues are in empirical tension with each other, given the tension between differing levels of emotional intensity typical of their emotional components. One obvious implication of this is that we, as humans, with the psychological and affective hardware that we have, may never be able to have all of the virtues, in all of their (including affective) components. This means that we will be unable to reach a state where the unity of virtues holds true for us. But this alone may not cause much hullabaloo, for one may (correctly) think virtue ethics was never committed to any of us humans actually attaining all of the virtues anyways. The virtuous person – who embodies the unity of virtues – is an ideal, and not any one of us (Annas, 2004, p. 67).

However, I don't think the tension of the virtues thesis merely impacts how we should think of humans who are simply en route to full virtue. Rather, if some of the virtues are in fact in tension with each other, this should also prompt us to modify our conception of the ideal virtuous person as well.

Before giving my argument for why this so, I want to first clarify the notion of the ideal virtuous person that I am working with: First, I take it that the ideal virtuous person is thought to be completely virtuous – the ideal virtuous person is presumed to have all of the virtues, in all of their components, including the required emotional sensitivities.

This means that the ideal virtuous person is one who would feel the right emotion given their circumstances, to the right degree or intensity.

Secondly, there are various senses of an ideal and the way that this ideal enters into the way we think about and pursue virtue. For instance, an ideal could be understood as something that we appreciate as excellent or deem as good; call this the *exemplary* sense. Another way of illustrating the *exemplary* sense of the ideal virtuous person is that, if it is true that some of the virtues are tension with each other and humans can never be fully virtuous, this would be unfortunate. There is something to mourn about the fact that we can never embody the ideal that we see as excellent. This exemplary sense of the ideal virtuous person is *not* one that I necessarily think we ought to modify, even if The Tension of the Virtues Thesis is true. Rather, there is another sense that I am concerned with: the *emulatory* sense – where the ideal virtuous person provides us with a goalpost to strive towards, and with a psychology that we strive to mold our own to be more like. This ideal of the virtuous person plays an important role as we try to approximate it and make our own characters more similar to it. Insofar as the tension of the virtues is true, it is this sense – the *emulatory* sense – of the ideal that I suggest we ought to modify.[53]

---

[53] Further empirical research needs to be done to consider whether one can psychologically hold on to the exemplary sense of the ideal virtuous person, where this ideal has all of the virtues, while simultaneously modifying the emulatory sense. It might be the case, psychologically speaking, that an ideal which we regard as excellent will automatically prompt us to also strive after or emulate it. Research on the emotion of admiration (see Onu et al., 2016, for a nice review) suggests that who we admire, or see as excellent, does often result in attempts to emulate. If this is true, then there might be an indirect reason to also do away

To show why we should modify the emulatory sense of the ideal virtuous person, I want to begin by making use of an *illusive ideal*, a notion which I get from Siversten (2019). In short, an illusive ideal is one where progress within one facet of the ideal undermines progress made in another facet. For example, Siversten argues that the informed, impartial spectator is an illusive ideal: the more informed one is, the less impartial one is, and vice versa. For a less academic example, take the ideal of an all-around excellent runner – one who is both an excellent endurance runner and also an excellent sprinter. It turns out that given the way our human muscular system works, we cannot be both, and that pursuit in one area undermines progress made in another. When one engages in sprinting-related training, her intermediate muscle fibers turn into fast twitch fibers, and so are necessarily not slow twitch. If this runner were to then aim at being an excellent endurance runner and so undergo endurance-related training, these intermediate muscle fibers would then become slow switch, and necessarily be no longer fast twitch. The training that helps her progress within the realm of endurance running undermines progress within the realm of sprinting, and vice versa. Thus, the ideal of an all-around excellent runner – at least insofar as it is an *exemplary* ideal – is an illusive one, for when we try to approximate it, our efforts are self-undermining.[54]

---

with the exemplary sense of the ideal virtuous person: insofar as holding on to this sense will interfere with the ideal that we seek to strive after (e.g. the emulatory sense), we might do better to change the ideal virtuous person – both in the exemplary and emulatory sense – to be one that does not necessarily have all of the virtues. I am grateful to Ethan Kross for bringing this potential problem to my attention.

[54] It should be noted that not *all* efforts in either case – both in striving to be an all-around excellent runner, or a virtuous person – are necessarily self-undermining. We can imagine the major couch potato who

Insofar as The Tension of the Virtues Thesis is true, trying to mold our psychologies to be like the ideal virtuous person – one who has all of the virtues, in all of their components – is on par with trying to undergo the physical training necessary to be an all-around excellent runner. Our efforts in both realms will often be self-undermining. Suppose that one slowly changes her emotional sensitivities such that she experiences general reductions in affective intensity, which enables her to better manifest the virtues of courage and self-control. But, if it's true that other virtues – like compassion or justice – involve a general increased affective intensity – then her efforts, if successful, will undermine the having of these other virtues. Obviously, it would be great if one could experience the reduced affective intensity necessary for courage when courage is called for, and the heighted affective intensity necessary for compassion when compassion is called for. And, likewise, it would be great if one's intermediate muscle fibers could be trained to be both fast and slow twitch, whichever being relevantly activated at a moment's notice, given the race she is partaking in. But that's just not how the human muscular system – and plausibly, not the human mind – works.

---

decides to get in shape by sprinting down his street every day. Given his current state, this minimal amount of sprinting will generally benefit his cardiovascular fitness, moving him closer to both being a better endurance runner and sprinter, and so an all-around excellent runner. So too, we might imagine the moral couch potato who one day decides to adopt the general moral motivation to do the right thing. This will likely move him closer to both being both courageous and compassionate, and so an all-around virtuous person. Nonetheless, in cases where is one beyond a mere (moral or physical) couch potato and is engaging in training that enhances one particular facet (the emotional sensitivities for a set of virtues; the muscle fibers for one kind of running), the ideal of the overall excellent runner and the ideal of the virtuous person becomes an illusive one.

Thus, if we do not modify the ideal of the virtuous person and continue to strive to inculcate all of the virtues, with all of their components, we will engage in a self-undermining pursuit, making our moral lives into frustrating and counterproductive ones. One modification, then, is that we should not try to pursue all of the virtues, in all of their components – in particular, having the correct emotional sensitivities that are called for. But this modification should not only be negative in nature, but rather should make room for something else to take its place: a positive modification whereby we include a social component – that of relying on another's moral testimony – within the ideal virtuous person.

Recently, several philosophers have claimed that relying on another person's moral testimony is at odds with being virtuous: it is a "moral defect of character" (Crisp, 2014, p. 132) and "a way of falling short" (Hills, 2009, p. 113). The virtuous person shouldn't need to rely on another person's moral say-so, because a virtuous person should be able to figure out what to do for herself. A virtuous person is her own moral compass (ibid, p. 112), having the ability to "see what is true in every set of circumstances, being like a carpenter's rule or measure" (Aristotle, *NE*, 1113a32–1113a34).

But I think this ideal of the virtuous person – at least insofar as it is an exemplary one that guides our own pursuits - is just an illusive one. This is because, given the empirical evidence, it is dubious that we humans could have the varying affective sensitivities necessary to figure out for ourselves what the right thing to do is, across a

wide variety of situations that call for a wide variety of virtues. Rather, one might do

well with respect to concerns of courage; but then not so well with respect to concerns of

compassion. And so to hold the exemplary ideal of the virtuous person as one who is her

own moral compass, across a variety of situations that call for a variety of virtues and

varying emotional sensitivities, is one that will likely prove to be self-undermining, futile,

and frustrating when we try to pursue it.

To illustrate how adding this social component of relying on another's moral

testimony would fit into our modified notion of the exemplary ideal virtuous person,

consider the following case of Jane, which I get from Claire Field (2022):

> JANE: Jane struggles with the ethics of asserting hurtful truths. She knows that
> some assertions can be upsetting, but she struggles to identify which these are.
> Explanations of why some personal truths can be upsetting strike her confusing
> and a little far-fetched. While she tries to be charitable, she struggles to believe that
> it could really be morally important to avoid such assertions, particularly when
> this comes at the expense of saying things that are relevant and true. Nevertheless,
> Jane wants to do the right thing, whatever that is. Aware of the ways in which she
> struggles with interpersonal interactions…when she notices that there is a risk of
> hurting others' feelings in a way that could be wrong, Jane asks a friend for advice,
> and does whatever she says is the right thing to do. (p. 2710-11)

We can further imagine that Jane struggles to have the necessary affective

sensitivities for virtues like kindness or compassion precisely because she *does* have the

necessary affective sensitivities for virtues like self-control and courage. In this case, then,

what is someone like Jane to do? Just what she in fact does – deferring to her trustworthy

friend's moral testimony. Note that, plausibly, if her friend has the necessary emotional

sensitivities to get things right in this scenario, her friend might likewise struggle with situations where courage or self-control are called for, and in those cases, Jane's friend would do well to defer to someone like Jane, who has the necessary affective sensitivities to navigate those situations well.

