SELF-BELIEFS AND EPISTEMIC JUSTIFICATION

SELF-BELIEFS AND EPISTEMIC JUSTIFICATION:

WHAT MAKES OUR SELF-BELIEFS ABOUT OUR PERSONALITY TRAITS

EPISTEMICALLY JUSTIFIED?

By SHAHDAH MAHHOUK

A Thesis Submitted to the School of Graduate Studies in Fulfilment of the Requirements

for the Degree Master of Arts McMaster University © Copyright by Shahdah Mahhouk,

August 2023

McMaster University MASTER OF ARTS (2023) Hamilton, Ontario (Philosophy)

TITLE: Self-beliefs and epistemic justification: What makes our self-beliefs about our personality traits epistemically justified? AUTHOR: Shahdah Mahhouk, B.A. (University of Toronto) SUPERVISOR: Professor Megan Stotts. NUMBER OF PAGES: vii, 124

Table of Contents

Abstract

I explore the epistemic justification of self-beliefs regarding personality traits within the internalism-externalism debate. Historically, the question of epistemic justification of self-beliefs has been discussed only with respect to our beliefs about our current mental states while the epistemic justification of our self-beliefs about our personality traits was assumed not to be any different from the justification of our beliefs about the external world. However, I use empirical psychology to highlight a few unique characteristics of our self-beliefs about personality traits that make the typical application of internalist or externalist standards less straightforward. These characteristics have to do with the biases and the self-verification that accompany our self-beliefs about our personality traits. I argue that externalism, in general, and virtue reliabilism, in particular, are more suitable to the context of our self-beliefs about our personality traits than other theories of justification. However, I contend that within the virtue reliabilism framework, a self-belief-forming process can become more competent if it generates self-belief from the instances where individuals manifest the trait in question while having the motivation and opportunity to do otherwise. I show how this condition makes the self-belief-forming process more competent and, therefore, makes the produced self-beliefs more epistemically justified.

Keywords: Self-Beliefs, Personality Traits, Epistemic justification, Externalism, Internalism, Mentalism, Process Reliabilism, Virtue Reliabilism, Empirical Psychology.

Introduction

Why Does Justification Matter?

In epistemology, it is generally endorsed that holding justified beliefs is more rational than holding unjustified ones. This idea has motivated several major debates throughout the history of epistemology. One example is the debates concerning skepticism.[1] As some philosophers suggest, the core of any potential responses to any skeptical problem is about providing a kind of *justification* for the beliefs that a skeptic may doubt (e.g., beliefs from the senses, beliefs about the external world...etc.).[2] In other words, a successful response to any skeptical problem must either show why such beliefs are justified, grounded, or warranted, or why we have some reasons to hold them despite the skeptical challenge. Thus, in many cases, the core of the challenges from skepticism can be somehow reframed in terms of challenges about justifications.[3]

Another example of how justification has motivated major debates in the history of epistemology is the debate concerning the traditional pre-Gettier conception of knowledge as a justified true belief.[4] Prior to t

---

[1] Declan Smithies, "Why Justification Matters," In *Epistemic Evaluation* (Oxford: Oxford University Press, 2015). 224–244.

[2] Richard H. Popkin, *History of Scepticism: From Savonarola to Bayle* (Cary: Oxford University Press, 2003) 8-34.

[3] Smithies, "Why Justification Matters," In *Epistemic Evaluation,* 224–244.

[4] Linda Zagzebski, "The Inescapability of Gettier Problems," *The Philosophical Quarterly* 44, no. 174 (1994): 65–73.

he publication of Gettier's paper, philosophers believed that justification is the link that would turn true beliefs into knowledge.[5] The pre-Gettier approach to justification can go back to Plato's *Theaetetus*, where Socrates argues that "all who have any right opinion will be found to have it with the addition of *rational explanation*, and there will henceforth be no possibility of right opinion apart from knowledge."[6] The basic idea, for Socrates, is that if one knows a certain proposition, then, in addition to having a true belief, one must have had a right, a warrant, or a good reason to believe in that proposition. The same approach to justification can be seen in the following example from the early twentieth century in C. I. Lewis, who argues that "knowledge is [the] belief which is not only true but also is justified in its believing attitude."[7] Thus, if knowledge is the core of the epistemological inquiries and if justification is the link that turns beliefs into knowledge, then it makes sense that figuring out the nature and criteria of that link was one of the core epistemological inquiries at that time.

However, in his influential article "Is Justified True Belief Knowledge?", Gettier illustrates various cases where someone does not have knowledge despite having true justified beliefs.[8] [9] Thus, since the publication of Gettier's cases, it becomes one of the very

---

[5] The question of justification is usually discussed in relevance to propositional knowledge (i.e., knowledge-that, where the that-clause expresses a proposition) rather than non-propositional knowledge (i.e., knowledge-how, where the how-clause expresses an emotion, a skill, or an experience). This approach to knowledge remains relevant in contemporary epistemology. Therefore, whenever I mention knowledge, I am specifically referring to propositional knowledge.

[6] Plato, "*Theaetetus*," (Oxford: Clarendon Press, 1973): 201d–210a.

[7] Clarence Irving Lewis, *An Analysis of Knowledge and Valuation*, (La Salle, Ill: The Open court publishing, 1946) 9.

[8] Edmund Gettier, "Is Justified True Belief Knowledge?" *Analysis (Oxford)* 23, no. 6 (1963): 121–23.

[9] Here is one example of a Gettier case: Smith and Jones have applied for the same job. Smith has strong evidence to believe that both he and Jones are highly qualified for the position. The hiring manager informs Smith that Jones will likely be offered the job, based on the evidence supporting Jones'

few consensuses among contemporary epistemologists that the traditional conception of knowledge as justified true beliefs fails.[10] Nevertheless, this did not make the question of justification any less important. Instead, numerous epistemologists shifted their attention from inquiries concerning knowledge to rationality, and, as a result, they focused more on justification since they perceived it to be more related to rationality than knowledge.[11] Thus, despite how knowledge is viewed, it is still generally endorsed that holding justified beliefs is more rational than holding unjustified ones.[12] As Wright puts it "We can live with the concession that we do not, strictly, know some of the things we believed ourselves to know, provided we can retain the thought that we are fully *justified* in accepting them."[13]

What Do We Mean by Epistemic Justification?

Epistemic justification has been generally conceptualized in history and contemporary epistemology as the right standing of a belief that makes believing rational

---

qualifications. From this information, Smith logically concludes that "the man who will get the job has ten coins in his pocket." However, in contrast to the hiring manager's prediction, Smith gets the offer instead of Jones. What Smith doesn't know is that, coincidentally, he himself also has ten coins in his pocket. Thus, his belief that "the man who will get the job has ten coins in his pocket" is still true and justified. But we do not think that Smith has knowledge in this case, and this is because luck and coincidence have interfered with the link between justification and truth.

[10] Zagzebski, "The Inescapability of Gettier Problems," 65–73.

[11] For example, see, Richard Foley, "Rationality and Perspective." *Analysis* 53, no. 2 (1993): 65–68 and Linda Zagzebski, "The Methodology of Epistemology," Chapter One. In *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*, (Cambridge: Cambridge University Press, 1996) 22-35.

[12] The relation between justification and rationality is controversial in contemporary epistemology. Certain philosophers, such as Cohen (1984), believe that a justified belief is equivalent to a rational belief. However, philosophers like Lyons (2016) argue that rationality and justification are distinct epistemic characteristics.

In this thesis, my focus will be solely on justification without addressing the connection to its relations to rationality.

[13] Crispin Wright, "Scepticism and Dreaming: Imploding The Demon*," Noûs (Bloomington, Indiana) 25, no.* 2 (1991): 88. Emphasis is mine.

(i.e., pertaining to reasons), responsible (i.e., expressing an epistemic virtue), or more likely to be true (i.e., aiming at the truth).[14] The word 'epistemic' comes from *episteme,* which is the Greek word for knowledge.[15] This distinguishes epistemic justification from non-epistemic justification. For example, epistemic justification can be distinguished from pragmatic justification, where beliefs or actions are evaluated based on practical considerations, such as the usefulness or effectiveness of a belief in achieving desired outcomes or goals.[16] Also, epistemic justification can be distinguished from moral justification, where beliefs or actions are evaluated based on moral principles, values, or ethical considerations.[17] Furthermore, epistemic justification can be distinguished from aesthetic justification, where the justification of judgments or evaluations related to beauty, harmony, creativity, art, or emotional impacts.[18]

The word 'justification', however, corresponds to a cluster of evaluative notions with respect to knowledge such as 'rationality', 'reasonability' and 'warrant'.[19] This evaluative aspect led many philosophers to argue that justification is inherently

---

[14] William Alston, "Epistemic Justification," in *Essays in the Theory of Knowledge* (Cornell: Cornell University Press, 1989), 11.

[15] Alston, "Epistemic Justification," 11-9.

[16] Epistemic justification is and has always been thought of as something that is intimately related to truth. However, a number of epistemologists argue that pragmatical factors do in fact play an essential role in determining whether a certain belief is epistemically justified. This viewpoint is commonly referred to as the 'pragmatic encroachment' thesis'. This is to say that by solely considering truth-related aspects, we overlook a crucial element in our understanding of knowledge or rationality. I will sidestep this issue in this thesis as I have not taken a definitive stance on it. However, for more on pragmatic encroachment, see Fantl & McGrath (2002), (2007), and Schaffer (2006).

[17] For example, see Haslett, D. W., *Equal Consideration: A Theory of Moral Justification*. (Newark: University of Delaware Press, 1987) 20-43.

[18] For example, see Stephen Halliwell, "Justifying The World As An Aesthetic *Phenomenon,"* *The Cambridge Classical Journal 64* (2018): 91–112.

[19] Alston, 13-9.

normative.[20] This means that determining whether a belief is justified or not is about deciding whether we should or should not believe a certain proposition. However, some philosophers have rejected the normative approach to justification in favour of a more naturalistic approach.[21] In short, while normative approaches to justification focus more on the subjective perspective of the believer as a responsible agent, naturalistic theories focus more on the objective relationship between a belief-forming mechanism and reality.[22] Although whether justification is normative or naturalistic is a meta-philosophical question of it is own, this thesis takes a more naturalistic approach in the sense that it focuses on the *mechanism* that makes a certain *belief* more justified instead of what makes *us* justified in having a certain belief.

To conceptualize the notion of epistemic justification in a more concrete scenario, consider the following example. Suppose a police officer knocks on Amanda's door and informs Amanda that her partner has been seriously injured in a car accident. However, instead of believing the officer, Amanda believes that her partner has somehow survived untouched. She has no grounds to believe so, yet she believes it anyway. Amanda's belief that her partner had survived is not epistemically justified. She might have a psychological reason as she sought immediate comfort in believing her partner is safe. She might have a prudential reason as it could be in her best interest to embrace the belief that her partner is safe, so she does not collapse. However, with respect to her epistemic stance, she has no

---

[20] For example, see Sanford C. Goldberg, *To the Best of Our Knowledge: Social Expectations and Epistemic Normativity,* (Oxford: Oxford University Press, 2018), 13–47.

[21] For example, see Helen De Cruz, Boudry Maarten, De Smedt Johan, and Blancke Stefaan, "Evolutionary Approaches to Epistemic Justification," *Dialectica* 65, no. 4 (2011): 517–35.

[22] Alston, 18-20.

epistemic justification for her belief that her partner has survived. She has no good reason, nor does she form her belief via a reliable belief-forming process. Even if she turned out to be correct, Amanda would not be said to *know* or to be *rational* in believing that her partner had survived.

This is not to say that having an epistemic justification would necessarily lead to truth. In fact, many philosophers argue that having an epistemic justification for believing a certain proposition is compatible with the potential falsity of that proposition.[23] For example, while Amanda's belief that her partner has been injured as reported by the police officer is epistemically justified, it could be the case that the police officer is mistaken (e.g., having the incorrect name/address) and that Amanda's partner is actually safe. The potential error in the report does not change the fact that Amanda's belief that her partner has been injured is epistemically justified. This is because the epistemic justification of a proposition does not guarantee its truth.[24] Thus, justification is compatible with fallibility: it is possible for a false belief to sometimes be epistemically justified. Nevertheless, having justified beliefs can increase the *likelihood* that we arrive at truth; therefore, it is still more rational to hold justified beliefs than unjustified ones.

There are many questions to be asked about the nature and the structure of epistemic justification. I focus on the criteria we employ to assess the epistemic justification of the self-beliefs we have about our personality traits. I delve into this inquiry within the context

---

[23] See, for example, Richard Feldman who provided a defence of fallibilism in the context of self-knowledge of one's current mental state in "Fallibilism and Knowing That One Knows." The *Philosophical Review* 90, no. 2 (1981)

[24] Alston, 20-3.

of the internalism-externalism debate. Internalists argue that justification relies solely on an individual's mental state while externalists assert that justification encompasses external factors, such as reliable processes.[25] I examine both views in relation to some features specific to the self-beliefs we have about our personality traits.

Epistemic Justification, Beliefs, and Self-Beliefs

Epistemic justification is a property of a belief[26], where a belief can be understood as a mental state or an attitude in which an individual holds a certain proposition to be true.[27] In other words, beliefs are the mental states that reflect the individual's stand on the truth or falsity of a particular proposition.[28] Thus, a belief refers to the attitude one might have concerning a proposition. A person has a belief that Ottawa is the Canadian capital city when the person has the attitude that such a proposition is true. In this sense, self-belief refers to the attitude one might have concerning a proposition about *oneself*. This includes a wide range of propositions about one's past, current, future, or ongoing emotions, attributes, thoughts, and desires. For example, you have the self-belief that you like coffee

---

[25] Kihyeon Kim, "Internalism and Externalism in Epistemology," *American Philosophical Quarterly (Oxford) 30*, no. 4 (1993): 303–16.

[26] Again, there are two approaches to the questions of epistemic justification: one approach views justification as a property of the belief while the other approach thinks of justification as a virtue of the believer. This thesis takes the formal approach and focuses only on the property that makes a certain belief justified.

[27] Eric Schwitzgebel, "A Phenomenal, Dispositional Account of Belief*," Noûs (Bloomington, Indiana) 36, no*. 2 (2002): 249–75.

[28] The relation between a belief and a mental state is controversial. Some philosophers such as Schwitzgebel (2002) argue that mental state must constitute some irreducible beliefs while others argue that beliefs can stand apart from other mental states of mind since there are other information-bearing mental states (like perceptual experience or sub-personal cognitive states) that fall short of being classified as belief. This thesis adopts Schwitzgebel's perspective that every mental state must contain some basic or irreducible beliefs.

when you have the attitude that the proposition 'I like coffee' is true.[29] In this thesis, I will focus merely on a narrower set of self-beliefs, which are the beliefs someone may have about their personality traits. In psychology, personality traits refer to a consistent and complex pattern of thoughts, feelings, and behaviours that define an individual's unique character over time.[30] A person has a self-belief about a certain personality trait (e.g., conscientiousness) when she has the attitude that at least some of the propositions related to this trait (e.g., I care about deadlines) are correct descriptions of herself.

The philosophical inquiry about the epistemic justification of our self-beliefs about our personality traits is often overlooked within the epistemological literature.[31] One potential reason for this is that the question of epistemic justification of self-beliefs about personality traits was not considered epistemologically relevant or meaningfully distinct from the question of epistemic justification of our beliefs about the external world or our beliefs about our current mental states. However, I maintain a different perspective on this matter. I think there are some features of our self-beliefs about our personality traits that make the question of their epistemic justification distinct from other kinds of beliefs. In other words, I think the features that are specific to our self-beliefs about our personality traits give us reason to examine the epistemic justification of such self-beliefs in isolation

---

[29] It is controversial whether beliefs need to be conscious to be classified as beliefs. Some philosophers such as Dennett (1991) argue that consciousness of a belief is not necessary and that a mental state can be classified as belief as long as it corresponds to a certain complex pattern of brain activities. This thesis adopts Dennett's perspective on the relationship between beliefs and consciousness.

[30] Gerald Matthews, Ian J. Deary, and Martha C. Whiteman, *Personality Traits*. Third edition (Cambridge: Cambridge University Press, 2009), 12-5.

[31] As will be discussed in Chapter 2, most of the philosophical literature on self-beliefs is concerned merely with our self-beliefs about our *current* mental state.

from the epistemic justification of our beliefs about the external world and our beliefs about our current mental states. To support my view, I will draw upon empirical studies in personality psychology that highlight a few unique characteristics of our self-beliefs about personality traits. These characteristics have to do with the biases and the self-verification that accompany some of our self-beliefs about our personality traits. I will show how these characteristics make the traditional understanding of epistemic justification less straightforwardly applicable and, therefore, highlight the need for a more nuanced approach to the conditions that make such a set of self-beliefs epistemically justified.

The Thesis Synopsis

My argument sympathizes with epistemic externalism rather than internalism. I argue that to be epistemically justified, a self-belief about a certain personality trait has to go beyond the person's mental resources. More precisely, for a self-belief about a certain personality trait to be epistemically justified, it has to be epistemically *reliable*. I illustrate two approaches to reliabilism: process and virtue reliabilism. While the two approaches argue that a belief is epistemically justified when it is produced by a reliable belief-forming process, process reliabilism cashes out its reliability condition in statistical terms while virtue reliabilism thinks of it as a collection of virtues (epistemic competency).[32] I argue that virtue reliabilism is more suitable to the context of self-beliefs about personality traits than process reliabilism.

---

[32] Ernest Sosa, "Goldman's Reliabilism and Virtue Epistemology," *Philosophical Topics* 29, no. 1/2 (2001): 383–400.

Thus, I situate my proposal to the epistemic justification of self-beliefs about personality traits within the virtue reliabilism framework. According to virtue reliabilism, a belief-forming process that is less affected by biases is more epistemically competent than the one that is more affected by biases. However, to reduce the effect of biases in the context of our self-beliefs about our personality traits, I argue that an epistemically competent self-belief-forming process has to rely on cases where the person manifests the trait in question while having the motivation and the opportunity to do otherwise. In other words, I argue that for a self-belief-forming process to be epistemically more competent, it has to generate self-beliefs from cases where the person manifests the trait in question despite having the motivation and the opportunity to do otherwise. I show how this can potentially reduce the impact of the biases inherent to some of our self-beliefs about personality traits; increase the competency of our self-belief-forming process; and, therefore, makes our self-beliefs about our personality traits more justified.

Chapter One illustrates the debate between internalists and externalists in more detail. Chapter Two illustrates how the epistemic justification of self-beliefs has been discussed in the philosophical literature. Chapter Three summarizes some relevant findings from empirical personality psychology in order to establish three facts about the nature, accuracy, and self-verification of our beliefs about our personality traits. Chapter Four illustrates an obstacle and a proposal. It shows that the features of self-beliefs illustrated in Chapter Three propose a hurdle to our straightforward approaches to justification. Then, it proposes a way out by arguing that a competent self-belief-forming process has to rely on

cases where the person manifests the traits in question while having the motivation and the opportunity to do otherwise.

Chapter One: The Internalism-Externalism Debate

There are many questions to ask about epistemic justification. For example, one might ask about the extent to which the justification of one belief depends on the justification of other beliefs.[33] One may also ask about the extent to which justification is universal or subject to contextual factors.[34] Additionally, one may ask whether epistemic justification is an all-or-none or whether a belief can provisionally have a certain degree of justification.[35] In this thesis, I investigate what would make a certain self-belief more justified. In other words, I ask what criteria, if any, would enhance the epistemic justificatory status of a particular self-belief. Or, put differently, I try to explore the factors

---

[33] This question asks whether the epistemic justification of a certain belief depends on the justificatory status of other related beliefs. There are two dominant views in the literature: coherentism and foundationalism. This thesis is not necessarily committed to either view. However, for more detail on the epistemic justification of systems of beliefs, see, for example, Ernest Sosa in "The Raft and the Pyramid: Coherence Versus Foundations in the Theory of Knowledge." *Midwest Studies in Philosophy* 5, no. 1 (1980): 3–26.

[34] The association of epistemic justification with truth is often seen as indicating its universality. However, alternative perspectives argue that contextual factors can influence the determination of justification, challenging the assumption of a universal link between justification and truth. While not extensively addressed in my thesis, it is worth exploring the works of Fantl and McGrath (2002, 2007) and Schaffer (2006) for a deeper understanding of the debate on the universality of epistemic justification.

[35] While some philosophers argue for limited variation in epistemic justification, contemporary thought tends to view it in terms of degrees, acknowledging that beliefs can have varying levels of justification. For further exploration of this perspective, refer to Jeff Dunn's "Reliability for Degrees of Belief", in *Philosophical Studies 172, no.* 7 (2015): 1929–52. However, for arguments in favor of limiting the degree of beliefs, see Nicholas J. Smith, "Degree of Belief Is Expected Truth Value," In *Cuts and Clouds*, (Oxford: Oxford University Press, 2010), 1- 18; and Peter Klein who argues that justification comes only in degree in his "No Final End in Sight," In *Philosophy and Phenomenological Research 58*, no. 4 (1998): 95-115.

that if present, they would make a self-belief about our personality traits more epistemically justified.

In this sense, I focus on doxastic justification rather than propositional justification.[36] It is important to distinguish between these two approaches: while propositional justification pertains to the grounds a person has for forming a certain belief, doxastic justification delves into what it takes for a belief itself to be justifiably held. [37] [38] In other words, the focus in doxastic justification is on evaluating the justification status of a belief itself rather than examining the reasons that support its formation. To see the difference between the two approaches to justification, consider a person who perceives a green triangular shape in her visual field. Initially, this experience lacks belief or judgment as it is non-propositional. However, once the person entertains the propositional claim that there is a green triangular shape in her visual field, she obtains propositional justification. This is to say that she has a *reason* to believe that there is a green triangular shape in her visual field, and this is despite whether she actually holds that belief. However, by accepting that propositional claim conveyed in her perception and forming the doxastic attitude that there is a green triangular shape in her visual field, she attains doxastic justification. [39] Her belief is said to *be justifiably held* as it passes a specific justification

---

[36] Contemporary views in justification make a sharp distinction between propositional and doxastic justification. See Luis Oliveira and Silva Paul, *Propositional and Doxastic Justification: New Essays on Their Nature and Significance*. (New York, NY: Routledge, 2022).

[37] Some epistemologists suggest a third kind, for example, see Clayton Littlejohn, *Justification and the Truth-Connection*, (Cambridge: Cambridge University Press, 2012) 5-23.

[38] Oliveira and Paul, *Propositional and Doxastic Justification 5-8*.

[39] Philosophers hold different perspectives on the relationship between propositional and doxastic justification. The dominant position posits propositional justification as more fundamental and seeks to reduce doxastic justification to propositional justification. This position is assumed in this example. For more detail on this, see, for example, Jonathan L. Kvanvig's "Theoretical Unity and the Priority of

standard (i.e., being presented in her visual field in this case). In other words, doxastic justification concerns whether the belief meets the standard to be justifiably held whereas propositional justification relates to the reasons someone has for holding a particular belief.

The main question of this thesis is to explore the standards that have to be present for a self-belief about our personality traits to be more epistemically justified. I investigate this question in relation to the standards suggested by the two dominant theories of epistemic justification: internalism and externalism. As many philosophers suggest, the internalism-externalism debate is one of the highly challenging issues related to the question of epistemic justification in contemporary epistemology.[40] In a nutshell, internalists argue that justification is solely dependent on an individual's mental state while externalists contend that justification extends beyond an individual's mental state and incorporates external factors, such as reliable processes and causal connections.[41]

In this chapter, I will first illustrate general remarks regarding the key concepts, different approaches to the distinction, and the dynamics of the debate between internalism and externalism. This will provide a comprehensive overview of the internalism-externalism debate. Then, I will delve into each horn of the debate, offering a more detailed analysis of the arguments in favour and against each position. In terms of internalism, I will

---

Propositional Justification," Chapter Two in *Propositional and Doxastic Justification*, (New York: Routledge, 2022) 27-39. Conversely, another viewpoint suggests doxastic justification as more fundamental and aims to reduce propositional justification to doxastic justification, see for example, Hilary Kornblith, " What Does Logic Have to do With Justified Belief?" Chapter Three in *Propositional and Doxastic Justification*, (New York: Routledge, 2022) 40-67. Alternatively, some argue that neither form of justification can be reduced to the other, see for example, Ram Neta, "Justification Ex Ante and Ex Post" Chapter Four in Propositional and Doxastic Justification, (New York: Routledge, 2022), 70-87.

[40] Kim, "Internalism and Externalism in Epistemology," 304.

[41] Kim, 304–6.

focus on the two main flavours of internalism: accessibilism and mentalism. In terms of externalism, I will focus on process reliabilism and virtue reliabilism. By exploring these various views, we can establish the framework for the subsequent examination of the epistemic justification of our self-beliefs about personality traits.

What Is at Stake?

There are many ways to frame the disagreement between internalists and externalists. Many of these ways assume that the internalism/externalism distinction is mutually exclusive and jointly exhaustive.[42] This means that any given epistemic position can be categorized as either internalist or externalist, where there is no overlap or middle ground between the two. Although this assumption is an open question in the current literature, as contemporary philosophers seek to challenge it by proposals that merge internalism and externalism, it is worth noting that such attempts have been shown to inevitably collapse into one or the other position.[43] Therefore, this thesis accepts the assumption that the question of epistemic justification is mutually exclusive and jointly exhaustive. The justification for accepting this assumption lies in the recognition that reconciling or merging internalism and externalism has been shown to inevitably collapse into either an internalist or an externalist stance.[44]

---

[42] Alston, 20-33.

[43] There have been many attempts to reconcile internalism and externalism in one view. For example, Goldberg (2018) introduces the term 'epistemically proper belief' to describe a belief that satisfies the reliability and responsibility requirements at the same time. However, many philosophers illustrate how such reconciliations fail inevitably. For a detailed argument on how such attempts may fail, see Amia Srinivasan, "Radical Externalism," *The Philosophical Review* 129, no. 3 (2020): 395–431.

[44] Srinivasan (2020) argues that most attempts to reconcile the two positions collapse into an externalist view: a reconciling theory is externalist once it accepts that there is something to the question of

There are many ways to capture the distinction between epistemic internalism and externalism. The following are three ways in which the internalism-externalism distinction can be framed. One way to frame the internalism-externalism distinction is to view it in terms of supervenience relations.[45] In this sense, internalism would be the thesis that facts about epistemic justification supervene upon one's mental state. This is to say that, for internalists, every change in the justificatory status of a belief must be accompanied by a change in the person's mental resources. In other words, internalists would argue that there can be no change in the facts about the justification of a belief without a change in the person's mental state. In contrast, externalism, in this sense, would be the thesis that facts about epistemic justification do not supervene upon facts about one's mental state. In other words, externalism would accept that facts about one's mental state do not exhaust facts about justification because there could be an alteration in the justificatory status of a belief without necessarily a corresponding change in the individual's mental resources. Put differently, for externalists, the doxastic justification of a belief can be modified without a necessarily concurrent transformation in one's mental state.

The internalist idea that facts about epistemic justification supervene upon one's mental state has an important implication. As Feldman and Conee put it, it implies that "mental duplicates in different possible worlds have the same attitudes justified for them."[46]

---

justification other than or in addition to one's mental resources. Based on this, this thesis accepts the assumption that the question of epistemic justification is mutually exclusive and jointly exhaustive.

[45] Richard Feldman and Earl Conee, "Evidence," In *Epistemology*, (Oxford: Oxford University Press, 2008), 83-104.

[46] Richard Feldman and Earl Conee, "Internalism Defended." *American Philosophical Quarterly* (Oxford) 38, no. 1 (2001): 2.

This is to say that if two individuals have exactly the same mental state in two different possible worlds, then their beliefs have the exact same level of justification. In other words, this implies that because the justification of a belief is determined solely by the person's mental state, two beliefs are justified to the same extent as long as they are presented by identical mental states. Externalists obviously refuse this implication, and as we will see later, the idea of the supervenience of justification will underline how each side responds to the famous thought experiments in the literature (e.g., the new evil demon problem).

Another way to frame the internalism-externalism distinction is to view it in terms of necessary and sufficient conditions for knowledge.[47] As put by Feldman and Conee, "Internalism and externalism are views about which states, events and conditions can contribute to epistemic justification the sort of justification that, in sufficient strength, is a necessary condition for knowledge."[48] In this sense, internalism would be the thesis that internal factors are both necessary and sufficient for epistemic justification that is needed for knowledge. This is to say that the justification needed for knowledge is obtained as long as certain internal factors (e.g., having the relevant evidence)[49] are satisfied. In contrast, externalism, in this sense, would be *either* the thesis that internal factors are necessary but not enough for justification;[50] or the thesis that internal factors are neither necessary nor

---

[47] Feldman and Conee, "Internalism Defended." 1-4.

[48] Feldman and Conee, 1.

[49] The question of what exactly counts as evidence is an open debate in the internalist literature. For our purposes, I will use Feldman's and Conee's conception where evidence is the sum of the propositions that indicate truth to a person (see, Feldman and Conee, 2001: 10–18).

[50] An example of this version of externalism can be seen in earlier versions of Goldman's causal theory. See, for example, Alvin I. Goldman "Strong and Weak Justification," *Philosophical Perspectives 2 (1988):* 51–69.

enough for epistemic justification.[51] In both cases, externalists would argue that the justification needed for knowledge is obtained through factors outside the person's mental state. Or as Srinivasan puts it: "What we might call a pure externalism says that the satisfaction of the externalist condition (e.g., reliability, safety) is both necessary and sufficient for justification."[52]

Additionally, the internalism-externalism distinction can be framed based on each theory's emphasis on the truth-conduciveness of justification.[53] While both theories agree that truth-conduciveness is essential to epistemic justification, they disagree on the weight they assign to it. In this sense, externalism would be the thesis that assigns more weight to the truth-conduciveness of epistemic justification.[54] For example, in his defense of externalism, Littlejohn argues that "the conditions that justify [a] belief have to be sufficiently indicative of the truth or make it sufficiently likely that the belief is true."[55] In other words, externalists like Littlejohn share the tendency to evaluate justification solely based on its truth-conducive features such as its ability to lead to truth.

However, internalism, in this sense would be the thesis that highlights features in justification other than its ability to lead to truth.[56] For example, internalists may argue that justification possesses an intrinsic value, independent of its ability to lead to truth.[57] In this

---

[51] An example of this version of externalism can be seen in Amia Srinivasan, "Radical Externalism." *The Philosophical Review* 129, no. 3 (2020): 395–431

[52] Srinivasan, "Radical Externalism." *404*

[53] Littlejohn, *Justification and the Truth-Connection*, 1-44.

[54] Littlejohn, 26-8.

[55] Ibid, 26.

[56] Madison, B. J. C. "Epistemic Value and the New Evil Demon." *Pacific Philosophical Quarterly 98, no.* 1 (2017): 89–107.

[57] Madison, "Epistemic Value and the New Evil Demon." 89–107.

regard, internalism would evaluate the justificatory status of a belief based on criteria other than (or at least in addition to) its truth-directed outcomes. These criteria may include considerations such as whether the belief reflects certain intellectual virtues such as integrity and open-mindedness.[58] For instance, in his remarks on the relationship between justification and truth, Montmarquet suggests that "persons who desire the truth would desire to have these traits [i.e., intellectual virtues], but they are not simply traits that are truth-conducive in any straightforward way. In this sense, epistemic virtues related to justification are valuable even when they are not truth conducive."[59] In other words, Montmarquet suggests that although certain intellectual traits may be associated with truth, their value stems simply from being virtues of the mind rather than solely from their capacity to lead to truth. The three approaches mentioned above are by no means an exhaustive list of the ways in which the internalism-externalism distinction can be presented.[60] The goal was to establish a theoretical framework in which concrete theories from internalism and externalism can be discussed. In the next paragraph, I will illustrate the overall dynamic in the internalism-externalism debate. Then, I will shift focus toward examining specific theories within the internalist and externalist camps.

The debate between internalism and externalism in epistemology has generated significant discourse. In most of the cases, the dynamics of these argument employs a

---

[58] Ibid

[59] James A. Montmarquet, *Epistemic Virtue and Doxastic Responsibility*, (Lanham, Md: Rowman & Littlefield Publishers, 1993), 25.

[60] For more approaches to the distinction between internalism and externalism, see Kim, "Internalism and Externalism in Epistemology," 303–4.

variety of thought experiments.[61] Understanding the dynamics of these thought experiments will become useful later when illustrating and situating the thesis' thought experiment in Chapter Four. The dynamic goes as follows: One side usually proposes a scenario and shows that our intuition about that scenario challenges the theory of the rival side. One example of these hypothetical scenarios includes the reliable clairvoyant, Norman, who can reliably tell where the president is without relying on any clear indications of some kind (e.g., the newspaper) and, more importantly, with no clue about how he has such true beliefs.[62] Does the fact that Norman has always true beliefs about where the president is make his belief justified, given that he has no idea how/why he has such true beliefs? Internalists answer no. They state that we have the intuition that although Norman has reliable true beliefs, his beliefs are not justified and that this is precisely because he has no relevant mental states whatsoever to how/why he has such true beliefs.[63] In reply, externalists have either to provide an error account in which they argue that Norman's beliefs are, in fact, justified and, as a result, explain why we have such an intuition (i.e., showing why we feel that his beliefs are unjustified)[64] or to reconstruct the scenario within their position (i.e., showing why the case of Norman does not, in fact, undermine the external position in the first place).[65]

---

[61] Richard Pettigrew, "Radical Epistemology, Structural Explanations, and Epistemic Weaponry." *Philosophical Studies 179,* no. 1 (2022): 289–304.