When we look to the exemplary ideal virtuous person, then, what we should strive for is not to cultivate all of the virtues in all of their components, including the necessary emotional or affective sensitivities. Rather, we should try to be aware where we lie on the scale of general emotional intensity and what situations – given the virtues they call for – we will be able to navigate for ourselves and what situations we will need to rely on others.

3.6: Lingering questions and concluding thoughts

Any account that makes use of moral testimony is going to face *The Credential Problem*: "The expert's expertise might best be judged by the moral advice she provides, but a non-expert is in no position to appraise the content of that advice" (Cholbi, 2007, p. 325). Consider a non-moral, analogous, example: You do not know how to solve a difficult math problem, but need the answer, so you seek out and try to identify an expert mathematician to give you the solution. However, to keep this case as analogous as possible to the moral one, in this world, there is no formal mechanism of granting mathematical credentials – no such thing as PhD in math, for instance – that you could use to guide your search. Instead, you are to identify the math expert by finding someone

who reliably arrives at accurate solutions to tough math problems. But there is no way you could know what are in fact the accurate solutions – this is the precise skill you lack – which is why you are seeking a math expert to begin with! So, the very ability that you need to correctly identify the expert is also the precise ability you lack, and the reason why you seek out an expert's testimony. Analogously, then, how – as I've suggested – will a person who lacks a given virtue or expertise also have the skill or ability to successfully identify the person with that virtue or expertise, such that they can rely on their moral testimony?

While this problem may be one that plagues moral testimony more generally, I don't think it is as great of a threat to my proposal. The cases that I'm concerned with are ones where a person has one (set of) virtue(s) but lacks another (set). What is needed to identify a reliable moral testifier in these cases is to identify someone who has the contrary set of virtues that you lack: Jane – who has general reduced affective sensitivity – is strong in virtues like courage and self-control, but weak in virtues like compassion and kindness. In order to identify a reliable moral testifier in her case, Jane needs to be able to pick out someone who has generally heighted affective sensitivity and is strong in virtues like compassion and kindness. Unlike general cases of moral testimony, where the non-virtuous tries to identify the virtuous by way of discerning whether the testifier's particular testimony is correct or not, the cases I'm concerned with involve identifying a person who has set of general character traits that you also lack. Furthermore, given the

tension between these sets of virtues, we might expect that the person who has the virtues you are looking for *also* lacks, or is weak in, the virtues that you have or are strong in. This means that there will be some ways of narrowing in on the necessary testifier which appeal to *your* moral expertise, rather than simply your lack of it. So, we need not identify the reliable testifier by the accuracy of their particular testimony, but rather by their general traits. We seem to do this effectively lots of the time: Mary knows she is headstrong and decisive. Her partner, Aiden, knows he is flexible and cautious. When it comes to low-risk decisions – like deciding where to order dinner from - Aiden defers to Mary (or else they'd never eat!) When it comes to high-risk decisions - like buying a house - Mary defers to Aiden (or else they'd end up in a house with a leaky roof and crumbling foundation!). As couples often say: "we balance each other out." We often operate day to day by deferring to those who have character strengths which we lack in particular circumstances, and vice versa. Aside from our own everyday experiences to confirm this, there is good empirical evidence to think that we aren't too shabby at identifying other people's traits in general (see Westra, 2019, for discussion of this literature).

Thus, I think there is reason to treat the identification of someone who has a particular (subset of) virtue(s) by one, who herself, has a (subset of) virtue(s) to be different than identifying a general moral expert by one who lacks general moral expertise. Nonetheless, I don't think this process is infallible, and surely it can go wrong.

Future research surely needs to look at how variation in general affective intensity, along with having certain virtues and lacking others, might impact our ability to identify those across the other side of the aisle when it comes those having opposing affective intensities and opposing virtues.

Nonetheless, I am hopeful that including the social component of relying on another person's moral testimony is a beneficial modification we can make, as we start to rethink our notion of the ideal virtuous person, at least insofar as this ideal is an exemplary one, or one that guides our pursuit and cultivation of virtue. In any case, I think The Tension of the Virtues is likely an empirical tension that we need to reckon with, both in further investigating all of the ways that certain virtues may be intension with each other and all of the possible empirical solutions that may be available to us.

Chapter 4: Understanding What Matters

4.1: Introduction

Literature on moral understanding often begins with a case that illustrates a

puzzle. The puzzle usually concerns the asymmetry between accepting moral and

nonmoral testimony. But consider a different case and another puzzle:

> FOOTBRIDGE: Jake believes that it is wrong to push the large man off the bridge
> to stop an oncoming trolley in order to save the five people. His reason for why it
> is wrong is that doing so would be using the large man as merely a means to an
> end, intentionally killing (or severely harming) him to save the lives of others. Like
> Jake, Judith also believes that it is wrong to push the large man off of the bridge.
> And, like Jake, Judith also believes that it would be wrong to do so for the same
> reason that Jake takes it to be wrong (e.g. using the large man as a mere means to
> an end). However, additionally, Judith also grasps the reasons that speak in favor
> of pushing the large man off of the bridge – namely the utilitarian reason of saving
> the greatest number of lives.

Moral understanding why p has been typically defined as grasping the

explanation, q, for p, where p is some proposition that Φ-ing is morally right (or wrong).

The explanation for p consists of the reasons that make Φ-ing right or (wrong). But, in

FOOTBRIDGE, both Jake and Judith have moral understanding according to this

definition; both Jake and Judith understand why p, in virtue of q (it is wrong to push the

man off of the bridge because it is wrong to use a human person as merely a means).[55]

However, it is also intuitive that Judith has *more* understanding, or understanding to a greater degree, than Jake. Contrary to knowledge, understanding is thought to come in degrees – one can have more or less of it. This paper gives one way in which Judith has a greater degree of moral understanding. In this paper, I put forth an account of moral understanding which includes an *understanding of what matters.* Understanding what matters involves grasping the reasons that not only explain why *p* is true, but also explain what would make *p* false; understanding what matters involves grasping the reasons for and against.

This paper proceeds as follows: In *Section 4.2* I consider three ways that moral understanding is thought to be valuable and argue that in order to secure these three values, one often needs to have an understanding of what matters, too. In *Section 4.3,* I consider one further value that has been attributed to moral understanding. This value concerns carrying out morally worthy actions – if one's action is to be of moral worth, one must act for the reasons that make the action right, and so one must also grasp those reasons as they support the rightness of the action in question. I use this to point out that current accounts of moral understanding might be mistakenly equating the reasons which, when acted on, make an action morally worthy, with the totality of reasons that are grasped when one has moral understanding. I use an empirical model of intentional

---

[55] I assume that it is in fact wrong to push the man off of the bridge in this case.

action to show that we often act on a *subset* of the reasons that we grasp or deliberate over. Thus, moral understanding plausibly involves *more* reasons than merely those reasons that are acted on and make an action a morally worthy one. The grasping of these further reasons is to have an understanding of what matters. Lastly, in *Section 4.4*, I examine the relationship between virtue and a having deep degree of moral understanding, which includes understanding what matters. I suggest that virtue might sometimes require one to *not* understand what all matters, and so to have some degree of deficiency in their moral understanding.

4.2: Three values of moral understanding – Why moral understanding involves an understanding of what matters

Recent literature on moral understanding often highlights the value of moral understanding and what is lost when one defers to another's moral testimony. What makes moral understanding valuable is often connected to the nature of moral understanding: if moral understanding is valuable because it is required for us to engage in justifying our actions to others, for instance, then moral understanding must consist of something that involves abilities to explain and give reasons. In this section, I begin with three ways that moral understanding is often thought to be valuable and argue that if such values are to be secured, then this reveals something about the nature of moral understanding – and understanding what matters must also typically be involved. In other words, if one is inclined to think that moral understanding is valuable for the

114

reasons outlined below, then they should also be inclined to incorporate understanding

what matters into their conception of moral understanding.

First, moral understanding has been said to be intrinsically valuable, for when one

has moral understanding, one's cognitive states mirror moral reality. True beliefs are

thought to be valuable because their contents mirror or reflect truths about the world.

But, if such mirroring is valuable when it comes to having true beliefs, then

understanding why p "must be valuable twice over" (Hills, 2015, p. 679), for through such

understanding "you can mirror the *structure* of the world within the *structure* of your own

thoughts as well as their content" (ibid).   Stephen Grimm (2012) likewise explains that

"the mind of someone who understands mirrors or reflects reality at a deeper level than

the mind of someone who merely propositionally knows" (p. 109). When one has moral

understanding, it's not only that your beliefs or thoughts mirror what is true, but the way

these mental states are structured or organized – such as being in justifying or supporting

relationships – also reflect something about the structure of moral reality.

If it is right that mirroring the structure of the moral world is valuable in this way,

then moral understanding that mirrors the full span of moral considerations or values –

for and against – must be *even more* valuable than a kind of moral understanding which

only reflects the subset of reasons that end up explaining the action in question. This is

perhaps why we think that Judith has deeper moral understanding than Jake: her moral

understanding is *more* reflective of moral reality, appreciating what is all at stake.