[62] Laurence Bonjour, "Externalist Theories of Empirical Knowledge." *Midwest Studies in Philosophy* 5, no. 1 (1980): 53–74.

[63] For a full analysis of the clairvoyance scenario and other dominant thought experiments of this kind, see Bonjour, "Externalist Theories of Empirical Knowledge." 50–84.

[64] For an example of a philosopher who took this approach, see Amia Srinivasan, "Radical Externalism," 400-12.

[65] For an example of a philosopher who took this approach, see Alvin Goldman. "A Causal Theory of Knowing." *The Journal of Philosophy* 64, no. 12 (1967): 357–72.

Theories of Epistemic Internalism

Internalist theories share the core idea that factors determining the epistemic justification of a single belief are internal to the believer.[66] However, they disagree on where exactly to draw the line between what counts as 'internal' and 'external'. The general assumption is that the property of being 'internal' is relevant to the human cognitive system.[67] In this sense, internalists assume that "X is internal to a cognitive agent S from an epistemic point of view if and only if x is something that happens within the cognitive system of S."[68] Nevertheless since we are not aware of all the things that happen within our cognitive system, internalists had to specify the amount of awareness/reflection that the person needs to have about a belief for that belief to be justified.[69]

Thus, internalism comes in two flavours: accessibilism (access internalism) and mentalism. Access internalism (accessibilism) is the view that highlights our ability to access, at least, some of the reasons that justify our belief.[70] Accessibilism places emphasis on the idea that the epistemic justification of a belief is contingent upon the individual

---

[66] Feldman and Conee, "Internalism Defended," 1–18.

[67] It is almost a consensus view among internalists that the property of being internal should be interpreted in relevance to human cognition (see, for example, Fumerton, 1995, 2006; Bonjour, 1976, 1980, 1985; Bergmann, 2006).

[68] Kim, "Internalism and Externalism in Epistemology," 305.

[69] Most epistemologists interpret epistemic internalism as a theory of propositional justification (e.g., BonJour, 1985). In this sense, epistemic internalism pertains to an individual's justification to believe, withhold, or disbelieve certain propositions. However, as stated earlier, I focus only on doxastic justification, which views justification as a property of the belief itself. Interestingly, some internalists frame their position in terms of doxastic justification (e.g., Conee & Feldman, 2001; McCain, 2016), and this is why I focus mainly on their version of internalism.

[70] Within the literature, accessibilists engage in a debate regarding the analysis of accessibilism, focusing on whether it should be understood in relation to only one's access to the justifying reasons for holding a belief or in relation to one's access to the belief itself as well. In this thesis, I will focus on Feldman's and Conee's view, where accessibilism is mainly concerned with having some access to the reasons for the belief. For more detail, see Feldman and Conee, "Internalism Defended," 1–34.

having a particular type of access to the factors that support that belief. In other words, for access internalists, a belief is justified (or subject to be justified) if a person has or can have some access to her basis for holding that belief. As put by Feldman and Conee "What we shall call 'accessibilism' holds that the epistemic justification of a person's belief is determined by things to which the person has some special sort of access."[71]

There are different ways to characterize the concept of 'access' in access internalism.[72] For example, BonJour argues that having access is about having "suitable awareness."[73] However, Audi suggests that the notion of access can be achieved through "introspection", where introspection means "looking within."[74] Some argue that the access must be "direct" in a sense that is not mediated by any other state of affairs.[75] The main idea is that the person must have some form of ability to access, at least, some of the reasons that support that belief in order for the belief to be justified. In short, accessibilism posits that a belief's justification depends to a great extent on the individual's cognitive capacity to access and reflect upon the belief in question.

In this sense, there are two kinds of accessibilism: strict and less strict accessibilism. Strict accessibilists argue that for a belief to be justified (or subject to the question of epistemic justification), a person must be aware of *most of* the reasons related to the belief

---

[71] Feldman and Conee, "Internalism Defended," 2.
[72] Ibid
[73] Laurence BonJour, "The Dialectic of Foundationalism and Coherentism." In *The Blackwell Guide to Epistemology*, 118.
[74] Robert Audi, *Epistemology: A Contemporary Introduction to Theory of Knowledge*, (London: Routledge, 1998) 7.
[75] Barnard Russell, "Knowledge by acquaintance and knowledge by description," 192–3.

in question.[76] For example, a belief that Cairo is Egypt's capital city is justified, according to the strict accessibilists, when the person is aware of most of the reasons that ground that belief (e.g., being aware of when and how exactly she learned that fact. . .etc.). Strict accessibilism is usually criticized for being almost impossible to attain.[77] Therefore, less strict accessibilists argue that a belief can be justified (or subject to the question of epistemic justification) as long as the person has some general or somehow-relevant awareness of at least some of the relevant reasons.[78] According to the less strict accessibilists, the belief that Cairo is Egypt's capital city can be justified as long as the person can, in principle at least, demonstrate some awareness of some of her grounds for that belief (e.g., remembering learning this fact at some point in school). In short, for a belief to be justified, according to accessibilism, the person must have *some* mental awareness of the reasons that ground that belief.

Indeed, accessibilism is, sometimes, criticized for being too demanding for its awareness requirement on beliefs.[79] Mentalism drops this requirement. Mentalism argues that a belief can be justified (or subject to be justified) even if a person has no access to her basis for holding that belief. This is because mentalists argue that a belief can be subject to epistemic justification as long as the belief stems from a mental state of some kind.[80] This

---

[76] Ali Hasan, "Access Internalism, Mentalism, and Reliabilism" In *A Critical Introduction to the Epistemology of Perception*, (United Kingdom: Bloomsbury Publishing Plc, 2017) 117-9.

[77] Hasan, "Access Internalism, Mentalism, and Reliabilism", 119-22.

[78] For a defence of a less strict accessibilism, see Roderick Chisholm in *The Foundations of Knowing*. (Minneapolis: University of Minnesota Press, 1982). Also, see Matthias Steup, in his *An Introduction to Contemporary Epistemology* (Upper Saddle River, NJ: Prentice Hall, 1996).

[79] Hasan, 122-5.

[80] Ibid

notion is nicely captured by Feldman and Conee who argue that "a mentalist theory may assert that justification is determined entirely by occurrent mental factors, or by dispositional ones as well. As long as the things that are said to contribute to justification are in the person's mind, the view qualifies as a version of mentalism."[81] In other words, mentalism holds that the internal factors that determine epistemic justification do not require explicit reflection or awareness. Thus, the belief that Cairo is Egypt's capital city can be justified, according to mentalists, as long as this belief stems somehow from any mental state (e.g., feeling like she somehow believes that Cairo is Egypt's capital city).[82]

Theories of Epistemic Externalism

Externalism is usually defined as the theory that rejects the internalists' demand of limiting justification to the mental state of the individual.[83] This rejection is expressed through various (and sometimes very different) proposals.[84] BonJour identifies two core components of an externalist position on justification.[85] The first component of an external theory, according to BonJour, is that it argues that the person's mental resources are not enough for epistemic justification. This is to say that, for an externalist, no matter how

---

[81] Ibid, 2.

[82] Some might question the relevance of the internalist theories in the context of self-beliefs, given the fact that we do not have access to every self-belief we have about ourselves. However, as we will see in Chapter Three, although people might lack access to some of their beliefs about their personality traits or their reasons for having such beliefs, psychologists suggest that people do have the ability to cite at least some relevant reasons to back up their self-beliefs. Thus, although we are not always aware of every self-belief we have about our personality traits or of every reason for having such a belief, we have the ability to cite at least some relevant reasons. Thus, internalist theories are still relevant in this sense.

[83] Armstrong, D. M., *A Materialist Theory of the Mind*, (London: Routledge, 1993) 20-6.

[84] Armstrong, *A Materialist Theory of the Mind*, 23-9.

[85] Laurence Bonjour, "Externalist Theories of Empirical Knowledge." *Midwest Studies in Philosophy* 5, no. 1 (1980): 53–74.

strong or 'vivid' the mental states are, at the end of the day, justification requires something *more than* mental resources. This also means that, for an externalist, the epistemic justification of a belief may or may not be accompanied by a first-person insight. In this sense, externalist proposals can be further classified into two categories: theories that argue that mental resources are necessary but not enough (e.g., Sosa's virtue reliabilism) and theories that argue that mental resources are neither necessary nor enough (e.g., Goldman's process reliabilism).[86]

The second component of an external theory, according to BonJour, is that it has to offer an account of what else is needed as an alternative (or in addition) to the internal factors.[87] This is where externalists' theories vary. These externalists' proposals, however, share (to some extent) the underlying assumption that what is needed as an alternative (or in addition) to the internal factors must be truth conducive. Thus, although externalists' accounts seem to offer different proposals for epistemic justification, their accounts focus primarily on what has been thought of as more likely to lead to true beliefs.[88] An example of an externalist's proposal can be seen in Goldman's early causal and, then, later reliable version of externalism (both theories significantly influenced the development of externalist perspectives in epistemology). For example, in his earlier theory, Goldman emphasized the causal history of a certain belief, arguing that a belief is

---

[86] For more examples of externalists who argue that internal factors are necessary but not enough for justification see Sosa, 1991. However, for more examples of externalists who argue that internal factors are neither necessary nor sufficient see Bergmann, 2006.
[87] Bonjour, "Externalist Theories of Empirical Knowledge," 53–74.
[88] Armstrong, *A Materialist Theory of the Mind*, 23-9.

justified if and only if it is causally linked to the truth of the matter.[89] In this sense, according to Goldman's causal account, the belief that Cairo is Egypt's capital city is justified as long as the causal history of that belief is linked to the fact of the matter. Later, Goldman modified his view and proposed another theory that goes under the name of process reliabilism. In his later theory, Goldman argues that a belief is justified if and only if it was produced by a reliable belief-producing process.[90] In this sense, according to Goldman's process reliabilism, the belief that Cairo is Egypt's capital city is justified as long as the 'machine' that produces that belief is reliable.

Within the context of externalist theories, I will exclusively concentrate on reliabilism. This is because, unlike other externalist theories, reliabilism focuses primarily on the belief-forming process, evaluating the methods or "cognitive mechanisms" involved in forming beliefs.[91] This emphasis makes reliabilism particularly relevant to self-beliefs as it focuses on addressing the underlying cognitive processes that generate self-beliefs. Thus, in what follows, I focus on two forms of reliabilism: process reliabilism and virtue reliabilism. Exploring these two forms of reliabilism (in addition to the already presented versions of internalism) will give us the framework we need to see the challenges that will arise regarding the justification of our self-beliefs about our personality traits.

---

[89] Goldman. "A Causal Theory of Knowing." 357–72.

[90] Alvin Goldman, "What Is Justified Belief?" In *Reliabilism and Contemporary Epistemology* (New York: Oxford University Press, 2012) 89-93.

[91] Goldman, "What Is Justified Belief?" 92.

Goldman's process reliabilism emphasizes the reliability of the cognitive processes involved in belief formation.[92] He argues that a belief is justified if and only if it has been produced by a reliable cognitive process, regardless of whether the person is aware of the process or the belief's justifying reasons.[93] Goldman views the reliability condition of the belief-forming process as both necessary and sufficient for justification.[94] This means that, for Goldman, the reliability of the belief-forming process is by itself enough while the internal factors are neither necessary nor sufficient for justification. He writes: "I do not even assume that when a belief is justified there is something possessed by the believer which can be called justification."[95]

Goldman and other process reliabilists think of reliability in terms of how likely a certain belief-forming process is to produce overall true beliefs compared to false beliefs.[96] If a certain belief-forming process has the overall tendency to produce more true beliefs than false beliefs, then that belief-forming process is considered reliable.[97] Interestingly, in order to explain what is meant by the term 'reliable', Armstrong, who is a defender of epistemic externalism, draws a parallel between a thermometer that consistently indicates temperature accurately and a belief-forming process that consistently produces true beliefs.[98] This analogy suggests that like the thermometer that is said to be reliable when it produces more accurate and true readings, the belief-forming process is also said to be

---

[92] Goldman, 91-4.
[93] Ibid
[94] Ibid
[95] Ibid, 90.
[96] Armstrong, D. M. *Belief, Truth and Knowledge*, (Cambridge: University Press, 1973) 166-8.
[97] Armstrong, *Belief, Truth and Knowledge*, 166.
[98] Armstrong, 166.

reliable when it tends to generate more true beliefs than false ones.[99] This analogy also suggests that like a reliable thermometer, a reliable belief-forming process is one that tends to consistently produce outcomes that align with reality without requiring any additional or internal access of some kind.[100] However, it's worth noting that neither Goldman nor other defenders of process reliabilism think of reliabilism solely in terms of the actual frequency of true versus false belief production or sustenance to determine reliability.[101] Instead, they see reliability as the overall tendency of a method itself to mostly produce or sustain true beliefs over false ones. In other words, Goldman thought of the reliability of the belief-forming process as a propensity: He argues that a belief is considered justified if it arises from a process that has a high propensity to produce true beliefs when given true beliefs as inputs.[102]

Following the development of process reliabilism, Ernest Sosa's virtue reliabilism, which is also a branch of virtue epistemology, shares the general assumptions with process reliabilism that justification depends on factors beyond internal resources.[103] However, unlike Goldman's process reliabilism, where internal resources are neither necessary nor sufficient, Sosa's virtue reliabilism argues that internal factors are necessary but not enough

---

[99] The infallibility thesis is not widely supported among epistemologists from both sides. The prevailing view is that a belief can be justified even if it turns out to be false. For more on this see, Jack C. Lyons, "What We Talk About When We Talk About Epistemic Justification," *Inquiry* (Oslo) 59, no. 7-8 (2016): 867–88.

[100] Armstrong, 166.

[101] Jack C. Lyons, "Should Reliabilists Be Worried About Demon Worlds?" *Philosophy and Phenomenological Research* 86, no. 1 (2013): 1–40.

[102] Goldman, "What Is Justified Belief?" 101.

[103] While there are multiple advocates of virtue reliabilism, such as Greco and Pritchard, this thesis specifically focuses on Sosa's version. For further exploration of various accounts of virtue reliabilism, refer to Greco and Kelp's book, "*Virtue-Theoretic Epistemology: New Methods and Approaches*" (Cambridge University Press, 2020, pp. 42-66).

for justification.[104] As Sosa puts it: "Suppose epistemic justification should indeed depend only on such internal matters independent of the subject's contingent emplacement. How, more positively and fully, should we then conceive of such justification . . . But this way of understanding epistemic justification takes us *only* so far and comes up short"[105] Sosa articulates the limitations of an understanding of epistemic justification that relies *solely* on internal factors independent of the subject's specific circumstances. He questions how such a limited understanding can fully and comprehensively account for justification. Thus, Sosa does not eliminate internal factors but, rather, argues that relying solely on them leads to a narrow perspective on epistemic justification.

In general, virtue reliabilism cashes out the notion of epistemic justification in terms of epistemic competence.[106] Sosa writes: "A belief or judgment is epistemically competent iff it is formed well enough in epistemic respects. This is how we shall think of what so often goes under the label epistemic justification."[107] However, to illustrate the idea of epistemic competency, Sosa uses the example of an archer aiming at a target.[108] Sosa argues while it makes sense to determine the archer's success in terms of how accurately he hits the bullseye, accuracy alone is not enough for a skillful shot. To demonstrate his point, Sosa asks us to imagine a situation where the shot is initially thrown off by the wind but is corrected by a second gust, eventually hitting the

---

[104] Sosa, "Goldman's Reliabilism and Virtue Epistemology," 383–400.
[105] Sosa, 388.
[106] Ernest Sosa, "How Competence Matters In Epistemology." *Philosophical Perspectives* 24, no. 1 (2010): 465–75.
[107] Ernest Sosa, *Epistemology*, (Princeton: Princeton University Press, 2017) 105.
[108] Sosa, *Epistemology*, 107-14.

bullseye.[109] Sosa argues that although this shot is accurate, it doesn't fully capture the archer's skill and competence; therefore, he concludes that this shot is less superior to a one that truly reflects the archer's mastery.[110] In other words, in his example, Sosa wants to highlight the importance of considering not only accuracy (as in the case of Armstrong's thermometer) but also the overall aptness and competence when evaluating epistemic justification.[111] Sosa concludes: "A performance X attains its aim P not just through the fact that P but through its bringing it about that P. The case of knowledge is just the special case in which the performance is cognitive or doxastic."[112] This is to say that like how competency matters in the example of the archer, competence also matters in the context of knowledge: it is not sufficient for a belief to simply align with the truth, but rather, the cognitive process itself must actively contribute to the realization of that truth.

To some extent, the concept of epistemic competence may resemble that of a reliable belief-forming process. This is because the exercise and the identification of epistemic competence in virtue reliabilism can be thought of (if not totally reduced to) as an instance of a reliable process. For example, one might argue that, at the end of the day, what makes the archer's skills in Sosa's example competent is the fact that they increase the likelihood that the archer makes a successful shot. In this view, epistemic competence and statistical reliability are closely intertwined (if not even totally reduced to each other):

---

[109] Sosa, 107-14.
[110] Sosa, 108-12.
[111] Sosa, "How Competence Matters In Epistemology." 465–75.
[112] Sosa, *Epistemology*, 113.

epistemic competence is competence precisely because it leads (or increases the propensity that s process leads) to truth.

However, there are, at least, two reasons why virtue reliabilism is not reduced to process reliabilism.[113] First, unlike process reliabilism which focuses solely on the statistical propensity of the belief-forming process, virtue reliabilism has a more expanded focus.[114] This can be seen in the notion of epistemic competencies itself. Sosa argues that the notion of epistemic competencies can apply to a person or to the belief-forming process.[115] For example, Sosa states that "competencies are dispositions of an agent to perform well."[116] In this context, he refers to the abilities, skills, and virtues that make the person more likely to arrive at the truth. At the same time, Sosa relativizes his notion of epistemic competence to a belief-forming process. For example, he states that good eyesight including good rods, cones, and good color vision are examples of perceptual competence.[117] Thus, the notion of epistemic competency applies to things that cannot be reduced to a belief-forming process of some kind such as skills, habits, and people's perceptual faculties.

Second, unlike process reliabilism which focuses merely on the truth-conduciveness of the belief-forming process, virtue reliabilism assigns a special value to the mere "manifestation of an epistemic competence"[118] For example, in the case of the

---

[113] Sosa, "Goldman's Reliabilism and Virtue Epistemology," 388-90.
[114] Ibid
[115] Sosa, "How Competence Matters In Epistemology." 465–75.
[116] Sosa, "465.
[117] Sosa, 467.
[118] Sosa, *Epistemology*, 113-4.

archer, Sosa argues that while the archer's shot is considered favorable if it has not fallen short, the specific act or shot may still be deemed excellent based on its own merits or technical proficiency.[119] This is because, according to Sosa, the focus of a virtue reliabilist should be given to the archer's skill and ability to execute the shot correctly instead of the overall accurate shorts or outcomes. He writes: "A shot might manifest an archer's competence without its accuracy doing so. The shot with the two intervening gusts is a case in point. How does that shot manifest the archer's competence? By having at the moment of release an angle, direction, and speed that would take it to the bullseye in relevantly normal conditions."[120]

The fact that the mere manifestation of epistemic competence holds an intrinsic value makes virtue reliabilism irreducible to process reliabilism. In his criticism of process reliabilism, Sosa writes: ". . . [A] generic understanding of "reliabilism" as a view that makes the epistemic normative status of a belief dependent [only] on its likelihood of being true given some feature of it. . . is laughably inadequate."[121] This is to say that, in Sosa's virtue reliabilism, the focus of any reliable belief-forming process extends beyond producing true beliefs into exercising the virtuous intellectual faculties that contribute to epistemic success.[122] Thus, while Goldman's notion of reliability is cashed out entirely in terms of its truth-conduciveness, for Sosa, the mere manifestation of epistemic competence holds an intrinsic value, irrespective of whether the beliefs formed through such

---

[119] Sosa, 113-4.
[120] Sosa, 113.
[121] Sosa, 40.
[122] Ibid, 44-9.

competence turn out to be true. For Sosa, this value stems from the idea that engaging in intellectually virtuous practices are valuable endeavours besides their truth-conducive value.[123] In short, because virtue reliabilism underscores the inherent significance of epistemic competence as a manifestation of a reliable belief-forming process, it cannot be reduced to the mere statistical propensity that is exhibited in process reliabilism.

This chapter has provided a comprehensive overview of the internalism-externalism debate by illustrating general remarks regarding key concepts, different approaches to the distinction, and the dynamics of the debate as found in the literature. In terms of epistemic internalism, the discussion focused on the two main flavors of internalism, namely accessibilism and mentalism. In terms of epistemic externalism, the discussion focused process reliabilism and virtue reliabilism. Through exploring these various perspectives, we have established the framework in which we will examine the epistemic justification of our self-beliefs concerning personality traits. The next chapter will provide a literature review of how self-beliefs have been discussed in the philosophical literature.

Chapter Two: Self-Beliefs and Epistemic Justification

Self-belief refers to the body of propositional beliefs we have about ourselves.[124] This includes the propositions we have about our perceptions, attitudes, and judgments regarding our abilities, worth, and potential. These propositions that we have about

---

[123] Ibid

[124] Annalisa Coliva, *The Self and Self-Knowledge*, (Oxford: Oxford University Press, 2012) 1-15.

ourselves greatly influence our reactions, emotions, and behaviours, shaping our self-image and guiding our choices. In a sense, our self-beliefs serve as the foundation upon which we construct our understanding of ourselves and our interactions with the world around us. In this chapter, I focus on self-beliefs. I will first distinguish the concept of self-belief from seemingly similar concepts such as self-awareness. Then, I will illustrate how several dominant theories that focus on self-beliefs approach the question of epistemic justification. As the reader will notice, these theories focus merely on the beliefs we have about our current mental states. This is basically how the topic of self-beliefs is discussed in the literature. However, I discuss the ways in which these approaches to epistemic justification could be relevant to the context of self-beliefs about personality traits. I also discuss some reasons why we need a fresh eye when it comes to context of self-beliefs about personality traits. This will pave the road for future chapters, where I illustrate some intricacies related to the epistemic justification of self-beliefs about our personality traits.

What Do We Mean by Self-Belief?

It is fair to say that the terms 'self-knowledge' and 'self-beliefs' are often employed interchangeably in the literature to describe the body of beliefs we have about ourselves. Philosophers might occasionally utilize the term 'self-knowledge' to catch a particular sense of certainty that typically accompanies certain types of self-beliefs, particularly those concerning our present mental states (such as the belief that I am currently experiencing pain).[125] However, this usage lacks systematic consistency and fails to establish a clear

---

[125] Minh Nguyen, "What Good Is Self-Knowledge?" in *Journal of Philosophical Research* (2015), 138.

distinction. Consequently, I will use both terms interchangeably. Thus, while I will primarily use the term 'self-belief' for consistency, it also encompasses the broader notion of self-knowledge as discussed in the literature.

Another point that is worth raising with respect to how 'self-belief'/'self-knowledge' is used in the literature is that the terms usually include a variety of self-ascriptions that do not necessarily take the self as its object. For example, in defining self-knowledge, Coliva states that "self-knowledge [is] characteristically expressed in self-ascriptions of mental states with propositional content, like 'I hope my flight is on time' and 'I was relieved to find my luggage waiting for me'."[126] While the subject in these examples is the self, the objects are about things other than the self (e.g., the flight and the luggage). Although this usage is quite common in literature, I will only stick with the self-beliefs that take the self as their object (e.g., I believe I *am* shy).

However, self-belief in this sense has also to be distinguished from self-awareness, or what is usually referred to as the knowledge of the self.[127] It is one thing to know the kind of things we are, yet it is another thing to know the content of a particular thought we have about ourselves.[128] Inquiries related to self-awareness include (but are not limited to) questions about the nature of the self; the awareness of its existence; the ability to distinguish the self from others; and the persistence of the self through time.[129] In this sense, the discussion of self-awareness takes a step back and asks more fundamental questions

---

[126] Coliva, *The Self and Self-Knowledge*, 2.
[127] Quassim Cassam, "*Self-Knowledge for Humans*," (Oxford: Oxford University Press, 2014), 2-5.
[128] Cassam, "*Self-Knowledge for Humans*," 5-7.
[129] Cassam, 5-13.

compared to the discussion of self-beliefs: we may explore the propositional content of a particular thought we have about ourselves (self-belief), yet we may go further and explore the very ability to bear thoughts and desires (self-awareness).

Nevertheless, the relationship between self-belief and self-awareness is controversial. Some philosophers argue that self-awareness is necessary and sufficient for all kinds of knowledge one might have about oneself.[130] For example, in his famous *cogito* argument, Descartes uses his awareness of his ongoing thinking activity as a necessary and sufficient foundation for various kinds of knowledge (including knowledge and beliefs about himself).[131] However, some philosophers argue that although some level of self-awareness might be necessary to make self-beliefs possible, self-awareness cannot, by itself, give us insights into the propositional content of our beliefs about ourselves.[132] For example, in his reply to Descartes, Gassendi argues that knowing that we are a 'thinking thing' gives us no substantial knowledge about the content of our thoughts.[133] For Gassendi, this is simply because someone might be aware that she undergoes a certain mental state (e.g., thinking) without necessarily knowing what she is thinking about.[134] These kinds of arguments led so many philosophers to separate discussions of self-beliefs from that of self-awareness.[135] Thus, it is fair to draw the relation as follows: While self-awareness focuses

---

[130] Ibid.

[131] For a typical example of this view, see René Descartes, "*Discourse on Method,*" (New York: Liberal Arts Press, 1950) 25-8. However, for a contemporary supporter of this view see Declan Smithies, *The Epistemic Role of Consciousness*, (New York, NY: Oxford University Press, 2019) 1-55.

[132] Thomas M Lennon, "The Battle of the Gods and Giants: The Legacies of Descartes and Gassendi, (Princeton: Princeton University Press, 1993) 235.

[133] Lennon, "The Battle of the Gods and Giants," 238-9.

[134] Ibid.

[135] Ted Parent, "Self-Knowledge and Externalism About Empty Concepts," in *Analytic Philosophy* (2015), 158–68.

merely on our awareness of the presence of thoughts of some kind, self-beliefs focus on our relationship to the propositional content of such thoughts. This thesis will only concentrate on self-beliefs, acknowledging that this inherently assumes the presence of, at least, some degree of self-awareness.

What Kinds of Questions Do Philosophers Ask about Self-Beliefs?

We ordinarily distinguish the body of beliefs we have about ourselves from the beliefs we have about the world: It is one thing to state that it is cold outside or that I believe it is cold outside, yet it is another thing to state that I feel cold myself. One philosophical question with respect to self-beliefs is to explain whether this distinction reflects real asymmetries between the two categories of beliefs.[136] Does the observation that we ordinarily discern self-beliefs from other kinds of beliefs reflect a real difference between our beliefs about ourselves and our beliefs about the world? Do these *prima facie* asymmetries withstand the rigorous philosophical analysis? In short, one philosophical inquiry with respect to self-beliefs is 1) to cash out these *prima facie* asymmetries and 2) to argue whether they withstand the rigorous analysis. I will present three examples of the ways these *prima facie* asymmetries were cashed out and discussed in the literature, and this is not by any means an exhaustive list.

---

[136] Richard Moran. *Authority and Estrangement: an Essay on Self-Knowledge*. (Princeton: Princeton University Press, 2001) 4-36.

One way to cash out the *prima facie* asymmetry is to think of it in terms of access.[137] In this sense, the claim that 'it is cold outside' would belong to the set of beliefs about the external world, to which we think we have equal access. However, the claim that 'I feel cold myself' would belong to the set of beliefs about us, to which we do *not* think we have equal access. Then, the philosophical challenge in this sense would be to determine whether this asymmetry withstands the rigorous philosophical analysis: Do I really have privileged access to the thought that I am feeling good? Can others have access to this thought? Does my access to this thought (the first-person access) have the same metaphysical or epistemological status as others' access (the third-person access)?

Another way to cash out the *prima facie* asymmetry is to think of it in terms of some epistemic features such as epistemic certainty and security.[138] For example, Smithies argues that "there is an epistemic asymmetry to be drawn between first-personal and third-personal ways of knowing what we believe. Each of us has some way of knowing what we ourselves believe that is peculiar in the sense that it is different from any of our ways of knowing what others believe."[139] This way of cashing out the *prima facie* asymmetry also stems from ordinary observation: we often feel more certain and secure regarding the beliefs we have about ourselves compared to the beliefs we have about the world. For example, it

---

[137] For more details on the accounts that cash out the asymmetries between our beliefs about the world and beliefs about ourselves in terms of access, I recommend Brie Gertler, "*Privileged Access: Philosophical Accounts of Self-Knowledge,*" (London: Routledge, 2016) 73-6.

[138] For more details on the accounts that cash out the asymmetries between our beliefs about the world and beliefs about ourselves in terms of some epistemic features, I recommend Declan Smithies, "Belief and Self-Knowledge: Lessons from Moore's Paradox: Belief and Self-Knowledge." *Philosophical Issues 26*, no. 1 (2016): 393–421.

[139] Smithies, "Belief and Self-Knowledge," 393.

makes sense to be more confident in stating that you feel warm yourself than stating that the weather or someone else is warm. Although all of these statements are fallible in the sense that they can be false, this does not change the observation that we do feel more *certain* and *secure* in relation to the beliefs we have about ourselves compared to the beliefs we have about the world.[140] [141]Again, the philosophical challenge in this sense would be to illustrate whether this asymmetry withstands the rigorous philosophical analysis: Does the feeling of certainty or security that accompanies some of our self-beliefs carry a distinctive epistemic status? What potential implications would it have on relevant epistemic concepts, such as the knowledge acquisition processes or the notion of epistemic justification?

Furthermore, another way to cash out the *prima facie* asymmetry is to think of it in terms of *authority*.[142] Likewise, this also stems from our ordinary observation: we feel that, among everyone else, we are in the *best position* to determine the truth or the falsity of statements about ourselves. For example, if I tell my friend that I feel cold, then it would

---

[140] Some philosophers advocate for unqualified infallibility with respect to self-beliefs (e.g., Descartes). However, although such claims were accepted in the past, almost no contemporary philosophers accept infallibility in its *unqualified* forms with respect to self-beliefs. Contemporary theories highlight that the sense of certainty we experience regarding our self-knowledge can coexist with the recognition that a significant portion of these beliefs about ourselves may be inaccurate. However, some philosophers defend a qualified infallibility thesis with respect to some of our beliefs about some of our occurrent mental states. For example, Chisholm states that some forms of self-knowledge are "self-presenting" in the sense that one can't be entirely wrong about them (136). For Chisholm, this qualified level of infallibility grounds the epistemic security we feel regarding some of the beliefs we have about ourselves compared to the beliefs we have about the world. The question of the infallibility of self-beliefs (both in its qualified and unqualified form) is not within my focus. However, for more detail see Thomas A Russman, "Roderick Chisholm: Self and Other," In *The Review of Metaphysics (1979),* 130-6.

[141] Interestingly, some philosophers (e.g., Wright) argue that the mere possibility of error with respect to self-beliefs is required to make such beliefs *genuine* in some sense. For more details on this interesting view, see Crispin Wright, Barry C. Smith, and Cynthia Macdonald, *Knowing Our Own Minds*, (Oxford: Clarendon, 1998) 13-46.

[142] For more details on the accounts that cash out the asymmetries between our beliefs about the world and beliefs about ourselves in terms of some epistemic features, I recommend Moran, *Authority and Estrangement: An Essay on Self-Knowledge*, 66-152.

be weird (if not even offensive) for my friend to reply: No! You do not. The same reply, however, wouldn't be as weird (nor as offensive) if I was making a statement about the weather or about someone else instead. Likewise, the philosophical challenge in this sense would be to explore the seeming authority that we have with respect to, at least, some of the beliefs we have about ourselves: Where does this sense of authority come from? What does it mean? Does it have any moral or social implications? And how does it impact our understanding of the personal subjective experience and its role in shaping the person's worldview?[143]

Whether any form of the *prima facie* asymmetries between self-belief and other kinds of beliefs holds water is beyond the focus of this thesis. What is relevant for our purposes is to provide a literature review of how epistemic justification was approached by, at least, some theories of self-knowledge.[144] Thus, in the next section, I will only focus on three examples of accounts that engaged with the attempt to cash out the *prima facie* asymmetry in an epistemic sense: the acquaintance theory, the inner sense theory, and the rationalist theory of self-knowledge.[145] I will first demonstrate these accounts and, then, discuss how they approach the question of epistemic justification of self-knowledge. The take-home point from the following section will be that each theory constructs its position

---

[143] While these questions were initially presented as independent, it is important to note that philosophers who endorse any of the aforementioned asymmetry theses often have arguments regarding other asymmetries as well. For more detail on, I recommend Brie Gertler, *Self-Knowledge*. (New York: Routledge, 2011) 2-23.

[144] The three theories that I will discuss in this section use the term self-knowledge instead of self-belief quite consistently. Thus, to avoid any potential confusion, I will use the term 'self-knowledge' instead of 'self-belief' in this section in order to align with the terminology used by the discussed theories.