Another way of supporting this point is to consider what several philosophers say in response to the McDowellian silencing thesis. According to the silencing account, the virtuous person is thought to ignore reasons that speak against doing what is morally right. Objections made against the silencing thesis often point to the fact that "virtue can have a cost, and a mark of the wise person is that she recognizes it." (Baxley 2007, p. 419; see also Stark, 2001 and Stohr, 2003, for similar observations). There is something *valuable* about appreciating what's at stake, even when it means attending to and grasping the reasons that run against doing the right thing. To fail to grasp such reasons is to fail to appreciate what is of value. If mirroring is valuable because one's thoughts (and their structure) reflect moral reality, then one's thoughts should also mirror all that is at stake, or all that morally matters. And this will at least sometimes require one to not only grasp reasons that explain why *p*, but also reasons that speak against the truth of *p*.[56]

In addition to mirroring, two further values of moral understanding are instrumental in nature: 1) it helps guide us to right action and 2) it enables us to justify our actions to others.

When it comes to guiding us to the right action, moral understanding is thought to play an important role because when one grasps why a particular action is wrong, one

---

[56] I will come back to this point in *Section IV*, for I think there are cases where virtue may demand that on fails to appreciate reasons that speak against the right. Interestingly, though, I don't think that this shows that mirroring all of the underlying reasons – for and against – which is constitutive of understanding what matters fails to be valuable. Rather, it is just that sometimes virtue may require you to fail to attain valuable epistemic states.

has a sense of the underlying reasons for what makes it wrong. These reasons are often general enough so that they can be applied to new, non-identical situations.[57] When one has understanding why *p,* one has certain abilities – those of 'cognitive control' – which enable one to manipulate and apply the underlying reasons to new cases (Hills 2016, p. 674). Woodward (2003) describes that the grasping involved in understanding as abilities which enable one to answer the "what-if-things-had-been-different?'' question. Skyrms (1980, p. 11) similarly explains that this grasping enables one to see what would happen if one feature was 'wiggled.' If one understands why lying is generally wrong, for instance, one would be able to 'wiggle' the conditions of the situation to apply their grasp of the reasons that make lying wrong to distinct, but similar cases – such as to cases of deceitful omissions.

But, in order to be able to secure this value of moral understanding, I think one actually needs to have a grasp of all of the moral reasons at play – for and against – not just the ones that actually support the truth of *p* in this particular case. In other words, in order to have the cognitive control necessary for wiggling the variables, such that one can answer the "what-if-things-had-been-different" question, one needs an understanding of

---

[57] The generalizability of the underlying reasons is seen in the following passages as the reasons are understood to be principle-like: Crisp (2007) explains "Normative principles are like typical natural laws. Each system helps us to understand—in one case, why something happened; in the other, why someone should do something" (p. 47); Wilkenfeld (2020) describes moral understanding in terms "getting at the principles" (p. 30), where these principles explain "why some actions are right and other actions are wrong" (ibid); Strevens (2013) describes moral understanding as "grasp[ing] a correct moral explanation of the rule (perhaps a derivation of the rule from fundamental moral principles) or to have the ability to use the rule to explain moral facts" (p. 515).

what matters. In the FOOTBRIDGE scenario, Judith has something that Jake lacks to at least some degree – cognitive control. If the conditions were wiggled (such that the large man say, gave his consent to be pushed), Judith would likely be better situated to grasp why it would be permissible to push the man in this non-identical, but similar case. If one does not grasp all of the relevant moral reasons, for *and against* the truth of *p*, it is unclear exactly whether one would grasp the reasons which would support *~p'* when *~p'* is true. If Jake fails to grasp that saving 5 people is morally important and is reason that weighs *against* the wrongness of pushing the large man, then it is unclear that he will now come to appreciate this reason as one that *favors* pushing the large man in a new scenario (e.g. when the large man gives his consent). To be able to have the cognitive control that allows us to figure out what to do in new situations, we need to not only grasp the reasons for why *p* but also the reasons that speak against *p*. We need an understanding of what matters.

Lastly, moral understanding has been said to be instrumentally valuable in one other way: "moral understanding is important in part because being in a position to justify yourself to others is morally important…[for a] core ethical practice is the exchange of reasons" (p. Hills, 2009, p. 106-7). If one doesn't grasp the moral reasons for why *p*, then one will be unable to explain or justify why she did the right thing. She will not be able to engage in the exchange of reasons but will instead put forth unsatisfying answers like "*S* told me so" or "I don't know, it just felt right."

Again, if one is inclined to think that being able to give justification for one's moral actions to others is a value of moral understanding, then they should embrace a deeper sense of moral understanding that includes understanding what matters. If moral understanding only involves a grasping of the reasons that make *p* true (but necessarily also the ones that speak against *p*), one will be worse off when it comes to exchanging reasons and giving a moral justification for one's actions. Consider the contrast between Jake and Judith when asked by Peter – who is a staunch utilitarian and advocator for pushing the man to save the 5 – why they think pushing the large man would be wrong. Peter might protest: *"But what about the 5 innocent folks who will die?! Don't they matter?"* Jake will not be able to say much more than the explanation he has already given – perhaps he'd just simply repeat *"like I said, it is wrong to use a human person as a means to an end."* Judith, on the other hand, seems much better situated to engage in an exchange of reasons with Peter. She might point out different situations where she sees it permissible push the man to save the 5 – situations where the man gives his consent or where he is guilty of previously murdering members of society. In doing this, Judith can respond to Peter's question about whether the lives of the 5 matters: Yes, they do matter. And in some cases, saving their lives is enough reason to make it permissible to push the man off of the bridge. But in this case, it's not; here are the differences; here are the different moral considerations that matter.

Judith's explanation, of course, might not be completely satisfying to Peter. Peter might persist, asking why it is only when those certain moral conditions are met will saving 5 be a sufficient reason to push? Why not in the original situation described? When – and why – does using a person as a means outweigh saving lives? To be sure, Judith and Peter may eventually, too, come to a standstill. Nonetheless, Judith still seems much better able justify her actions or decisions to Peter when compared to Jake. And I think it is because Judith understands what matters – she grasps reasons both for and against and can appreciate the reasons that Peter is concerned with.

Thus, if one is sympathetic to thinking that moral understanding is valuable because it 1) mirrors moral reality, 2) helps guide one to right actions, or 3) enables one to engage in an exchange or moral reasons and justify their actions to others, then they should be eager to incorporate understanding what matters into their conception of moral understanding.

4.3: Where current accounts of moral understanding might have gone wrong

In addition to the three values of moral understanding highlighted in the previous section, one further value of moral understanding points to the fact that grasping the reasons which make an action right is that such grasping is thought to be necessary for our actions to be of moral worth. In order for our actions to be of moral worth, we need to act *for the right reasons.* This means that we must also grasp what those reasons are. It might seem intuitive then to think that since grasping the reasons which make an action

right is required for an action to be of moral worth that this also means that moral

understanding consists only in grasping that set of reasons. But thinking this would be

mistaken.

Consider the illustration in Figure 12: If your Φ-ing is of moral worth, then you Φ

because you grasp that r is what makes Φ the right thing to do. I'll grant that it's correct

that the reasons that you act on must be grasped in order for an action to be of moral

worth. But it is not obvious that grasping *those* reasons are sufficient for having moral

understanding. Moral understanding might be *necessary* for grasping the reasons why

make Φ right – which would explain why it is valuable in this sense – but that does not

mean that moral understanding *consists of* only these reasons grasped.
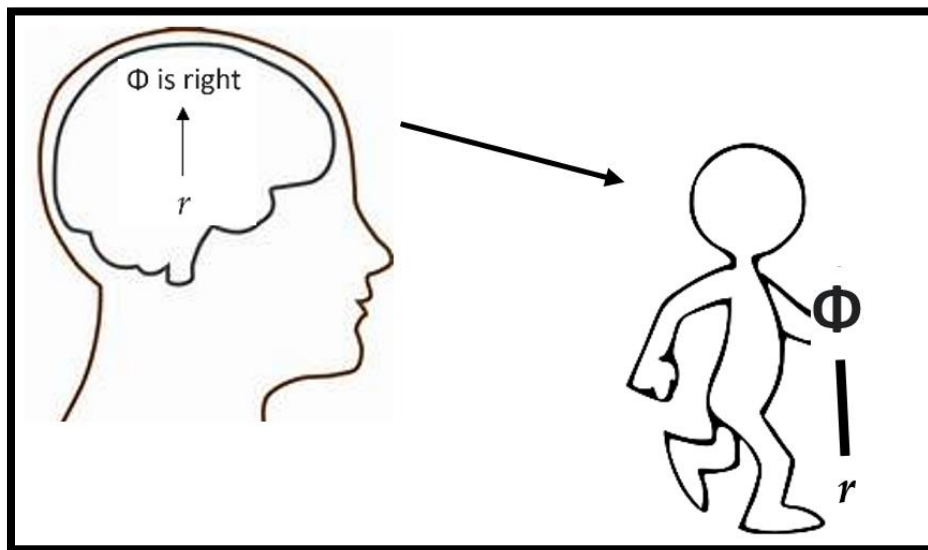
*Figure 12. Figure 12 illustrates the structure of mental contents that must
be had for one's actions to be of moral worth.*