[145] Of course, this does *not* mean that these accounts *only* engaged with the prima facie asymmetries in the epistemic sense.

on epistemic justification based on its view on the nature of self-knowledge. Theories that took an empirical approach to self-knowledge (i.e., the acquaintance theory, the inner sense theory) took a passive approach and view epistemic justification as either internal (i.e., the acquaintance theory) or external (i.e., the inner sense theory). However, theories that took a rational approach to self-knowledge (i.e., the rationalist's theory of self-knowledge) highlight our active role in our beliefs about ourselves and view epistemic justification in a complex way (i.e., locally external and globally internal). It is important to stress that all these theories focus mainly on the beliefs we have about current mental states (except the rationalists who employ wider examples of self-beliefs). The final section of this chapter will discuss the extent to which these views are relevant to the beliefs we have about our personality traits.

Three Theories of Self-Knowledge

In the following brief presentation, I will illustrate three theories of self-knowledge: the acquaintance theory, the inner sense theory, and the rationalist theory. Each of these theories offers a unique perspective into the nature of our beliefs about our mental states and how such beliefs are justified. Thus, through an examination of these theories, we can develop a more profound understanding of how each theory's conception of the nature of self-knowledge influences its stance on the epistemic justification of our beliefs concerning our occurrent mental states.

1. **The Acquaintance Theory of Self-Knowledge**

There are many versions of the acquaintance theory, and according to Gertler, these versions share the key assumptions that the epistemic *prima facie* asymmetry indicates a real metaphysical distinction between the beliefs we have about, at least, some of our current mental states and other kinds of beliefs.[146] [147] This is to say that, according to acquaintance theories, the difference between our beliefs about some of our current mental states and other kinds of beliefs is an actual difference in *kind* rather than in *degree*: our beliefs about, at least, some of our current mental states are *direct* while beliefs about the world are necessarily mediated to us.[148]

The most famous version of the acquaintance theory is associated with Bertrand Russell.[149] In his earlier philosophical work, Russell writes: "We shall say that we have acquaintance with anything of which we are directly aware, without the intermediary of any process of inference….Thus in the presence of my table, I am acquainted with the sense data that make up the appearance of my table …. My knowledge of the table as a physical object is not direct knowledge. Such as it is, it is obtained through acquaintance with the sense-data that make up the appearance of the table…. We have seen that it is possible,

---

[146] The reason why this distinction is considered metaphysical rather than solely epistemological, according to Gertler, is due to the fact that, according to the acquaintance theory, the asymmetry goes beyond mere differences in certainty or epistemic security. They suggest that there is something inherent in the nature of reality itself that accounts for the observed asymmetry in our beliefs. According to the acquaintance theory, this asymmetry implies that there is a fundamental disparity in the ontological status of our beliefs concerning our current mental states in comparison to other types of beliefs. In other words, according to the acquaintance theory, the asymmetry reflects a distinction in the underlying fabric of existence rather than in how we know or how we conceptualize things.

[147] Gertler, *Self-Knowledge*, 90-8.

[148] Gertler, 90-8.

[149] Bertrand Russell, "Knowledge by acquaintance and knowledge by description,"191–220.

without absurdity, to doubt whether there is a table at all, whereas it is not possible to doubt the sense-data." [150] [151]This is to say that, according to Russell, the *direct* input of the sense-data that we are acquainted with are direct in the sense that we cannot doubt them while the other inputs are indirect in the sense that we can doubt them.[152]

In terms of our beliefs about some of our current mental states, Russell argues that the high sense of epistemic security that accompanies some of our mental states indicates that at least some of these mental states are direct in the sense that there is no gap between these mental states and us.[153] In other words, he suggests that our grasp of, at least, some of our mental states is immediate and not mediated by any state of affairs. To see what Russell means, take the following example. If after reading a weather report I believed that it was snowing outside, then my belief about the weather is mediated by the weather report. However, if I believe I feel itchy, then, according to Russell, my belief is direct and not mediated by any other state of affairs.

Gertler argues that the Russellian idea of the directness of some mental states has two implications. First, it suggests that such knowledge is indubitable.[154] This is to say that

---

[150] Russell, 191.

[151] It is important to highlight that while the mention of acquaintance often evokes the sense-datum theory, the commitment to acquaintance does not inherently imply a commitment to it.

[152] There is ongoing debate regarding the interpretation of Russell's concept of direct acquaintance. For instance, philosophers like Fumerton argue that direct acquaintance is non-inferential, whereas others like Gertler argue that acquaintance is not a causal relationship.

[153] Russell, 191. This is related to the idea that direct acquaintance is identified with what the thought is about. See Richard Fumerton, "Acquaintance: The Foundation of Knowledge and Thought", in *Acquaintance: New Essays*, (Oxford; Oxford Academic, 2019), 245-459. Fumerton argues that the idea of direct acquaintance should be the basis of all knowledge and thought. He argues that our ability to directly perceive facts enables us to acquire truths without relying on inference from other known propositions.

[154] Gertler, 90-8.

while a skeptic might find something to say about the gap that always separates our perceptual state from the world, the skeptic would find nothing to say about our knowledge of some of our current mental states. Second, it suggests that certain forms of self-knowledge are obtained through introspection, which is a special first-person method of grasping or looking inward.[155] Introspection, in this sense, allows us to grasp and become aware of things we are acquainted with.

Although contemporary acquaintance theorists may not accept all elements of the Russellian picture, they mostly agree that while most of our knowledge of the world comes from perception, at least some forms of self-knowledge (mostly the knowledge of some of our current mental state) are more immediate, direct, and arrived at by introspection.[156] For example, Smithies argues that "just as I can know by introspection that I am in pain when I am, so I can know by introspection that I'm not in pain when I'm not."[157] He argues that if we carefully reflect on a sensation we are experiencing, such as the pain of biting our lips, the pain becomes part of our introspective experience rather than simply a cause of it. Thus, Smithies argues that, unlike perceptual experiences, which are mediated by processes such as light reflecting off an object and reaching the retina, introspective awareness is not similarly mediated.[158] In this sense, introspective awareness is considered metaphysically direct as there is no intermediary between one's awareness of the mental state and the state itself.

---

[155] Gertler, 90-8.
[156] Smithies, "Belief and Self-Knowledge," 393-8.
[157] Smithies, 393.
[158] Ibid

2. The Inner Sense Theory of Self-Knowledge

Unlike the acquaintance theory, the inner sense theory argues that the *prima facie* epistemic security of self-knowledge does not necessarily indicate a real metaphysical distinction between self-knowledge and other kinds of knowledge. Thus, the key assumption for the inner sense theory is that the epistemic *prima facie* asymmetry between our beliefs about our current mental states and other kinds of beliefs only indicates a difference in *degree* rather than an actual difference in *kind*.[159] This is to say that, for inner theories, both our beliefs about ourselves and our beliefs about the world fall along a continuum. Even though certain forms of self-beliefs may possess a greater level of certainty or epistemic security, inner theories suggest that this does not make these beliefs distinct from beliefs about the world.[160]

The inner sense theories' idea that beliefs about our mental states differ from other beliefs in *degree* rather than in *kind* has two implications. First, it rejects the idea that there could be a direct connection between us and some forms of occurrent mental states and posits that all relations are causal.[161] Second, it rejects the idea that there could be a first-person method such as introspection and argues that whatever is called introspection must be essentially similar to perception in that both involve a causal relationship between the mental state and its object.[162] For example, if I see a squirrel through the window, then (assuming my perception is healthy) my visual state is caused by the squirrel's location.

---

[159] Peter Carruthers, "Inner Sense Theories." In *The Opacity of Mind*, (United Kingdom: Oxford University Press, 2011) 192–222.
[160] Carruthers, "Inner Sense Theories."203-12.
[161] Carruthers, 212.
[162] Ibid

Similarly, the inner sense theories argue that if I feel pain, then this feeling is caused by my inner perception of the relevant biological and psychological clues in my body. In other words, the inner sense theories argue that like perception, introspection depends on a causal process that connects us to the world, and, therefore, all forms of introspection are to be cashed out somehow in terms of perception.[163]

The most famous version of the inner sense theory is associated with John Locke.[164] Locke argues that some forms of our beliefs about some of our current mental states seem more secure and certain because they result from rigorous monitoring or a scanning mechanism. For Locke, this scanning is conducted by what he called "our inner faculty".[165] He writes: "This Source of Ideas, every Man has wholly in himself … might properly enough be called internal Sense."[166] For Locke, the inner faculty does not employ a method that is distinctive in kind from other methods. He makes it clear that such a monitoring mechanism is crucially similar to our five senses in the sense that all of these faculties are part of a causal framework.[167] For Locke, like our senses, the inner faculty takes mental states as input and represents them as output, and this allows us to perceive our own mental states and have awareness of our own thoughts, feelings, and experiences.[168]

Contemporary inner sense theories share the emphasis that our beliefs about our current mental states are neither direct (non-causal) nor metaphysically distinct from other

---

[163] Ibid, 220-5.
[164] John Locke, "An Essay Concerning Human Understanding," (London, 1694) II.1.iv.
[165] Locke, "An Essay Concerning Human Understanding," II.1.iv.
[166] Locke, I.1.iv.
[167] Ibid
[168] Locke, I.1.iv

kinds of beliefs. They also accept that introspection is fundamentally similar to perception in that they both depend on a causal process, regardless of the person's awareness of such causal connections.[169] For example, Armstrong argues that our beliefs about our current mental states ought to serve a biological value of permitting a sophisticated response to stimuli, and this is done "by organizing the mere flow of the information that comes from our mental states as inputs and outputs."[170] Armstrong argues that this biological function is harder to be conceived under the acquaintance picture of introspection. He writes: "if introspection is conceived of as 'acquaintance' with mental states. . . it is difficult to see how all it can yield is information of such highly abstract nature about inner causes or potential inner causes. But if introspection as well as perception is conceived of as a mere flow of information or beliefs, then there is no difficulty."[171] Thus, for Armstrong, whatever is thought of as a first-person method is better understood as a complicated form of a causal chain of "self-scanning."[172]

However, if our beliefs about certain mental states are mediated by a causal chain, like our beliefs about the external world, then they become susceptible to skeptical doubts. While inner sense theories do not completely dismiss skepticism in such cases, they acknowledge that beliefs about certain mental states possess a higher *degree* of epistemic certainty. However, to account for this higher level of epistemic certainty, the inner sense theories argue that such certainty comes from the fact that the process of scanning one's

---

[169] Carruthers, 220-9.
[170] Armstrong, *"A Materialist Theory of the Mind,"* 324.
[171] Armstrong, 326.
[172] Ibid, 327-9.

mental state often occurs unconsciously or without the involvement of inferential reasoning.[173] For the inner sense theories, although this does not make some of our beliefs about our mental state indubitable, it explains at least, why such beliefs enjoy a higher level of epistemic certainty.[174]

Before delving into the rationalist theories of self-knowledge, it is important to consider a useful classification. In her book *Self-Knowledge*, Gertler categorizes theories of self-knowledge into two primary perspectives based on their approach to the nature of self-beliefs.[175] These perspectives encompass theories adopting an *empirical* viewpoint and theories embracing a *rational* standpoint. Gertler observes that empirical theories often adopt a naturalistic/reductive stance or, at least, take a passive approach to self-beliefs.[176] In this sense, Gertler argues that both the acquaintance theory and the inner sense theory qualify as examples of this category.[177] This is because whether one is directly acquainted with a certain mental state or caused somehow to have it, one's belief about that mental state, in both cases, would be a mere *reporting* of a phenomenon of some sort.[178] This

---

[173] Armstrong, 326-9.

[174] In essence, both the acquaintance theory and the inner sense theory concur that our beliefs about some of our mental states do not rely on inference. However, they diverge in their reasoning. Acquaintance theorists argue that introspection grants direct awareness of mental states, obviating the need for inference from the available evidence. On the other hand, inner sense theorists propose that the causal process of scanning one's mental state may occur without conscious awareness or evidence, eliminating the necessity of relying on inferential reasoning. Consequently, both theories agree that some forms of our beliefs about some of our mental states are distinct epistemically from other types of beliefs, at least in virtue of the higher epistemic certainty that accompanies such beliefs. For more comparison between the two theories, see Gertler, *Self-Knowledge*, 87-165.

[175] Gertler, *Self-Knowledge,* 165-9.

[176] Gertler, 166.

[177] For the scope of our discussion, empirical theories of self-knowledge would specifically pertain to the acquaintance theory and the inner sense theory.

[178] Gertler, 165-9.

categorization will become clearer when contrasted with rational theories, which criticize the passive approach and advocate, instead, for accounting for our sense of agency in many of our self-beliefs.

3. Rationalist Theories of Self-knowledge

The rationalist theories of self-knowledge disagree with the acquaintance and the inner sense theories' basic approach to self-knowledge. The rationalist theories argue that framing the debate merely around specific epistemic features such as epistemic security misses what really is distinctive about self-knowledge.[179] For rationalist theorists, self-knowledge is distinct by virtue of its "active state of normative commitment."[180] This normative commitment includes (but is not limited to) the self's ability to actively reflect, shape, make decisions, and take responsibility for its own self-beliefs.[181] Thus, rationalist theorists advocate for a normative perspective on self-knowledge, prioritizing the role our self-beliefs play in promoting rationality and maintaining our sense of agency.

Although there are different versions of rationalism, they all share the criticism of the passivity of empiricist approaches and argue that we play an active role in many of the beliefs we have about ourselves.[182] This active role becomes apparent when considering, for example, the difference between asking someone about the weather and asking them whether they want a cup of coffee. Rationalists argue that when people are asked about

---

[179] Tyler Burge and Christopher Peacocke, "Our Entitlement to Self-Knowledge." *Proceedings of the Aristotelian Society 96, no. 1* (1996): 91–116.

[180] Christine Korsgaard, "The Activity of Reason" in *Proceedings and Addresses of the American Philosophical Association* (2009), 39.

[181] Burge and Peacocke, "Our Entitlement to Self-Knowledge," 91–116.

[182] Burge and Peacocke, 91–116.

external matters, people simply report their observations; however, when they are asked about their own beliefs about themselves (such as whether they want a cup of coffee), people make decisions rather than merely reporting their observations. In this sense, the rationalist theory diverges from the inner sense and acquaintance accounts as it highlights the significance of control and deliberation in relation to, at least, some forms of our self-beliefs.

Although rationalists stress the active role we play in our beliefs about ourselves, rationalists do not think all self-beliefs are subject to deliberation.[183] Rationalists agree that some forms of our beliefs about ourselves can be beyond our control, and they have no problem accepting that these beliefs are obtained through self-introspection or self-scanning.[184] Thus, rationalist theorists usually adopt a hybrid view of the nature of self-knowledge, and this makes them compatible, in some respect, with the inner sense or the acquaintance account.[185]

Rationalist theories of self-knowledge are generally inspired by Immanuel Kant, who emphasizes the role of agency in self-knowledge.[186] Kant talks about the active role of the self in gaining knowledge and constituting us as rational individuals.[187] He writes: "Consciousness of itself is the simple representation of the I, and if all of the manifolds in the subject were given self-actively through that alone, then the inner intuition would be

---

[183] Ibid.

[184] Ibid.

[185] Gertler, 165-9.

[186] Dina Emundts, "Kant's Ideal of Self-Knowledge," In *Self-Knowledge*, (New York: Oxford University Press, 2017), 185-6.

[187] Emundts, "Kant's Ideal of Self-Knowledge," 185-8.

intellectual."[188] The main point here for Kant is that our consciousness would not enjoy its current intellectual status if there is no activity of the self. For Kant, this activity of the self is precisely a rational and intellectual one. He argues that we are the way we are as intellectual and rational beings not by virtue of this simple and passive presentation of the I, but, rather by virtue of the active role we play in determining many of the things that are internal to us.[189] In this sense, many forms of self-knowledge are distinctive because they include self-constitution. In other words, many forms of self-knowledge are distinctive because the self plays an active role in constituting and shaping its beliefs, intentions, and the criteria it uses to make decisions and take actions.[190] In short, self-knowledge in this sense entails recognizing oneself as a rational, self-governing doer of actions rather than just a passive recipient or observer of information.

Contemporary rationalist theorists share the Kantian emphasis on the role of agency in self-knowledge. For example, Moran writes: "The phenomena of self-knowledge, not to mention the wider spectrum of asymmetries between the first- and third-persons, are themselves based as much in asymmetries of responsibility and commitment as they are in differences in capacities, or in cognitive access."[191] This is to say that, according to Moran, the *prima facie* asymmetry of self-knowledge is not to be solely rooted in differences in cognitive capacities or access, but rather in asymmetries

---

[188] Immanuel Kant, Paul Guyer, and Allen W. Wood. "*Critique of Pure Reason,*" (Cambridge: Cambridge University Press, 1998) B 60, p.189.
[189] Emundts, 186-9.
[190] Ibid
[191] Richard Moran, "*Authority and Estrangement: an Essay on Self-Knowledge,*" (Princeton: Princeton University Press, 2001), 64.

with respect to our responsibility and commitment. Moran summarizes what he thought of as asymmetries with respect to our responsibility and commitment in the following three points.[192] First, he argues that some self-beliefs are subject to revision through deliberate thought and consideration in a way that is not applicable to our beliefs about the world.[193] Secondly, he argues that at least some self-beliefs are obtained in a non-observational manner, whereas our beliefs about the world often rely on observational evidence.[194] Lastly, he argues that we perceive some of our self-beliefs as being particularly transparent to us, as we can directly comprehend them by reflecting on their objects, whereas this level of transparency is not typically found in our beliefs about the world.[195] In the next section, I illustrate how each of the three theories of self-knowledge perceives the question of epistemic justification based on its way of viewing the nature of self-beliefs.

Epistemic Justification in Theories of Self-Knowledge

Let's take the case of Adam, who wants to know what anniversary gift to get for his wife. To decide which gift to buy, Adam may take any of the following steps. He may simply ask his wife what she wants; observe her behaviours and attitudes in response to some advertisements; or pay attention to the things she talks more about. In all these cases,

---

[192] Moran, "*Authority and Estrangement*," 64-75.
[193] Moran, 70-2.
[194] Ibid
[195] Ibid

if Adam forms the belief that his wife wants X as a gift, then whether his belief is justified can be straightforwardly evaluated by any of the internalist or externalist theories discussed in Chapter One.[196] For example, for internalists, Adam's belief that his wife wants X is justified as long as his belief stems from a relevant mental state of some kind (e.g., feeling like he saw something like X in his wife's wish list).[197] However, for an externalist, Adam's belief that his wife wants X is justified as long as the belief is produced by a reliable belief-forming process such as asking her trustworthy best friend who is reliably right about what his wife wants.[198]

However, if Adam wants to know what he *himself* would want as an anniversary gift, he need not take any of the steps mentioned above. He need not observe his behaviour, take notes of the advertisements that grab his attention, nor pay attention to things he talks more about. He simply (and perhaps directly) just knows what he wants as a gift. If he forms the belief that he wants Y as a gift, then there is a good chance that such a belief is just correct. The interesting question at this point is the following: in virtue of what would Adam's belief that he wants Y be justified?

This question is particularly pressing given that internalist and externalist theories of justification cannot be easily applicable, or at least, as straightforwardly as in the case of

---

[196] Or according to any theory of epistemic justification in general.

[197] Not all internalists would agree that this is what it takes for Adam's belief to be justified. For example, a strict internal accessibilist would require Adam to have a clear and vivid awareness of most of the relevant mental states. However, as I stated in Chapter One, mentalism is more relevant for our purposes. Thus, I only present what a mentalist would say about the justification of Adam's belief.

[198] Again, many things can go wrong and Adam's belief that his wife wants X could end up false. However, as discussed earlier, this does not change the fact that Adam's belief is justified, at least according to those theories of epistemic justification.

his belief that his wife wants X. On one hand, in terms of internalism, it is less clear what kind of reasons/evidence that supports Adam's belief that he wants Y as a gift. As Goldberg nicely puts it: "The problem with accounting for the epistemic status of first-person opinions is [that] one's judgements regarding one's own standing attitudes are often formed without any evidence at all, and so by extension cannot be represented as the conclusion of a process of acceptable reasoning."[199] Thus, while Goldberg would acknowledge that individuals, like Adam, may engage in deep thinking and gather some relevant evidence when forming their beliefs, Goldberg argues that "though these types of reasoning-based cases do exist, they are far from the standard way in which we form first-person opinions."[200] This is because, for Goldberg, ".. in most cases, I just know what I believe, without having to think about it or consider my recent behaviours."[201]

On the other hand, in terms of externalism, it is also less clear what truth tracking process is involved in Adam's belief that he wants Y as a gift. As Goldberg puts it, "the epistemic warrant for first-person opinions is not seen as deriving from some kind of 'truth-tracking' manner in which first-person judgments are formed. Rather, what the facts themselves are is itself defeasibly determined by the judgments that the person is disposed to form about these facts: one's first-person opinions are held to stand by default."[202] This is to say that because self-beliefs are initially granted a default standing, it is unclear, according to Goldberg, how the justification of such beliefs can be attributed to a 'truth-

---

[199] Sanford C. Goldberg "The Psychology and Epistemology of Self-Knowledge," *Synthese* (Dordrecht) 118, no. 2 (1999): 166

[200] Goldberg "The Psychology and Epistemology of Self-Knowledge," 166.

[201] Goldberg, 166.

[202] Goldberg, 173.

tracking' process of some kind. In other words, rather than being a product of a 'truth-tracking' process, Adam is granted a default standing on his belief that he wants Y as a gift, and this is why he is correct that he wants Y as a gift.

But what would the inner sense, the acquaintance, and the rationalist theories say about the epistemic justification of Adam's belief that he wants Y as a gift? In other words, in virtue of what would such a belief be justified according to each one of these theories? There is no clear-cut answer to this, and obviously one can defend a variety of different combinations. However, Gertler argues that although "there are no strict entailment relations between these competing conceptions of justification and the three theories of self-knowledge outlined above. . . the acquaintance theory is strongly associated with epistemic internalism, and inner sense theorists are generally epistemic externalists. Rationalists tend to adopt a combination of internalism and externalism."[203] This is to say that although no definitive position has been explicitly taken by any of these theories, Gertler suggests that it makes sense to think of the acquaintance theory as aligned with epistemic internalism; the inner sense theorists as leaning towards epistemic externalism; and the rationalists, on the other hand, as a complex combination of the internalist and externalist perspectives. In the following few paragraphs, I briefly illustrate why Gertler made these associations. From my illustration, it would be clear that I agree with Gertler's association between the acquaintance theory and epistemic internalism as well as her association between the inner sense theorists and externalism. However, I use the case of

---

[203] Gertler, *Self-Knowledge,* 12.

Adam to raise some concerns about the rationalist's approach to the question of epistemic justification as illustrated by Gertler.

In terms of the acquaintance theory, there are many key elements in the acquaintance theory that make it plausible to think of the acquaintance theory as aligned with epistemic internalism. Both the idea that some mental states can be direct and the idea that we have a special method by which we access such states make the reliance on internal mental resources sufficient for justification. Why would anyone look outside their mental realm 1) if they take their internal resources to be sufficient and 2) if looking outside opens the door for skepticism, and, therefore, less epistemic security? In fact, some philosophers (e.g., Chisholm) argue that being in a certain mental state is sufficient to justify the person's belief that he is in that mental state.[204] [205] In other words, according to this view, the person needs no further justification for his belief that he is in a certain mental state beyond the very experience of being in that mental state.[206] Thus, Adam's belief that he wants Y as a gift is justified, in this sense, as long as he experiences the very mental state of wanting Y as a gift.

---

[204] Russman, "Roderick Chisholm: Self and Other," 130-6.

[205] Chisholm advocates for one version of the qualified infallibility thesis about mental states (see footnote number 142). The idea is that if a person is qualifiedly infallible about a certain mental state, then the very fact that she is in that mental state is sufficient to justify her belief that she is in that mental state. If I feel cold, then, according to the qualified infallibility thesis, this is enough to justify my belief that I feel cold. In other words, I need no further justification for my belief that I feel cold besides the fact that I am in the mental state of feeling cold. Proponents of the qualified infallibility thesis propose some condition that makes a certain state qualifiedly infallible, and this is usually proposed by foundationalists (philosophers who believe that there are some basic beliefs that make the foundation for other non-basic beliefs) against the question of infinite regress. For more detail on this see Russman, 130-6.

[206] Russman, "Roderick Chisholm: Self and Other," 130-6.

In contrast, there are many elements in the inner sense theory that make it plausible to think of it as aligned with epistemic externalism. Both the idea that our self-beliefs do not differ in kind from other beliefs and the idea that all forms of introspection can be reduced into forms of causal perception make the reliance on mental resources alone uncompelling. In other words, why would anyone stick with their mental resources if they take these resources to be no different from any other kind of beliefs? In this sense, Adam's belief that he wants Y is justified by the same standards that justify his belief that his wife wants X. For inner sense theorists, the fact that we cannot 'identify' the truth-tracking process that is involved in the production of Adam's belief does not mean there is none. If we would take reliabilism as an example of externalism in this case, then Adam's belief that he wants Y is justified as long as the truth of the belief can be tracked by a reliable belief-forming process, where the reliable belief-forming process, in this case, would be Adam's inner sense perception.

However, things get trickier when it comes to the rationalist theories of self-knowledge. Gertler suggests that the prevalent forms of rationalism appear to be "locally externalists" yet "globally internalists."[207] This is because, as suggested earlier, the key element of the rationalist thesis is that we play an active role with respect to, at least, some of our self-beliefs while, at the same time, they accept that this role can be out of the question with respect to other self-beliefs. Gertler argues that it makes sense to pose that rationalists are externalists with respect to the self-beliefs that are *out* of our reach as agents and that this externalism is 'local' since these beliefs are mostly concerned with biological

---

[207] Gertler, 15.

sensations.[208] At the same time, she argues that it also makes sense to pose that rationalists are internalists with respect to the self-beliefs that are *within* our reach as agents and that this internalism is 'global' since these beliefs are, according to rationalists, necessary for our rational nature.[209] As Moran puts it: "A sensation of pain or vertigo really is something one may be passively subject to, something that just happens. . . . Beliefs and other attitudes, on the other hand, are stances of the person to which the demand for justification is internal. And the demand for justification internal to the attitudes involves a sense of agency and authority that is fundamentally different from the various forms of direction or control one may be able to exercise over some mind or another."[210]

But what would the rationalists say about the justification of Adam's belief that he wants Y as a gift? First, a rationalist has to determine whether Adam's belief belongs to the beliefs that are *within* or *outside* the reach of his agency. If Adam's belief belongs to the set of beliefs that is *within* the reach of his agency, then rationalists would take this belief to be justified according to internalist standards.[211] However, if Adam's belief belongs to the set of beliefs that is *outside* the reach of his agency, then rationalists would take this belief to be justified according to externalist standards. What I take as a problem here is that even if Adam's belief that he wants Y as a gift belongs to the set of beliefs that is *within* the reach of his agency, such belief must consist of (if not totally supervene upon) at least some mental states that are beyond his agency (e.g., basic biological sensations). For

---

[208] Ibid

[209] Ibid

[210] Moran. *Authority and Estrangement*, 114.

[211] This is most likely not only in the weak mentalist sense but in the accessibilists' sense as rationalists highlight the importance of reflection.

example, let's say Adam believes that he wants a watch as an anniversary gift. The mental state that he wants a watch may consist of (or, at least be influenced by) other mental states such as his memories of being bothered by not knowing the time, being irritated by frequently asking others for the time, or being envious of others who have stylish watches. These additional mental states may implicitly contribute to the fact that Adam prioritized wanting a watch over other things (e. g., a book). These additional mental states may not be completely *within* the reach of Adam's agency. For example, Adam's tendency to get irritated about not knowing the time might be connected to basic biological disposition of some kind, his envy might arise from internalizing a societal norm or cultural expectations, and his preference for a watch might be linked to an inherent fascination with status symbols. Therefore, these additional mental states are to be justified externally since they are *outside* the reach of Adam's agency. In other words, even if we take Adam's belief that he wants Y as a gift to be a belief *within* his agency, there are additional mental states that play a role in cementing Adam's belief and these mental states are justified externally since they are *outside* the reach of Adam's agency. This is problematic because it seems that the very mental state that is within Adam's sense of agency somehow emerges from (or was bolstered by) the mental states that are *outside* the reach of his agency. Simply put, it is inconceivable to comprehend how agentive mental states that are internally justified emerge from non-agentive mental states that are externally justified.[212]

---

[212] This is merely a preliminary observation, as a thorough critique of the rationalists' perspective on justification falls outside the scope of this thesis. Indeed, it is also important to direct the reader's attention that the issue I take with the rationalists' perspective on justification stems from my acceptance of the premise that the internalism-externalism distinction is mutually exclusive and jointly exhaustive (as I suggested in Chapter One).

In short, each theory takes its stance on epistemic justification based on its view of the nature of self-beliefs about our current mental states. Empirical theories, such as the acquaintance theory and the inner sense theory, adopt a passive approach to self-knowledge and consider epistemic justification as either internal or external. In contrast, rationalist theories emphasize our active role in our self-beliefs and propose a hybrid view of epistemic justification. The point of this chapter was to provide a brief review of how the question of self-beliefs and epistemic justification was approached in the literature, at least with respect to how the topic of self-belief is mostly discussed (i.e., with respect to the beliefs we have about our current mental states).

However, this is not irrelevant to our beliefs about our personality traits. In one sense, our current mental states can provide some information and insights into our personality traits. For example, a person who consistently experiences feelings of anger and frustration in various situations is more likely to be higher in neuroticism than someone who does not experience these mental states on a regular basis. Similarly, the person who consistently experiences such a mental state is more likely to form the self-belief that she is higher in neuroticism than someone who does not experience these mental states very often. In this sense, our self-beliefs about our current mental states can be thought of as the subjective indicator or the first step in which we interpret or evaluate our self-beliefs about our personality traits. If a person would be asked to what extent they believe they are highly neurotic, it is likely that, among many other things, the person would appeal to their memories of the times they were irritated. In other words, these memories are, to some extent, the sum of the times in which the person had self-beliefs about mental states of

specific content (i.e., irritation). In this sense, the question of the epistemic justification of our self-belief about our current mental state can complement our inquiry into the epistemic justification of our self-belief about our personality traits.

At the same time, our current mental states (and our beliefs about them) might not always convey our personality traits. This is because our current mental states can be influenced by various transient factors such as stress, fatigue, or situational context, which may not necessarily reflect our underlying personality traits. Additionally, situational context can play a significant role in shaping our immediate mental states, causing us to behave differently than we typically would. For example, a typically outgoing and extroverted individual might exhibit introverted behaviors when feeling fatigued or stressed. In such instances, their current mental state may not accurately reflect their actual extroverted personality trait. Thus, while such a person might have a perfectly justified self-belief about her current introverted mental state, this does not mean that the person would necessarily have a justified self-belief about her personality trait. Therefore, although we do get valuable perspectives from examining the epistemic justification of our current mental states, this, by no means, diminishes the need for a fresh eye when it comes to the philosophical inquiry of the epistemic justification regarding our beliefs about personality traits.

In this chapter, my focus has been on self-beliefs. I started by distinguishing the concept of self-belief from seemingly similar notions such as self-awareness. Next, I explored how several prevailing theories that center on self-beliefs tackle the issue of epistemic justification. As we have observed, these theories primarily revolve around the

beliefs we hold about our current mental states. However, I ventured into discussing the potential relevance of these approaches to the context of self-beliefs about personality traits. Furthermore, I highlighted the reasons why we need a fresh perspective when it comes to the justification of our self-beliefs about personality traits. In the next chapter, I will shift focus and illustrate some *empirical* facts from contemporary personality psychology with respect to our self-beliefs about our personality traits. Then, I will discuss how these facts propose some obstacles to some of our theories of epistemic justification.

Chapter Three: Our Beliefs About Our Personality Traits: Facts from Empirical Psychology

In this chapter, I will summarize some relevant findings from empirical personality psychology to establish three facts about our beliefs about our personality traits.[213] The first fact is that people do hold some true beliefs about, at least, some of their personality traits. The second fact is that such self-beliefs are not infallible because many factors can interfere with their accuracy. The third fact is that some of these beliefs can be self-verifying since they both shape and are shaped by our behaviours. Establishing these facts is central to the philosophical argument of this thesis. In the next chapter, I will show

---

[213] It is important to make notes about the replication crisis in social and personality psychology, which is a recurring concern that is relevant to this chapter. The replication crisis refers to the difficulties researchers face in reproducing the results of previously conducted studies, opening the door for some skepticism regarding the robustness of many findings. While this crisis does not undermine the entire field of social and personality psychology, it highlights the need for critical reflection in dealing with such research. For more on the replication crisis in social and personality psychology, see Harold Pashler, "Is the Replicability Crisis Overblown? Three Arguments Examined." *Perspectives on Psychological Science: a Journal of the Association for Psychological Science. 7, no.* 6 (2012): 531–36.

how these facts make some of our theories of epistemic justification less straightforwardly applicable.

What Does Psychology Tell Us about Our Beliefs About Some of Our Personality Traits?

In psychology, personality traits refer to a consistent and complex pattern of thoughts, feelings, and behaviours that define an individual's unique character over time.[214] Psychologists distinguish personality traits, which are generally stable through time, from transient properties, such as temporary mood states.[215] Psychologists study personality traits to understand how they influence people's perceptions, motivations, actions, and their overall interactions with others as well as the world around them.[216] These traits can be broad such as efficiency versus emotionally orientated traits (i.e., the tendency to prioritize industriousness over sympathy or vice versa) or specific such as openness, and neuroticism.[217] The Big Five personality traits, also known as the Five-Factor Model, is one of the widely used frameworks for understanding and organizing the many facets of human personality.[218]

The Big Five represents a structure of a complex hierarchy of attributes that tend to be distributed normally in the population.[219] The five attributes are openness, conscientiousness, extraversion, agreeableness, and neuroticism. Each one of these five

---

[214] Gerald Matthews, Ian J. Deary, and Martha C. Whiteman, *Personality Traits*. Third edition (Cambridge: Cambridge University Press, 2009), 12-5.