To illustrate, consider the Rubicon model of action – an empirical model which

illustrates how we go about carrying out intentional actions across four phases. It is in

the third phase is when one actually acts on or carries out one's intention (See Figure 13

below).[58] It is here in the third phase where one would carry out a morally worthy action

by Φ-ing for the reasons make Φ right. But notice that in additional to this actional phase,

the first phase – that of deliberation – seems to be where the epistemic stuff happens: one

considers the pros and cons, or reasons that tell in favor and against Φ-ing. Eventually,

one "crosses the Rubicon" and decides what to do, because one sees that certain pros

outweigh other cons. When one crosses the Rubicon and develops an intention to act,

they have selected the reasons which they are about to act on. Figure 12 (above) depicts

the state one after one has finished deliberating and has crossed the Rubicon. But what

is going on in Figure 12 fails to account for the reasons that one considers and weighs

when one is in the deliberation phase. But grasping all of the reasons – for and against

– is an essential epistemic and cognitive step in this process. And, plausibly, doing this

well – having a grasp or understanding of *what matters* – can also be of moral worth. It's

just that the moral worth of understanding is not identical to the moral worth of acting

for the right reasons. The moral worth of the understanding involves a broader scope of

potential reasons, for and against; the moral worth of right action involves a narrower

scope – a zeroing in on the reasons that actually ground the particular action that one is

carrying out (see Figure 14 below). But just because the moral worth of the latter involves

a more restricted set of reasons does not mean that the latter does as well. The reasons

---

[58]Achtziger & Gollwitzer (2018)

that one ends up acting on are usually a subset of the reasons that one deliberates over.

To fail to appreciate the reasons that speak against Φ, even if Φ is the right thing to do,

seems to be to restrict the domain of reasons that a good deliberator, a wise reasoner, or

an insightful understander ought to take into account. In many (even most) cases, moral

understanding involves a grasping of the reasons that explain why Φ is right, but also

would explain why ~Φ is right, if it were to be right. Moral understanding ought to

include understanding what matters, even if what matters doesn't end up making a

difference to what is right when it comes to Φ.
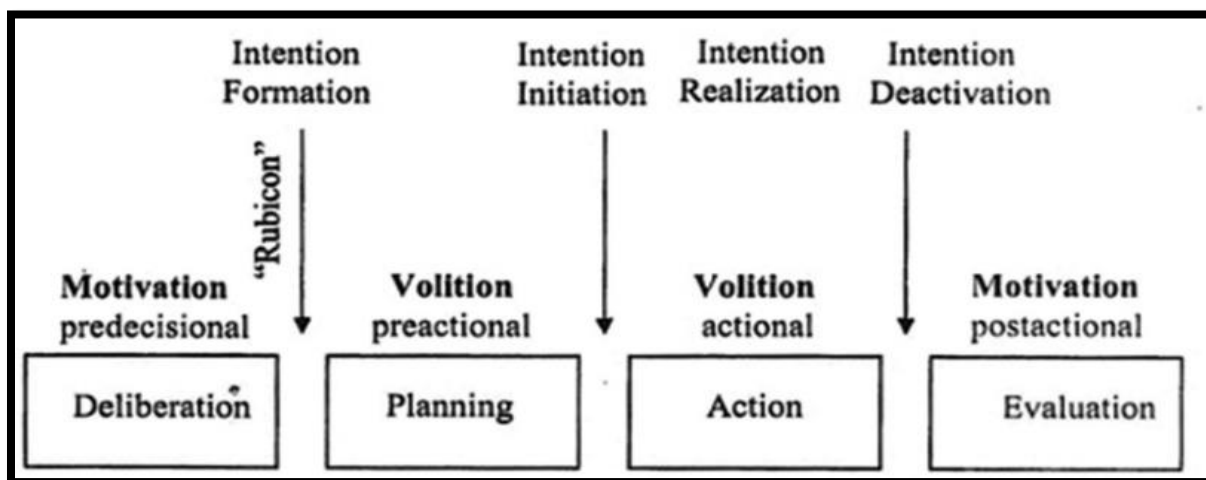
Figure 13. *Taken from Gollwitzer & Achtziger (2018), Figure 13 illustrates the four phases of the Rubicon Model of Action.*
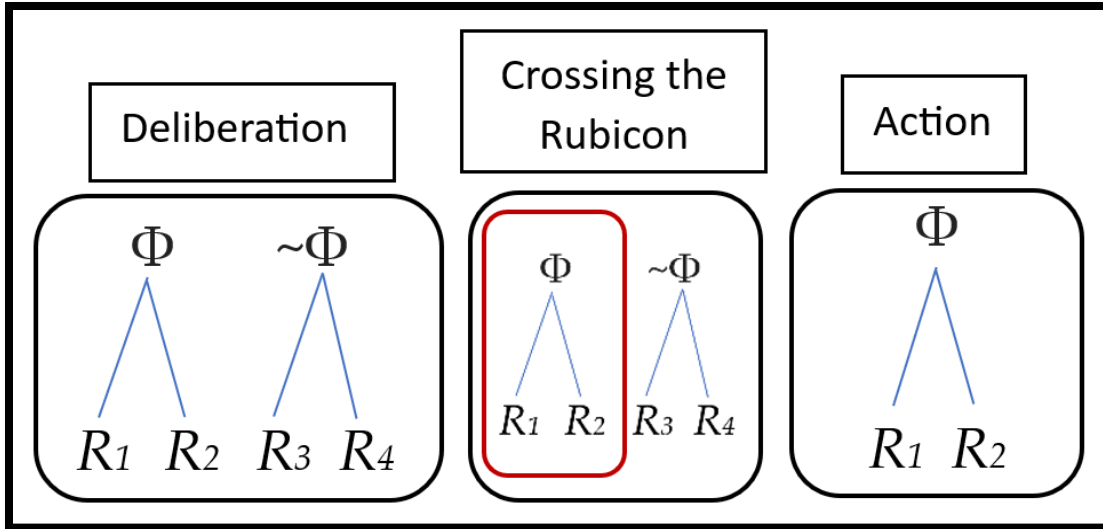
Figure 14. *Figure 14 illustrates how the reasons acted upon during the Action phase are a subset of the reasons that one initially grasps and considers in the Deliberation phase.*

4.4: Understanding what matters and its relationship to virtue

Thus far, I have argued that understanding what matters is an important part of having a greater degree of moral understanding. One needs to not only grasp the reasons that in fact explain *p*, but also appreciate the larger set of morally relevant reasons which tell in favor, and against, the truth of *p*.

As mentioned, moral understanding has been commonly thought to be connected to virtue, both for instrumental and intrinsic reasons. Deficiency in moral understanding will likely impact one's ability to judge and act correctly; without moral understanding, one will be unable to apply moral reasons to new situations (Crisp, 2014, p. 130; Driver, 2006, p. 638; Howell, 2014, p. 403; Hills 2016, p. 151-53).[59] And when it comes to the latter,

---

[59] Although see Wiland's (2021, p. 61) point on this depending on particularism being false.

moral understanding has been said to be intimately tied up with having virtue. To fail to have moral understanding has been said to indicate a "defect of character" (Crisp 2014, p. 132) and is "a way of falling short" (Hills, 2009, p. 113).[60] Howell (2014) explains that being virtuous consists of having an integration around the good, including having an "integration of reasons" (p. 410). This integration means that one's moral reasons for a particular action are not isolated from other beliefs or cognitive states, and that one's cognitive states are not isolated from other aspects of one's character. The way a virtuous person morally moves and exists in the world involves a kind of coherence or integration.

If it is right that virtue requires moral understanding, and that a deeper sense of moral understanding includes understanding what matters, then an interesting question arises – what does this mean for the silencing thesis of virtue? According to McDowell (1978), a virtuous person 'silences' – or does not attend to – reasons that run contrary to the virtuous act. McDowell claims that the virtuous person is simply concerned with "keeping [one's] attention firmly fixed on what Aristotle calls 'the noble'" (p. 27), such that other considerations will not enter into one's practical reasoning. If this is right, then it seems that the virtuous person would be at odds with having deep degree of moral understanding which includes understanding what matters, because the virtuous person

---

[60] Although some have argued that in certain circumstances, virtue or what is ideal might be consistent with a deficiency in moral understanding (see McShane, 2018). While this seems to be the minority position, it is not completely unheard of.

should not attend to or appreciate morally relevant reasons that tell against p (when p is true).

One way around this is to take a more moderate view of silencing: For instance, Schuster (2020) proposes that the silencing of reasons which the virtuous person engages in makes moral reasons *motivationally* inert, but not *evaluatively* inert. The virtuous person ought to not experience motivational conflict about what to do, but will still experience evaluative conflict. The virtuous person will still appreciate that two values or moral considerations pull in opposite directions given one's particular circumstances, even if it doesn't affect one's motivations. Under this reading of silencing, the virtuous person *does* engage in silencing, but the silenced reasons are motivational, not evaluative. This still makes room for the virtous person to appreciate and grasp all of the relevant moral reasons for and against p, or to have a greater degree of moral understanding which includes understanding what matters.

As interesting and attractive as Schuster's interpretation is, I nonetheless still wonder if virtue might sometimes require one to silence even evaluative reasons which speak against doing the virtous action. This would mean that the virtuous person, at least sometimes would not have an understanding of what matters and so being virtuous might require that one has a deficiency in moral understanding.

One place where virtue and having a greater degree of moral understanding might be at odds is when virtue is had within a certain role, relationship, or set of commitments.

126

Such contexts might make it less-than-virtous to recognize and appreciate all morally relevant reasons, or to understand what matters. Consider the case Orthodox priest, Fr. Patrick Reardon who sees the role of the priest ('Father Confessor') as seeing people and their wrongful actions through the lens of mercy: "the Father Confessor is always *on the side of* the sinner… [the Father Confessor] is a man whose habitual mindset is formed in a forum where he functions as the minister of divine mercy…[he] looks at sin and crime through sacramental eyes" (2013). Reardon makes these claims within the context of arguing that priests should never serve on jury duty for they should never be put in the position where they are asked to try to condemn another or find one guilty. Despite there likely being *morally relevant reasons that speak in favor of the defendant's guilt*, Reardon seems to be suggesting that, because of the nature of one's role as a priest, it would not be good or virtuous for a priest to appreciate or grasp these reasons. The virtuous priest evaluatively silences moral reasons that speak against granting mercy. To appreciate or grasp these reasons means that the priest is no longer seeing how to respond to another's wrongdoing "through sacramental eyes."

If Reardon is right, then there is an important implication for moral understanding and virtue: extending our conception of moral understanding to include understanding what matters opens up the door to, at least given one's role, being virtuous may be at odds with having (a deepened sense of) moral understanding.

4.5: Conclusion

127

Debate about the nature of moral understanding has largely presumed that it involves grasping the reasons or ground which make the action under consideration right. But including the reasons that speak against the rightness of an action within one's conception of moral understanding have largely been ignored. This is surprising, given that the value of having moral understanding seems to also pertain to grasping both sets of reasons – those for and against. I have argued that (at least, a deeper degree of) moral understanding ought to include not only grasping the reasons that explain why *p* but also the reasons that speak against the truth of *p.* I have called this kind of understanding *'understanding what matters.'* Interestingly, once we incorporate this into our conception of moral understanding, further implications arise concerning the nature of virtue and moral understanding. I have pointed out a case which illustrates how these two might come apart, but my discussion of this is only cursory and rightfully deserves much more ink.

References

Aars, H. & Dijkersterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology, 78*, 53-63.

Achtziger, A., Bayer, U., and Gollwitzer, B. (2012). Committing to implementation intentions: Attention and memory effects for selected situational cues. *Motivation and Emotion, 36*, 287- 300.

Achtziger, A., & Gollwitzer, B. (2018). Motivation and volition in the course of action. In *Motivation and Action*, 3rd Edition. Eds. J. Heckhausen and H. Heckhausen. New York: Springer, 485-528.

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057-8.

Alderson-Day B, Lima, C F., Evans, S, Krishnan, S., Shanmugalingam, P. & Fernyhough, C. (2017). Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain, 140,* 2475–2489.

Annas, J. (2004). Being virtuous and doing the right thing. *Proceedings and Addresses of the American Philosophical Association*, *78*(2), 61-75.

Aquinas, T. (1920). *The Summa Theologiæ.* Fathers of the English Dominican Province (Trans). Retrieved from https://www.newadvent.org/summa/ on May 5, 2023.

Aristotle. (1984). *The Complete Works of Aristotle: The Revised Oxford Translation.* J. Barnes (Ed.) Princeton, NJ: Princeton University Press.

Audi, R. (2013). *Moral Perception.* Princeton, NJ: Princeton University Press.

Avarguès-Weber, A., Finke, V., Nagy, M., Szabó, T., d'Amaro, D., Dyer, A. G., & Fiser, J. (2020). Different mechanisms underlie implicit visual statistical learning in honeybees and humans. *Proceedings of the National Academy of Sciences, 117*(41), 25923-25934.

Bargh, J. (1990). Auto-Motives: Preconscious determinants of social interaction. In *Handbook of Motivation and Cognition: Foundations of Social Behavior*. Eds. E. T. Higgins and R. M. Sorrentino. New York: Guildford, 93-130.

Bargh, J., Chen, M. & Burrows, L. (1996). Automaticity in social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230-44.

Bargh, J. & Gollitzer, P. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situation and behavior. In *Integrative Views of Motivation, Cognition, and Emotion.* Ed. W. D. Spaulding. Lincoln: University of Nebraska Press, 71-124.

Bargh, J., Gollwitzer, P., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81*, 1014-1027.

Batson, C. D. (2011). *Altruism in humans*. Oxford: Oxford University Press.

Baxley, A.M. (2007). The price of virtue. *Pacific Philosophical Quarterly, 88*, 403-23.

Beckman, J., & Heckhausen, H. (2018). Situational determinants of behavior. In *Motivation and Action*, 3rd Edition. Eds. J. Heckhausen and H. Heckhausen. New York: Springer, 113-62.

Bélanger, J.J., Vallerand, R.J., Kruglanski, A.W. (2013). When passion makes the heart grow colder: the role of passion in alternative goal suppression. *Journal of Personality and Social Psychology, 104,* 126-47.

Belyavsky-Bayuk, J., Janiszewski, C., & Leboeuf, R. A. (2010). Letting good opportunities pass us by: Examining the role of mind-set during goal pursuit. *Journal of Consumer Research, 37*, 570-583.

Besser-Jones, L. (2008). Social psychology, moral character, and moral fallibility. *Philosophy and Phenomenological Research, 76*, 310–332.

Bengson, J. (2015). The intellectual given. *Mind, 124*(495), 707-760.

Bieleke M., Gollwitzer P. M., Oettingen G., Fischbacher U. (2017). Social value orientation moderates the effects of intuition versus reflection on responses to unfair ultimatum offers. *Journal of Behavioral Decision Making, 30,* 569- 581.

Blum, L. (1994). *Moral Perception and Particularity.* New York: Cambridge University Press.

Brady, M. (2013). *Emotional Insight: The Epistemic Role of Emotional Experience.* Oxford: Oxford University Press.

Bruers, S.. (2016). In search of moral illusions. *Journal of Value Inquiry,* 50, 283–303.

Bry, C., Büttner, O.B., Wieber, F., Schulz, A.M., Bayer, U.C., Florack, A., & Gollwitzer, P.M. (2014). Visual attention and goal pursuit: Deliberative and implemental mindsets affect breadth of attention. *Personality and Social Psychology Bulletin, 40*(10), 1248–1259.

Cameron, C.D., Payne, B.K., Sinnott-Armstrong, W., Scheffer, J.A., Inzlicht, M. (2017). Implicit moral evaluations: A multinomial modeling approach. *Cognition 158,* 224-241.

Carlson, E.N. (2013). Overcoming barriers to self-knowledge: Mindfulness as a path to seeing yourself as you really are. *Perspectives on Psychological Science, 8*, 173-86.

Carver, C. S., and Scheier, M. F. (2001). *On the Self-Regulation of Behavior.* Cambridge: Cambridge University Press.

Carver, C. S., and Scheier, M. F. (2003). Self-regulatory perspectives on personality," In *Handbook of Psychology, Vol 5: Personality and Social Psychology*. Eds. Millon, Theodore and Lerner, Melvin J. Hoboken, NJ: John Wiley & Sons, Inc., 185-208.

Cassidy, C. M., Balsam, P. D., Weinstein, J. J., Rosengard, R. J., Slifstein, M., & Daw, N. D. (2018). A perceptual inference mechanism for hallucinations linked to striatal dopamine. *Curr Biol, 28*, 503–514.e4.

Chen, X. E., & Latham, G. P. (2014). The effect of priming learning vs performance goals on a complex task. *Organizational Behavior and Human Decision Processes, 125*, 88–97.

Cholbi, M. (2007). Moral expertise and the credentials problem. *Ethical Theory & Moral Practice*, 10: 323–334.

Chudnoff, Eli. (2020*). Forming Impressions: Expertise in Perception and Intuition.* Oxford: Oxford University Press.

Clark, A. (2019). Perception as controlled hallucination: Predictive Processing and the nature of conscious experience. *Edge*. Accessed February 13, 2023. https://www.edge.org/conversation/andy_clark-perception-as-controlled-hallucination.

Colzato L. S., Ozturk A., & Hommel B. (2012). Meditate to create: the impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Frontiers in Psychology, 3*, 1-5.