[215] Matthews, Deary, and Whiteman. *Personality Traits*, 3.

[216] Matthews, Deary, and Whiteman, 3-5.

[217] Ibid, 12-4.

[218] Thomas A. Widiger, *The Oxford Handbook of the Five-Factor Model of Personality* (New York, NY: Oxford University Press, 2015), 11-9.

[219] Widiger, *The Oxford Handbook*, 16.

traits is an umbrella for other related and more specific traits.[220] For example, people who tend to score high in openness tend also to be curious, creative, and open-minded.[221] They are often interested in art, culture, and new experiences. People who score high in conscientiousness are also typically organized, responsible, and reliable.[222] They have a strong work ethic and are often very goal oriented. People who score high in extraversion also score high in outgoing, sociability, and assertiveness.[223] They enjoy being around others and tend to be very talkative. People who score high in agreeableness also score high in cooperation, compassion, and kindness.[224] They value harmony and often put the needs of others before their own. People who score high in neuroticism tend to register higher anxiety, stress, and mood fluctuation.[225] They are more sensitive to criticism and rejection and prone to negative thinking. In all cases, psychologists agree that people have the tendency to exhibit varying degrees of these traits, with only a small minority of the population falling to the extremes.[226]

Psychologists argue that for a person to hold a self-belief about some of her personality traits, the person has to identify a stable *pattern* in her thoughts, emotions, and behaviours. [227] Merely forming a belief based on a single situation does not constitute a self-belief about personality traits. For example, to have a self-belief that she has a good sense of humor,

---

[220] Widiger, 17-9.
[221] Widiger, 11-9.
[222] Ibid, 21-6.
[223] Ibid, 30-4.
[224] Ibid 36-8
[225] Ibid
[226] Ibid, 16-39.
[227] Simine Vazire and Erika N. Carlson, "Self-Knowledge of Personality: Do People Know Themselves?" In *Social and Personality Psychology Compass* 4, no. 8 (2010): 606.

Fiona must identify a pattern of thoughts, emotions, and behaviour that goes beyond getting everyone to laugh once at a good joke. To have the self-belief that she is the kind of person who has a good sense of humor, Fiona has to identify varieties of situations where she thought, felt, and acted like a person with a good sense of humor.

Psychologists who study the Big Five traits examine how the various aspects of these traits, including individuals' perceptions of them, may influence people's behaviours and their overall well-being. The following sections provide a summary of several relevant empirical studies about the beliefs people have about their personality traits. All these studies focus primarily on the personality traits as classified by the Big Five.[228] By delving into these empirical findings, I aim at establishing three facts about our beliefs about our personality traits. I will illustrate that we do have some true beliefs about our personality traits; however, such beliefs are not perfectly accurate and some of them are inherently self-verifying.

Do We Have True Beliefs About Our Personality Traits?

We have a strong intuition that we know the kind of person we are. In fact, every aspect of our life, from everyday mundane moves to major life-changing decisions, is both a

---

[228] It is worth highlighting that within the field of psychology, the term "self-knowledge" is often employed to describe a wide range of beliefs individuals hold about their personality traits. When using this term, psychologists disregard the nuanced epistemological distinction between belief and knowledge. In fact, much of what psychologists identify as knowledge in this context would be clearly classified by epistemologists as mere beliefs. Thus, while philosophers also use the two terms interchangeably to describe one set of beliefs (our beliefs about our current mental states), this could be attributed to specific features of our beliefs about our current mental states (See footnotes 146). Thus, because much of what is discussed in this chapter would clearly fall under what epistemologists view as beliefs, I would rather maintain consistency with the distinction made by epistemologists and, therefore, utilize the term 'self-beliefs' instead of 'self-knowledge' throughout this discussion.

reflection and an exercise of that intuition. The intuition that people demonstrate some idea about the kind of person they are, is not only documented in many studies[229] but is also used to justify the whole idea of relying on self-report, which is the main and wildly used tool in personality psychology.[230] Self-report is essential for personality research as the vast majority of contemporary personality psychologists use people's self-reports to measure personality traits.[231] Vazire (2006) argues that around seventy percent of studies published in the Journal of Research in Personality in 2003 involved personality assessments that relied *exclusively* on self-reports.[232] The significance of self-report personality research is captured by Swann and Pelham who state that "if we are to question the self-reports of participants, much, if not most, of results of research on the self, becomes suspect."[233]

Nevertheless, many studies have tried to challenge the validity and reliability of people's self-report of their personality traits, and this has been usually done by comparing participants' self-reports to one or more of the following three independent measures.[234] First, by comparing participants' self-reports to the reports of participants' well-acquainted others such as close friends and family members. Second, by comparing participants' self-

---

[229] For example, see Pronin, Emily, Kruger, Savitsky, and Ross, "You Don't Know Me, But I Know You: The Illusion of Asymmetric Insight," and Vazire and Mehl, "Knowing Me, Knowing You: The Accuracy and Unique Predictive Validity of Self-Ratings and Other-Ratings of Daily Behavior," who discussed this spontaneous intuition in detail.

[230] Simine Vazire and Timothy D. Wilson, *Handbook of Self-Knowledge*, (New York, N.Y: Guilford Press, 2012) 130-52.

[231] Arthur A Stone, *The Science of Self-Report: Implications for Research and Practice*, (Mahwah, NJ: Psychology Press, 2000) 3-12.

[232] Simine Vazire, "Informant Reports: A Cheap, Fast, and Easy Method for Personality Assessment." *Journal of Research in Personality* 40, no. 5 (2006): 6-9.

[233] William B. Swann and Pelham Brett, "Who Wants Out When the Going Gets Good? Psychological Investment and Preference for Self-Verifying College Roommates." In *Self and Identity* 1, no. 3 (2002): 228.

[234] Stone, *The Science of Self-Report*, 5.

reports to the participants' behaviours and implicit measures (e.g., nonverbal cues such as gestures, body language, and eye movements). Third, by comparing participants' self-reports to their reports of how they think they are seen by others. The main assumption is that if all these measures correlate with the person's self-report, then there must be some credibility to what the person believes about themselves. For example, if Emy states that she is an introvert, and, if her view of herself fits with how she navigates her life (e.g., the number of friends she has, the likelihood that she laughs loudly in public or the tendency that she looks directly into someone's eyes…etc.) as well as with the views of people who know her closely (e.g., her family and coworkers), then Emy probably has some true beliefs with respect to her view of herself as an introvert.

A number of meta-analyses and individual studies have examined whether self-reports of personality traits do in fact correlate with any of the measures above.[235] Results demonstrate some link between the person's self-report of personality traits and each one of the previously mentioned measures.[236] For example, Vazire and Carlson show a correlation of .40 between the participants' self-reports of their personality traits and the reports of others who were chosen by the participants themselves.[237] Vazire and Carlson

---

[235] In personality psychology, correlation or "R coefficient" is a statistical measure that ranges from -1.00 to 1.00, indicating negative and positive associations between two variables. A correlation close to 0.00 indicates little association, while values nearing -1.00 or 1.00 suggest strong connections, where variables change closely together. A correlation of 0.40, for example, represents a moderate relationship, signifying meaningful but not remarkably strong connections. Falls between 0.20 and 0.39 are considered fair, showing noticeable yet relatively modest associations. Correlations below 0.20 are considered weak, indicating a less discernible relationship. For more detail, see Hofmann, "A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures*," Personality & Social Psychology Bulletin*, 31, no. 10 (n.d.): 1369–85.

[236] Vazire and Carlson, "Self-Knowledge of Personality": 616.

[237] Vazire and Carlson, 616.

suggest that the correlation becomes even higher for traits that are easier to observe (such as extraversion) compared to the traits that are harder to observe (such as neuroticism).[238]

Indeed, expanding on this notion, Biesanz et al. suggested that the strength of the correlation between self-reports and others' reports of personality traits increases to .45 with the duration of acquaintance. [239] Biesanz et al argue that individuals who have known someone for an extended period tend to provide reports that align more closely with the self-report of the individual in question, compared to those who have known the individual for a shorter period of time.

Furthermore, a correlation of .35 was also illustrated between the person's self-report and their actual behaviour and implicit measures. For example, Back et al show that both the participants' self-reports of their Big Five and their implicit measures can predict participants' actual behaviour in a range of social situations.[240] In other words, research tend to support the predictive validity of self-reports in relation to actual behavioral tendencies.[241]

Finally, a correlation of .35 was also demonstrated between people's beliefs about their personality traits and their perceived reputation. For example, Carlson and Furr show that

---

[238] Vazire and Carlson, 616-9.

[239] Jeremy Biesanz, Stephen G West, and Allison and William G Graziano, "Moderators of Self-Other Agreement: Reconsidering Temporal Stability in Personality*." Journal of Personality and Social Psychology 75*, no. 2 (1998): 467–77.

[240] Mitja D. Back and Stefan C Schmukle, and Boris Egloff, "Predicting Actual Behavior From the Explicit and Implicit Self-Concept of Personality," In *Journal of Personality and Social Psychology* 97, no. 3 (2009): 533–48.

[241] Back and Schmukle, and Egloff, "Predicting Actual Behavior From the Explicit and Implicit Self-Concept of Personality," 538–46.

people's self-report on how others see them does in fact correlate with the actual report of others, which means that people do have some ideas about their general reputation.[242]

Collectively, these findings emphasize the fair association between self-reports of personality traits and a range of measures, including reports from others, actual behaviour, implicit measures, and perceived reputation. Although the correlation between self-report and these measures is not perfect, psychologists agree that the association is good enough to establish that people do have some true beliefs about their personality traits.[243] As Vazire and Carlson put it, "people's perceptions of their own personality are certainly more accurate than random guesses would be, but they are substantially far from perfect."[244]

In other words, the take-home point from this section is that the convergence of correlations from diverse sources further supports the idea that people are not clueless about who they are and that at least some of people's beliefs about their personalities are true. Thus, given these correlations, it is not unreasonable to assume that people's beliefs about their personalities are at least somehow tied to reality. The next section discusses the factors that might affect the accuracy of people's beliefs about their personality traits.

Do We Have Accurate Beliefs About Our Personality Traits?

If 'accurate' means perfect or infallible, then the short answer is no. Psychologists agree that although we do have some true beliefs about our personality traits, these beliefs are far

---

[242] Erika N. Carlson and Michael R. Furr, "Evidence of Differential Meta-Accuracy: People Understand the Different Impressions They Make," In *Psychological Science* 20, no. 8 (2009): 1033–39.
[243] Vazire and Wilson, *Handbook of Self-Knowledge*, 130-52.
[244] Vazire and Carlson, 615.

from perfect.[245] As Hansen puts it, "The lenses through which we look [at ourselves] are far from objective, and they can distort, cloud, and colour what we see in ways that, for example, sometimes present us with overly positive self-views and other times present us with overly negative ones."[246] In fact, one of the main questions in the study of personality traits is to examine the factors that affect the accuracy of people's beliefs about their personality traits.[247]

Psychologists identify several biases that may distort people's views of themselves. These biases can be mainly organized into biases related to a general misuse of information (e.g., lacking, misperceiving, or misinterpreting information) and biases driven by certain motivations (e.g., the desire to conform to socially desirable traits).[248] Interestingly, research suggests that the effect of both biases decreases with awareness: People are more likely to adjust their own views about themselves when they realize that their previous views were lacking critical information or were subject to social desirability. [249]

People's beliefs about their personality traits can be affected by biases related to their misuse of available information. One example of this type of bias is the "above-average effect" also known as the "better-than-average effect" or " illusory superiority bias."[250] The "above-average effect" refers to the tendency of individuals to perceive themselves as being

---

[245] Vazire and Wilson, 130-52.

[246] Katherine E. Hansen, "Illusions of Self-knowledge" in *Handbook of Self-Knowledge* (New York, N.Y: Guilford Press, 2012), 345.

[247] Hansen, "Illusions of Self-knowledge", 345-7

[248] Ibid, 345-60.

[249] Mark D. Alicke, "Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives," *Journal of Personality and Social Psychology 49,* no. 6 (1985): 1621–30.

[250] Alicke, "Global Self-Evaluation," 1621–30.

above average or better than the average person on various positive traits or abilities.[251] This cognitive bias can manifest in different domains, including personal attributes. Thus, despite the statistical impossibility that everyone is above average, many individuals consistently think and rate themselves as superior to the average person. The above-average effect is considered a common phenomenon that highlights how people can have self-beliefs that are not only inaccurate but also (collectively) statistically impossible.[252] In this sense, we would have an above-average effect with respect to people's beliefs about their personality traits if a significant number of people in a certain group viewed themselves as higher than the average on the personality traits that are perceived more positively (e.g., conscientiousness).[253]

Many psychologists argue that the "above-average effect" may have less to do with people's motivation than with their misuse of information.[254] For example, Moore and Small (2007) suggest that people have the tendency to feel more confident and overestimate their own abilities relative to others for easy tasks (e.g., a trivia quiz) compared to difficult ones (e.g., a challenging crossword) because when judging themselves in easy tasks, people are more likely to focus on their capability and fail to note that others are also capable; however, when judging themselves in difficult tasks, people are more likely to focus on

---

[251] Alicke, "1621–30.

[252] Ibid.

[253] It's important to note that the above-average effect doesn't mean that everyone necessarily believes they are above average in a certain personality trait, but rather it just highlights the general tendency for individuals to rate themselves more favourably compared to the average when assessing their own personality traits.

[254] See for example, Vazire and Wilson, 348; Alicke, 1621–30; and Don A Moore and Deborah A Small, "Error and Bias in Comparative Judgment: On Being Both Better and Worse Than We Think We Are," *In Journal of Personality and Social Psychology* (2007): 972–89.

their incompetence and fail to recognize that most people are incompetent at these tasks too.[255] Moore and Small (2007) also show that people did correct their self-views when they were confronted with accurate information about how others did (e.g., when they were shown actual statistics).[256]

Indeed, people's beliefs about their personality traits can be affected by biases related to certain desires and motivations that arise in particular contexts.[257] One example is the "self-enhancement phenomenon", wherein individuals tend to inflate their self-views.[258] This phenomenon is often driven by the *desire* to prove, gain, protect, and maintain a sense of self-worth, particularly in contexts where there may be perceived social or environmental threats.[259] For example, consider the case of Jamal, who firmly believes that he possesses an exceptional level of openness to experience. According to the self-enhancement phenomenon, if someone questions or challenges Jamal's perception of his openness, he would be motivated to adopt a more extreme or intensified view of his own openness. For example, rather than considering alternative viewpoints or acknowledging potential limitations, Jamal might feel compelled to defend his self-belief by polarizing his own perspective of his openness. This could be achieved, for instance, by emphasizing his eagerness to explore and highlighting instances where he has actively sought out unfamiliar situations.

---

[255] Moore and Small, "Error and Bias in Comparative Judgment," 972–89.
[256] Moore and Small, ", 979–83.
[257] Vazire and Wilson, 100-9.
[258] Taylor and Brown, "Illusion and Well-Being: a Social Psychological Perspective on Mental Health," *Psychological Bulletin 103*, no. Mar 88 (1988): 193–210.
[259] Taylor and Brown, "Illusion and Well-Being": 193–210.

The self-enhancement phenomenon is not equally demonstrated by everyone nor in every situation.[260] Research links individual differences in self-enhancement to differences in people's "ego-involved domains."[261] This is simply to say that people are more likely to inflate their self-views both in their own as well as others' eyes when the thing in question matters to them. Indeed, Paulhus and John have identified two types of self-enhancement: egoistic and moralistic self-enhancement.[262] Egoistic self-enhancement is observed in individuals who have a stronger desire for power and influence and, therefore, tend to inflate their self-views of certain traits, such as extraversion and conscientiousness. On the other hand, moralistic self-enhancement is observed in individuals who have a higher desire for social approval.[263] Paulhus and John argue that individuals with moralistic self-enhancement are more likely to inflate their self-views with respect to traits like agreeableness and moral uprightness. Similarly, drawing particularly from the relationship between narcissism and aggression, Baumeister, Smart, and Boden argue that people are especially prone to self-enhancement when their favourable views about themselves are threatened.[264] Baumeister, Smart, and Boden argue that threat, in this case, can range from being totally confronted and debunked to implicitly challenged or put into question.[265] These different forms of self-enhancement highlight how people's motivations and desires can shape their self-beliefs, leading them to emphasize positive attributes that align with

---

[260] Taylor and Brown, 193–8.