Costa, P. T., Herbst, J. H., McCrae, R., Samuels, J., and Ozer, D. J. (2002). The replicability and utility of three personality types. *European Journal of Personality*, *16*, S73–S87.

Cowan, R. (2015). Perceptual intuitionism. *Philosophy and Phenomenological Research, 90*, 164-193.

Crisp, R. (2007). Ethics without reasons. *Journal Moral Philosophy, 4*(1), 40-49.

Crisp, R. (2014). Moral testimony pessimism: A Defence. *Proceedings of the Aristotelian Society Supplementary Volume,* 88(1), 129–43.

Crocq, M. (2013). Milestones in the history of personality disorders. *Dialogues in Clinical Neuroscience, 15*(2): 147-153.

Cullison, A. (2010). Moral perception. *European Journal of Philosophy, 18*(2), 159-175.

Danziger, N., Prkachin, K. M., Willer, J.C. (2006). Is pain the price of empathy? The perception of others' pain in patients with congenital insensitivity to pain. *Brain, 129*(9), 2494-507.

Darnell, C., Gulliford, L., Kristjánsson, K., & Paris, P. (2019). Phronesis and the knowledge-action gap in moral psychology and moral education: A new synthesis? *Human Development, 62*, 101–29.

Delgado-Pastor L. C., Perakakis, P., Subramanya, P., Telles, S., and Vila, J. (2013). Mindfulness (Vipassana) meditation: Effects on P3b event-related potential and heart rate variability. *International Journal of Psychophysiology, 90*, 207-214.

Demarree KG, Loersch C, Briñol P, Petty RE, Payne BK, Rucker DD. (2012). From primed construct to motivated behavior: validation processes in goal pursuit. *Personality & Social Psychology Bulletin, 38*(12), 1659-1670.

Deonna, J. A. & Teroni, F. (2015). Emotions as attitudes. *Dialectica*, *69*(3): 293–311.

Dijkersterhuis, A., Spears, R., Postmes, T. Stapel, D.A., Koomen, W., van Knippenberg, A. & Scheepers, D. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology, 75*, 862-871.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology, 74*, 865-877.

Dijksterhuis, A., & van Knippenberg, A. (2000). Behavioral indecision: Effects of self-focus on automatic behavior. *Social Cognition, 18*, 55-74.

Dijksterhuis, A., van Knippenberg, A., Holland, R. W. (2014). Evaluating behavioral priming research: Three observations and a recommendation. In *Understanding Priming Effects in Social Psychology*. Ed. D. C. Molden. New York: The Guilford Press, 205-17.

Doris, J. (1998). Persons, situations, and virtue ethics. *Noûs, 32*, 504-30.

Doris, J. (2002). *Lack of Character*. New York: Cambridge University Press.

Doyen, S., Klein, O., Pichon, C. L. & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE, 7.*

Driver, J. (2006). Autonomy and the asymmetry problem for moral expertise. *Philosophical Studies, 128*(3), 619-644.

Favre, M. R., Markram, H., Markram, K. (2019). Individual differences in sensory sensitivity: Further lessons from an Autism model. *Cognitive Neuroscience, 10*(3), 171-173.

Feldman Barret, L. (2016). The theory of constructed emotion: An active inference account of interoception and categorization, *Social Cognitive and Affective Neuroscience, 12*(1), 1–23.

Feldman Barret, L. (2017). *How Emotions are Made: The Secret Life of the Brain.* Boston: Houghton Mifflin Harcourt.

Field, C. (2022). At least you tried: The value of de dicto concern to do the right thing. *Philosophical Studies*, *179*, 2707–2730.

Fiser, J. (2009). Perceptual learning and representational learning in humans and animals. *Learning & Behavior, 37*(2), 141-53.

Fitzgibbon, B.M., Giummarra, M.J., Georgiou-Karistianis, N., Enticott, P.G., & Bradshow, J.L. (2010). Shared pain: From empathy to synaesthesia. *Neuroscience and Biobehavioral Reviews, 34,* 500-512.

Flanagan, O. (1991). *Varieties of Moral Personality*. Cambridge, MA: Harvard University Press.

Foot, P. (1983). Moral Realism and moral dilemma. *The Journal of Philosophy, 80*(7): 379-398.

Freund, W., Weber, F., Billich, C., Birklein, F., Breimhorst, M., Schuetz, U. H., (2013). Ultra-marathon runners are different: Investigations into pain tolerance and personality traits of participants of the TransEurope FootRace 2009. *Pain Practice, 13*(7), 524-532.

Friedman, M., & Roseman, R. (1974). *Type A Behavior and Your Heart.* New York: Alfred A. Knopf.

Frijda, Nico H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Fujino, M., Ueda, Y., Mizuhara, H., Saiki, J., & Nomura, M. (2019). Open monitoring meditation reduces the involvement of brain regions related to memory function. *Scientific Reports, 8*, 9968.

Galinsky, A. D., Wang, C. S., & Ku, G. (2008). Perspective-takers behave more stereotypically. *Journal of Personality and Social Psychology, 95*, 404-19.

Ganegoda, D. B., Latham, G. P., & Folger, R. (2011). The effect of subconscious goal setting on organizational justice. In *Meeting of the Society for Industrial and Organizational Psychology*. Chicago, IL.

Garland, E. L., & Howard, M. O. (2013). Mindfulness-oriented recovery enhancement reduces pain attentional bias in chronic pain patients. *Psychotherapy and Psychosomatics, 82*, 311-318.

Garland, E. L., Farb, N. A., Goldin, P., & Fredrickson, B. L. (2015). Mindfulness broadens awareness and builds eudaimonic meaning: A process model of mindful positive emotion regulation. *Psychological Inquiry, 26*, 293–314.

Geisler M, Ritter A, Herbsleb M, Bär KJ, and Weiss T. (2021). Neural mechanisms of pain processing differ between endurance athletes and nonathletes: A functional connectivity magnetic resonance imaging study. *Human Brain Mapping, 42*(18), 5927-5942.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Goldie, P. (2007). Seeing the kind thing to do: Perception and emotion in morality. *Dialectica, 61*(3), 347-361.

Gollwitzer, P. W. (2003). Why we thought the action mindsets affect illusions of control. *Psychological Inquiry, 14*, 261–269.

Gollwitzer, P. W. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*(7), 493-503.

Gollwitzer, P. W., Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology, 73*, 186-199.

Gollwitzer, P. M., & Sheeran, P. (2009). Self-Regulation of consumer decision making and behavior: The role of implementation intentions. Journal of Consumer Psychology, 19, 593–607.

Gregory, R. L. (1968). Visual illusions. *Scientific American,* 21966–76.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, 352, 1121–1128.

Gregory, R. L. (2009). *Seeing Through Illusions*. Oxford: Oxford University Press.

Grimm, S. (2012). The value of understanding. *Philosophy Compass 7/2:* 103-117.

Haddock, G. Macrae, C. N., & Fleck, S. (2002). Syrian science and smart supermodels: On the when and how of perception-behavior effects. *Social Cognition, 20*(6), 461-79.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol Rev, 108*(4), 814-34.

Hansen, J., & Wanke, M. (2009). Think of capable others and you can make it! Self-efficacy mediates the effect of stereotype activation on behavior. *Social Cognition, 27*, 76-88.

Harman, G. (1977). Ethics and observation. *The Nature of Morality*. Oxford: Oxford University Press, pp. 3–10.

Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society, 99*, 315-331.

Harman, G. (2000). The nonexistence of character traits. *Proceedings of the Aristotelian Society, 100*, 223- 226.

Harman, G. (2003). No character or personality. *Business Ethics Quarterly, 13*, 87-94.

Harris, C. R., Coburn, N., Rohrer D., & Pashler, H. (2013). Two failures to replicate high performance-goal priming effects. *PLoS ONE 8*(8), e72467.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion, 11*, 101-20.

Hills, A. (2009). Moral testimony and moral epistemology. *Ethics,* 120(1), 94-127.

Hills. A. (2015). Understanding why. *Noûs 49*(2), 661-688.

Hills, A. (2016). Virtue and cognition. In I. Fileva (Ed.) *Questions of Character,* pp. 147-157.

Hodgins, H. S., & Adair, K. C. (2010). Attentional processes and meditation. *Conscious Cognition, 19*, 872-878.

Hong, N.T. (2019). Unintentional unethical behavior: The mediating and moderating roles of mindfulness. *International Journal of Ethics and Systems, 36*(1), 98-118.

Howell, R. (2014). Google morals, virtue, and the asymmetry of deference. *Noûs 48*(3), 389-415.

Howes, R. J., & Carskadon, T. G. (1979). Test-retest reliabilities of the Myers-Briggs Type Indicator as a function of mood changes. *Research in Psychological Type*, *2*: *7*.

Hursthouse, R.. (2006). "Are virtues the proper starting point for morality?" In *Contemporary Debates in Moral Theory,* Ed. J. Dreier. Blackwell Publishing, pp. 99-112.

Jackson, P. L., Meltzoff, A. N., Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage, 24*(3), 771-9.

Jacobson, D. (2005). Seeing by feeling: Virtues, skills, and moral perception. *Ethical Theory and Moral Practice, 8*(4), 387-409.

James, W. (1884). What is an emotion? *Mind, 9*:34, 89–90.

Janczyk, M., Dambacher, M., Bieleke, M., Gollwitzer, P.M. (2015). The benefit of no choice: Goal-directed plans enhance perceptual processing. *Psychological Research, 79*(2), 206-20.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research*. New York: The Guilford Press, 114-158.

Kamtekar, R. (2004). Situationism and virtue ethics on the content of our character. *Ethics, 114*, 458-491.

Klatzsky, R. L., & Creswell, J. D. (2014). An intersensory interaction account of priming effects-and their absence. *Perspectives on Psychological Science, 9*, 49-58.

Kleiman-Weiner, Max, Saxe, Rebecca, & Tenenbaum, Joshua B. (2017). Learning a commonsense moral theory. *Cognition, 167*, 107-123.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. *Advanced Experimental Social Psychology, 34*, 331–378.

Lamm, C., Decety, J.,& Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage, 54*, 2492- 2502.

Landy, Michael S., Banks, Martin S., and Knill, David C. (2011). Ideal-observer models of cue integration. In J. Trommershauser, K. Kording, and M. S. Landy (eds.), *Sensory Cue Integration*, pp. 5-29. New York: Oxford University Press.

Larsen, R. J. & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality, 21*(1), 1-39.

Latham, G. P. (2018). The effect of priming goals on organizational related behavior: My transition from skeptic to believer. In G. Oettingen, A. T. Sevincer, & P. M. Gollwitzer (Eds.), *The Psychology of Thinking About the Future.* New York: Guilford Press, 392-404.

Layton, R. L. & Muraven, M. (2014). Self-control linked with restricted emotional extremes. *Personality and Individual Differences, 58:* 48-53.

LeBoeuf, R. A., & Estes, Z. (2004). "Fortunately, I'm no Einstein": Comparison relevance as a determinant behavioral assimilation and contrast. *Social Cognition, 22*, 607-636.

Lippelt, D. P., Hommel, B., Colzato, L. S. (2014). Focused attention, open monitoring and loving kindness meditation: effects on attention, conflict monitoring, and creativity – A review. *Frontiers in Psychology, 5,* 1083.

Liu, Y., Meng, J., Yao, M., Ye, Q., Fan, B., & Peng, W. (2019). Hearing other's pain is associated with sensitivity to physical pain: An ERP study. *Biological Psychology, 145:* 150-8.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task performance. *American Psychologist, 57*, 705–17.

Locke, E. A., & Latham, G. P. (2013). *New Developments in Goal Setting and Task Performance*. New York, NY: Routledge.

Locke, E.A., & Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *APA – Legacies in Motivational Science, 5*(2), 95-105.

Loersch, C., & Payne, B. K. (2011). The situated inference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science, 6*, 234-52.

Loersch, C., & Payne, B. K. (2014). Situated inference and the what, who, and where of priming. *Social Cognition, 32*, 137-151.

Lowery, B. S., Eisenberger, N. I., Hardin, C. D., & Sinclair, S. (2007). Long-term effects of subliminal priming on academic performance. *Basic and Applied Social Psychology, 29,* 151-157.

Masicampo, E. J., & Baumeister, R. F. (2012). Committed but close-minded: When making a specific plan for a goal hinders success. *Social Cognition, 30*, 37-55.

McBrayer, J. (2010). A limited defense of moral perception. *Philosophical Studies, 149:* 305–320

McDowell, J. (1978). Are moral requirements hypothetical imperatives? *Aristotelian Society Supplementary, 52*(1), 13–29.

McDowell, John. (1998). *Mind, Value, and Reality.* Cambridge: Harvard University Press.

McGrath, S. (2018). Moral perception and its rivals. In R. Cowan, A. Bergqvist (Eds.), *Evaluative Perception.* pp. 161-182.

McLaughlin, B. P. (2010). The Representational Vs. The Relational View of Visual Experience. *Royal Institute of Philosophy Supplement* 85(67), 239–262.

McShane, P. J. (2018). Moral testimony and moral understanding. *Philosophical Studies, 175*(3), 629-647.

Millar, Boyd. (2015). Naïve Realism and illusion. *Ergo,* 24(2), 607-625.

Miller, C. B. (2003). Social psychology and virtue ethics. *The Journal of Ethics 7*(4), 365-392.

Miller, C. B. (2010). Character traits, social psychology, and impediments to helping behavior. *Journal of Ethics and Social Philosophy 5*(1), 1-36.

Miller, C. B. (2013). *Moral Character: An Empirical Theory.* Oxford: Oxford University Press.

Miller, C. B. (2014). *Character and Moral Psychology.* Oxford: Oxford University Press.

Mischkowski, D., Crocker, J., Way, B. M. (2016). From painkiller to empathy killer: acetaminophen (paracetamol) reduces empathy for pain. *Soc Cogn Affect Neurosci, 11*(9), 1345-53.

Moore, A., & Malinowski, P. (2009). Meditation, mindfulness, and cognitive flexibility. *Conscious Cognition, 18*, 176-186.

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology, 38*, 397-404.

Moss, J. (2009). Akrasia and perceptual illusion. *Archiv für Geschichte der Philosophie, 91*(2), 119-156.

Mukherjee, S., Srinivasan, N., Kumar, N., & Manjaly, J. A. (2018). Perceptual broadening leads to more prosociality. *Frontiers in psychology, 9*, 1821.

Nussinson, R., Seibt, B., Hafner, M., & Strack, F. (2010). Come a bit closer: Approach motor actions lead to feeling similar and behavioral assimilation. *Social Cognition, 28*, 40-58.

Onu, D., Kessler, T., Smith, J. R. (2016). Admiration: A conceptual review. *Emotion Review, 8*(3).

Ordóñez, L., Schweitzer, M., Galinksy, A., Bazerman, M. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives, 23*(1), 6-16.

Papies, E. K. (2016). Health goal priming as a situated intervention tool: How to benefit from nonconscious motivational routes to health behaviour. *Health Psychology Review, 10*(4), 408- 424.

Parks-Stamm, E. K., Gollwitzer, P. M., & Oettingen, G. (2007). Action control by implementation intentions: Effective cue detection and efficient response initiation. *Social Cognition, 25*, 247- 264.

Payne, B. K., Brown-Iannuzzi, J. L., & Loersch, C. (2016). Replicable effects of primes on human behavior. *Journal of Experimental Psychology: General, 145*(10), 1269–1279.

Pellicano, E., Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends Cognitive Science, 6*(10), 504-10.

Pettersen, S. D., Aslaksen, P. M., Pettersen, S. A. (2020). Pain processing in elite and high-level athletes compared to non-athletes. *Frontiers in Psychology, 11*, 1-9.

Pettinger, D. J. (1993). The utility of the Myers-Briggs Type Indicator. *Review of Educational Research, 63*(4), 467-488.

Powers A. R., Mathys C., and Corlett P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science, 357*, 596–600.

Prinz, J. (2004). *Gut Reactions: A Perceptual Theory of Emotion.* Oxford: Oxford University Press.

Prinz, J. (2006). Is emotion a form of perception? *Canadian Journal of Philosophy, 36*, 137-160.

Rabinoff, E. (2018). *Perception in Aristotle's Ethics.* Evanston, IL: Northwestern University Press.

Railton, P. (2011). Two cheers for virtue: Or, might virtue be habit forming? In *Oxford Studies in Normative Ethics, Vol 1*, ed. M. Timmons. New York: Oxford University Press.

Railton, P. (2014). The affective dog and it's rational tail: Intuition and attunement. *Ethics, 124*(4), 813-859.

Railton, P. (2017). Moral Learning: Conceptual foundations and normative relevance. *Cognition, 167,* 172-190.

Ratnayake, S. (2019, July 25). The problem of mindfulness. *Aeon.* https://aeon.co/essays/mindfulness-is-loaded-with-troubling-metaphysical-assumptions.

Reardon, P. H. (2013). Jesus, jail, and jury duty. *Preacher's Institute.* Retrieved from https://preachersinstitute.com/2013/07/23/why-confessors-must-never-serve-on-a-jury/#:~:text=When%20he%20hears%20Confessions%2C%20the%20priest%20functions%20in,is%20always%20on%20the%20side%20of%20the%20sinner on April 28, 2023.

Ricciardelli, P., Carcagno, S., Vallar, G., & Bricolo, E. (2013). Is gaze following purely reflexive or goal-directed instead? Revisiting the automaticity of orienting attention by gaze cues. *Experimental Brain Research, 224*: 93.

Roberts, R. C. (1989). Aristotle on virtues and emotions. *Philosophical Studies 56*(3), 293 – 306.

Roberts, R. C. (2003). *Emotions: An Essay in Aid of Moral Psychology.* Cambridge: Cambridge University Press.

Roberts, R. C. (2013). *Emotions in the Moral Life.* Cambridge: Cambridge University Press.

Rodgers, T. J., and Warmke, B. (2015). Situationism against Situationism. *Ethical Theory and Moral Practice, 18*(1), 9-26.

Ross, L. & Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology.* Mcgraw-Hill Book Company.

Ruedy, N.E. & Schweitzer, M.E. (2010). In the moment: The effect of mindfulness on ethical decision making. *Journal of Business Ethics, 95*, 73-87.

Russell, D. (2009). *Practical Intelligence and the Virtues*. Oxford: Oxford University Press.

Russell, D (2014). Aristotelian Virtue Theory: After the person-situation debate. *Revue Internationale de Philosophie* 267(1), 37-63.

Scarantino, A. (2014). "The Motivational Theory of Emotions", in Justin D'Arms and Daniel Jacobson (eds.), *Moral Psychology and Human Agency*, Oxford: Oxford University Press, 156–185.

Scarantino, A. (2015). "Basic Emotions, Psychological Construction and the Problem of Variability", in L. F. Barrett and & J.A. Russell (Eds.), *The Psychological Construction of Emotion*, pp. 334-376.

Scholl, Brian J., & Gao, Tao. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197-230). Cambridge, MA: MIT Press.

Schröder, T., & Thagard, P. (2013). The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review, 120*, 255-80.

Schubert, T. W., & Hafner, M. (2003). Contrast from social stereotypes in automatic behavior. *Journal of Experimental Social Psychology, 39*, 577-584.

Schuster, N. (2020). Complex harmony: Rethinking the virtue-continence distinction. *The Journal of Ethics, 24*, 225–240.

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M. & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different methods of measurement. *Journal of Cross Cultural Psychology, 32*(5), 519-42.

Segall, Marshall. H., Campbell, Donald. T. & Herskovits, Melville. J., (1966). *Influence of Culture on Visual Perception* Indianapolis, IN: Bobbs-Merill.

Seidman, J. (2005). Two sides of 'silencing'. *The Philosophical Quarterly, 55*(218), 68-77.

Seriès, P. & Seitz, A. (2013). Learning what to expect (in visual perception). *Frontiers in Human Neuroscience, 7*, 668.

Seth, A. K. & Critchley, H. D. (2013). Extending predictive processing to the body: emotion as interoceptive inference. *Behavioral and Brain Sciences, 36*(3), 227-8.

Shah, J., Friedman, R., & Kruglanski, A. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology, 83*(6), 1261-1280.

Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: bottom-up priming of goals by means and its effects on self-regulation. *Journal of Personality and Social Psychology, 84*, 1109– 1122.

Shams, L. and Kim, Robyn. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews, 7*: 269–84.

Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L, Cenac, Z., et al. (2013). Priming intelligent behavior: An elusive phenomenon. *PLoS ONE, 8,* 256515.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frifth C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303*, 1157-1162.

Siversten, S. S. (2019). On the practical impossibility of being both well-informed and impartial. *Erasmus Journal for Philosophy and Economics, 12*(1), 52-72.

Skyrms, B. (1980). *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws.* New Haven: Yale University Press.

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J., Davidson, R. J. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology, 5*(6) e38.

Slote, M. 'Sentimentalist Virtue Ethics', in Nancy E. Snow (ed.), *The Oxford Handbook of Virtue*, Oxford Handbooks (2018; online edn, Oxford Academic, 6 Dec. 2017), https://doi.org/10.1093/oxfordhb/9780199385195.013.47, accessed 17 Mar. 2023.

Smith, A. D. (2002). *The Problem of Perception.* Cambridge, Mass.: Harvard University Press

Smith, A. D. (2010). Disjunctivism and illusion. *Philosophy and Phenomenological Research*, 80(2), 384–410.

Snow, N. (2006). Habitual virtuous actions and automaticity. *Ethical Theory and Moral Practice, 9*, 545- 561.

Snow, N. (2010). *Virtue as Social Intelligence: An Empirically Grounded Theory.* New York: Routledge.

Snow, N. (2016). How habits make us virtuous. In *Developing the Virtues: Integrating Perspectives*, eds. J. Annas, D. Narvaez, and N. E. Snow. Oxford: Oxford University Press.

Stajkovic, A. D., Latham, G. P., Sergent, K., & Peterson, S. (2018). Prime and performance: Can a CEO motivate employees without their awareness? *Journal of Business Psychology*, 1-12.

Stark, S. 2001. Virtue and emotion. *Noûs, 35*(3), 440-55.

Staw, B. M., & Boettger, R. D. (1990). Task revision: A neglected form of work performance. *Academy of Management Journal, 33*(3), 534-559.

Stewart, V. M. (1973). Tests of the "carpentered world": Hypothesis by race and environment in America and Zambia. *International Journal of Psychology, 8*(2), 83–94.

Stichter, M. (2018). *The Skillfulness of Virtue: Improving Our Moral and Epistemic Lives.* Cambridge: Cambridge University Press.

Stocker, A. A. & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*, 578–585.

Stohr, K., (2003). Moral cacophony: When continence is a virtue. *Journal of Ethics, 7*, 339-363.

Strevens, M. (2013). No understanding without explanation. *Studies in History and Philosophy of Science Part A, 44*(3), 510-515.

Tappolet, C. (2016). *Emotions, Values and Agency.* Oxford: Oxford University Press.

Tessman, L. (2005). *Burdened Virtues: Virtue Ethics for Liberatory Struggles.* Oxford: Oxford University Press.

Upton, C. (2017). Meditation and the cultivation of virtue. *Philosophical Psychology, 30*, 373-394.

Valentine, E. R., & Sweet, P. L. (1999). Meditation and attention: a comparison of the effects of concentrative and mindfulness meditation on sustained attention. *Mental Health, Religion, & Culture, 2*, 59-70.

Van Boxtel, J. J., Lu, H. (2013). Impaired global, and compensatory local, biological motion processing in people with high levels of autistic traits. *Front Psychol, 23* (4), 209.

Van der Laan, L. N., Papies, E. K., Hooge, I. T. C., & Smeets, P. A. M. (2017). Goal-directed visual attention drives health goal priming: An eye-tracking experiment. *Health Psychology, 36*(1), 82– 90.

Van Leeuwen, K., De Fruyt, F., and Mervielde, I. (2004). A longitudinal study of the utility of the resilient, overcontrolled, and undercontrolled personality types as predictors of children's and adolescents' problem behaviour. *International Journal of Behavioral Development, 28*(3), 210–220.

Veling, H. & van Knippenberg, A. (2006). Shielding intentions from distraction: Forming an intention induces inhibition of distracting stimuli. *Social Cognition, 24*(4), 409–425.

Vigani, D. (2019). Virtuous construal: In defense of silencing. *Journal of the American Philosophical Association,* 229-45.

Walker, A. M. D. (1989). Virtue and character. *Philosophy, 64*(249), 349-362.

Ward J. (2019). Individual differences in sensory sensitivity: A synthesizing framework and evidence from normal variation and developmental conditions. *Cognitive Neuroscience, 10*(3), 139-157.

Webb, T. L., & Sheeran, P. (2007). How do implementation intentions promote goal attainment? A test of component processes. *Journal of Experimental Social Psychology, 43*, 295-302.

Webber, J. (2016). Instilling Virtue. In *From Personality to Virtue: Essays in the Philosophy of Character*, eds. A. Masala & J. Webber. Oxford: Oxford University; pp. 135-154.

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin, 142*(5), 472-497.

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin, 142*(5), 472–497.

Weiss, Y., Simoncelli, E. P. & Adelson, Edward H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604.

Werner, P. J. (2016). Moral perception and the contents of experience. *Journal of Moral Philosophy 13*(3), 294-317.

Werner, P. J. (2020). Which moral properties are eligible for perceptual awareness? *Journal of Moral Philosophy, 17*(3), 290-319.

Westra, E. (2018). Character and theory of mind: An integrative approach. *Philosophical Studies, 175*(5), 1217-1241.

Westra, E. (2020). Getting to know you: Accuracy and error judgments of character. *Mind & Language, 35*(5), 583-600.

Westra, E. (2022). In defense of ordinary moral character judgment. *Erkenntnis, 87*, 1461-1479.

Wieber, F., & Sassenberg, K. (2006). I can't take my eyes off of it – Attention attraction effects of implementation intentions. *Social Cognition, 24*, 723-752.

Wiland, E. (2021). *Guided by Voices: Moral Testimony, Advice, and Forging a 'We'*. Oxford: Oxford University Press.

Wilkenfeld, D. (2020). Moral understanding and moral illusions. *Thought, 9*(1), 25-44.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322*(5901), 606–607.

Wolf, S. (2007). Moral psychology and The Unity of the Virtues. *Ratio*, *20*(2), 145-167.

Woodward, J. (2003*). Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. & Allan, J. (2007) Moral intuition: Its neural substrates and normative significance. *Journal of Physiology – Paris, 101,* 179-202.