[261] Vazire and Wilson, 109.

[262] Delroy L. Paulhus and Oliver P. John, "Egoistic and Moralistic Biases in Self-Perception: The Interplay of Self-Deceptive Styles With Basic Traits and Motives." *Journal of Personality* 66, no. 6 (1998): 1025–60.

[263] Paulhus and John, "Egoistic and Moralistic Biases in Self-Perception" 1025–60.

[264] Roy F. Baumeister, Laura Smart, and Joseph M. Boden, "Relation of Threatened Egotism to Violence and Aggression: The Dark Side of High Self-Esteem," *Psychological Review* 103, no. 1 (01, 1996): 5-33.

[265] Baumeister, Smart, and Boden, "Relation of Threatened Egotism " 5-33.

the desired self-image while downplaying or ignoring aspects that do not align with their self-concept.

It is important to note that these biases are not necessarily pathological or maladaptive since they may serve important adaptive functions in certain contexts.[266] One example of adaptive behaviour related to the lack of information is the "self-signaling phenomenon."[267] The main idea is that when people are not sure or lack information about their personality traits, they are more likely to pick the behaviour that is more likely to create a more positive image of themselves (e.g., prosocial behaviours).[268] For example, Savary and Goldsmith argue that when people are uncertain about their level of altruism, they are more likely to donate privately than publicly since private donations are more likely to be perceived by the people themselves as genuine acts of altruism.[269] Nevertheless, obviously, when any bias is taken to extremes or used inappropriately, it can lead to distorted perceptions of oneself and others, and, as a result, hinder personal growth and social relationships. The next section discusses another relevant feature of people's perception of their own personality traits, which is the susceptibility to self-verification.

Are Our Beliefs About Our Personality Traits Self-Verifying?

---

[266] Taylor and Brown, 193–210.

[267] Roland Bénabou and Tirole Jean, "Intrinsic and Extrinsic Motivation," *The Review of Economic Studies* 70, no. 3 (2003): 489–520.

[268] Bénabou and Jean, "Intrinsic and Extrinsic Motivation,": 489–520.

[269] Jennifer Savary and Kelly Goldsmith, "Unobserved Altruism: How Self-Signaling Motivations and Social Benefits Shape Willingness to Donate," *Journal of Experimental Psychology*. Applied 26, no. 3 (2020): 538–50.

The short answer is yes! A person's self-belief shapes their choices, actions, and interactions with others. The self-verification arises when we consider that the choices, actions, and interactions that were shaped by belief 1 at time 1 would be the *evidence* that the person relies on to justify forming belief 2 at time 2. The following example illustrates how self-verification makes self-beliefs prone to be endlessly reinforced. First, let's illustrate a case where a belief is not self-verifying. Dana may look at her window and correctly form the belief that it is raining outside. Dana's belief that it is raining may affect her behaviours (e.g., her belief may make her decide to bring an umbrella or not go out). However, neither Dana's actions nor decisions can affect the state of affairs that originally produced Dana's belief that it is raining outside (i.e., the rain itself). Thus, Dana can always take a fresh look at her window to form a new belief about the weather without necessarily relying on her previous beliefs and actions. In other words, her previous beliefs about the weather may play no role in her new and fresh beliefs. This is a case where beliefs are not self-verifying. Second, let's illustrate how things are different when self-belief is involved. If Dana has the self-belief that she is a hard-working individual, then psychologists argue that it is highly probable that Dana will behave in accordance with her self-belief (e.g., she might pick competitive situations/careers where she would likely feel the urge to work hard). Interestingly, Dana's behaviour would function as evidence for her future self-beliefs. This is to say that she will use the fact that she has worked hard in certain situations to back up her future self-belief that she is a hard-working person. This suggests that self-beliefs can get endlessly reinforced because the person's self-beliefs affect her actions, and then, these actions affect future self-beliefs. In other words, the very state of affairs that

originally produced Dana's belief that she is a hard-working individual is continuously shaped by her self-beliefs and actions.

Psychologists discuss multiple ways in which our self-beliefs shape and get shaped by our behaviours. One way for this is the "self-verification phenomenon."[270] Psychologists argue that people have a fundamental need to maintain a stable and coherent sense of self, and, therefore, people feel the constant urge to behave in accordance with their stable preexisting self-views.[271] One important aspect of the self-verification phenomenon is that, unlike the self-enhancement phenomenon discussed earlier, people who self-verify are not necessarily motivated to be seen in a positive light by others. Rather, people who self-verify are seeking to maintain their own views of themselves even if this means being seen in a negative light. Thus, the self-verification phenomenon can be either positive or negative.[272] For example, if Suzan believes that she is low in conscientiousness, then although she does not think this is a positive thing, Suzan might find herself acting more in a way that verifies her self-belief.

Psychologists provide three ways in which someone may self-verify their preexisting self-beliefs, and these ways do not necessarily require people's awareness.[273] First, a person may verify her preexisting self-beliefs by selecting the environment that

---

[270] William B. Swann, "*Handbook of Theories of Social Psychology*," (London: SAGE Publications Ltd, 2012): 2-33.

[271] Swann, "*Handbook of Theories of Social Psychology*," 2-33.

[272] Swann, 2-33.

[273] William B. Swann, Richard M. Wenzlaff, Douglas S. Krull, and Brett W. Pelham, "Allure of Negative Feedback: Self-Verification Strivings Among Depressed Persons," *Journal of Abnormal Psychology* (1965) 101, no. 2 (1992): 293–306.

would most likely give the feedback that confirms the person's preexisting belief.[274] For example, to verify her belief that she is low in conscientiousness, Suzan would (unconsciously) pick situations where she would likely have a harder time completing tasks (e.g., a noisy place). Second, a person may verify her self-beliefs by evoking others to give feedback that confirms the person's preexisting self-beliefs.[275] For example, Suzan might directly ask her partner if he noticed how disorganized she is. Third, a person may verify her self-beliefs by giving more credibility to whatever fits with their preexisting self-beliefs.[276] For example, Suzan may give more weight to a comment that confirms her procrastination than a comment that disconfirms it.

In this chapter, I summarized relevant findings from empirical personality psychology to establish three facts about our beliefs about our personality traits. The first fact was that people were not completely clueless about who they were, as they held some true beliefs about, at least, some aspects of their personality traits. The second fact indicated that such self-beliefs were not infallible, as various factors could interfere with their accuracy. The third fact highlighted the self-verifying nature of some of these beliefs, as the previous self-beliefs and behaviors can endlessly enforce new ones. Establishing these facts will be central to the philosophical argument of this thesis. In the next chapter, I will argue that because of these facts, it is less straightforward how we can apply some of our theories of epistemic justification. Then, I will emphasize that although virtue reliabilism provides more suitable framework than the other theories, there are some elements within

---

[274] Swann et all., "Allure of Negative Feedback,": 293–306.
[275] Swann et all., 293–306
[276] Ibid.

the virtue reliabilism framework that a virtue reliabilist needs to highlight to do justice to the complexity related to our beliefs about our personality traits.

Chapter Four: The Epistemic Justification of Self-beliefs About Personality Traits: An Obstacle and a Proposal

In the previous chapter, I illustrated three empirical facts about our self-beliefs about our personality traits. The first fact was that people have some true beliefs about their personality traits. The second fact was that people's self-beliefs about their personality traits are not infallible because many factors may systematically interfere with the accuracy of these beliefs. The third fact was that some of our self-beliefs about our personality traits are inherently self-verifying since previous self-beliefs and behaviours can endlessly enforce new ones.

This chapter includes a challenge and a proposal. In terms of the challenge, I present a hypothetical scenario of Noor and Grace whose cases reflect some of the complexity related to our self-beliefs about our personality traits (at least as inspired by some of the empirical facts discussed in the previous chapter). Although there are nuanced differences between the two cases, both Noor and Grace end up forming the self-belief that they are genuinely highly agreeable. This means that they believe that they have the disposition of being highly empathetic, compassionate, and considerate of others' needs and feelings. Indeed, this also means that they think of their agreeable nature not just as superficial, but, rather, as a consistent, authentic, and integral part of their characters.

I illustrate what it takes for Noor's and Grace's self-belief to be justified from the internalist and the externalist points of view.[277] I then show that while the application of theories of justification in this context is less straightforward than their application in the context of our beliefs about the world, one might see how Noor's and Grace's self-beliefs can be considered justified according to the internalist standard. At the same time, one might see how both Noor's and Grace's self-beliefs can be considered unjustified according to the externalist's standard. In favor of externalism, I highlight a difference in the way we perceive the justificatory status in the two cases. I argue that while both self-beliefs are unjustified according to the externalist's standard, we do, to some extend, perceive Grace's self-belief to be, relatively speaking, somehow in a better position than Noor's self-belief. I argue that this difference in the way we perceive the justificatory status of Grace's and Noor's self-belief supports the externalist position that things beyond the person's mental resources such as the nuanced difference in the environment plays a role in increasing/decreasing the justificatory status of a certain belief.

Then, I zoom in into two externalist theories: process reliabilism and virtue reliabilism. I illustrate why they both consider Noor's and Grace's self-beliefs to be unjustified. For process reliabilism, this is because Noor's and Grace's self-belief-forming processes suffer from biases that affect the overall production of epistemically true self-beliefs. However, for virtue reliabilism, this is because Noor's and Grace's self-belief-

---

[277] It is important to remind the reader that the concern in this context is doxastic justification instead of propositional justification. In other words, the focus is on the conditions necessary for a belief to be considered justified, rather than on the conditions necessary for an individual to have reasonable grounds for holding that belief.

forming processes suffer from biases that make their processes less competent than how they potentially could. In light of this, I illustrate some reasons that make virtue reliabilism more promising as a theory of epistemic justification in the context of our self-beliefs about our personality traits.

In terms of the proposal, I highlight some elements in line with the virtue reliabilist framework. I argue that a virtue reliabilist may need to consider one condition that would potentially enhance the epistemic competency of Noor's and Grace's self-belief-forming process. This condition has to do with the features of the cases that a self-belief-forming process must rely on when generating self-beliefs. A virtue reliabilist may need to add that in the context of self-beliefs about personality traits, a competent self-belief-forming process has to generate self-beliefs from the cases where the person exhibits the personality trait in question despite having the opportunity and the motivation *not* to do so. Put differently, a virtue reliabilist may need to argue that, for more competency, a self-belief-forming process should not generate beliefs solely from cases in which the person simply manifests a certain personality trait. Instead, a virtue reliabilist may need to argue that, for more competency, a self-belief-forming process should consider the context in which a certain trait is manifested.

My proposal is in line with the virtue reliabilist framework. This is because a self-belief-forming process that is less subject to biases would have more epistemic competency than a self-belief-forming process that is more subject to biases. A trait that is manifested despite the person's motivation and opportunity to do otherwise is less likely to be subject to biases and the self-verifications discussed in the previous chapter. A trait that is

manifested despite the person's motivation and opportunity to do otherwise is more likely to reflect a genuine expression than traits that arise in response to certain environmental needs. Thus, even if the focus of virtue reliabilism is not the production of more true self-beliefs, a virtue reliabilist would assign an intrinsic value to the self-beliefs that seem to be more authentic and less impacted by biases or external pressure. Therefore, for virtue reliabilism, a self-belief-forming process that is sensitive to such cases would have more epistemic competency, and, therefore, more epistemic justification, according to the virtue reliabilist framework.

Two Hypothetical Case Scenarios: Noor and Grace

Let's introduce Noor. Noor is a self-identified woman who grew up in a society that values kindness, politeness, and accommodation, particularly in females. She has noticed that more assertive, confident, or outspoken women are often criticized in her social circles. Throughout time, Noor has learned to internalize those societal standards, prioritize being agreeable and tone down her assertiveness to fit in and avoid conflict. Despite this, Noor feels confident in her belief that she is an authentically highly agreeable person. She believes that she is naturally prone to be highly empathetic, compassionate, and considerate of others' needs and feelings. She believes that this agreeable nature is an integral part of her character.

Noor justifies her self-belief based on her own reflection on her experiences, actual behaviours, and the feedback she receives from others. For example, she recalls several instances where she behaved in a way that is considered to be "agreeable," such as volunteering extra hours for community projects, changing her plans for the day to listen

attentively to her friend's problems, and agreeing many times to substitute for her coworkers' absence. Additionally, she cites the feedback that supports her belief that she is highly agreeable. For example, her boss at work praises her for being a team player and for being willing to help her colleagues, and her family members express their gratefulness for her flexibility and easiness on holiday plans.

Indeed, Noor's self-belief about her agreeableness is further reinforced by her everyday choices. For example, one time her spouse advised her that she should consider saying no to coworkers who asked her to substitute for their absence. However, she thought that such a refusal would not align with her view of herself as a highly agreeable person. Hence, she keeps accepting her coworkers' requests, which ultimately added more items to her list of proofs of her agreeableness. Sometimes, Noor wonders if she is overestimating her self-belief about her level of agreeableness. She worries that she might be ignoring instances where she was not so agreeable, or that her friends and family were just being nice and polite when they praised her. Furthermore, she sometimes tries to imagine how her personality would be if things were different (e.g., if she had grown up in a different place). However, when she reflects again on her evidence, Noor thinks her self-belief is justified as she feels that she would always be genuinely highly agreeable, despite where she lives.

Now let's introduce Grace, who is also a self-identified woman. Grace grew up in a society that is relatively neutral about the level of agreeableness one ought to demonstrate. Thus, unlike Noor's society, Grace's environment had no explicit expectation that certain people should be accommodating or highly agreeable. This means that when it comes to

the formation of her self-belief about her level of agreeableness, Grace is less likely to be driven by the desire to conform to the societal standards that Noor had to face. As an example, since her childhood, Grace has had many opportunities to explore her level of agreeableness in various settings with no pressure or judgment. For instance, when she was in school, she participated in a debate club where she was able to practice and hone her argumentative skills. She found that she did well at defending her point of view and speaking up for what she believed in. However, when it came to conflict resolution, she couldn't handle it and felt so uncomfortable and inauthentic. Over the years and with many experiences like this, Grace formed the belief that she is *genuinely* highly agreeable.

Like Noor, Grace's self-belief is also based on her reflection on evidence including her experiences, actual behaviours, and the feedback she receives from others. Indeed, like Noor, Grace also wonders sometimes if she is overestimating her agreeableness. Moreover, like Noor, she also reassures herself after reflecting on her evidence that she is authentic in being highly agreeable. Furthermore, like Noor, Grace's belief is also influenced by her previous beliefs and behaviours. She acts in a way that confirms her view of herself as a highly agreeable person, and then, her actions count as new proof for her future self-beliefs. In short, the sum of Grace's evidence is very similar to that of Noor, and Grace's self-belief is also very similar to that of Noor. However, the fact that Grace had a variety of options in which her disposition could get manifested makes her agreeableness less affected by the previously discussed biases than Noor's. Similarly, the fact that Noor had very limited options in which her disposition could get manifested makes her agreeableness more affected by such biases than Grace's. As we will see later, this subtle distinction in their

environments would have consequences that carry on to a distinction we perceive in the justificatory status of their self-beliefs.

For clarification, there are various questions that can be asked about 'justification' in Noor's and Grace's hypothetical scenarios. For example, one may ask whether their self-belief is justified in the psychological, pragmatic, epistemic, or ethical sense. I focus only on justification in the epistemic sense. Indeed, one may ask whether Noor's and Grace's self-beliefs are epistemically justified in the doxastic or the propositional sense. I focus only on doxastic justification: I ask whether their self-beliefs are justified, and I do not ask whether they are justified in having such self-beliefs. Finally, one may ask about the justification of the self-belief that each one of them might have had if each of them had lived elsewhere (e.g., the justification of Noor's or Grace's self-belief that she is highly agreeable if either of them had lived in a less/more judgmental society). I only focus on the justification of their *current* self-beliefs that they are highly agreeable, given their current situation. The questions in short are as follows: Is Noor's/Grace's self-belief that she is genuinely highly agreeable justified? Do their self-beliefs reflect our intuition regarding what it means for a belief to be justified? The next section answers this question from an internalist and an externalist point of view.

Part 1: Why an Internalist Account About Noor's and Grace's Self-belief Is Unsatisfying

As illustrated in Chapter 1, internalism is a family of theses that share the idea that the epistemic justification of a belief depends solely on things internal to the person, such

as the person's evidence, experiences, and reasons for having a certain belief.[278] As Conee and Feldman argue, "the justificatory status of a person's doxastic attitudes strongly supervenes on the person's occurrent and dispositional mental states, events, and conditions."[279] Although internalists disagree on how much access/awareness the person needs to have (e.g., strict and less strict accessibilism requires the person to be somehow aware of her evidence while mentalism requires the person to have at least some relevant mental state), they agree that a belief B at time T is justified as long as B is grounded on reasons in the believer's relevant mental states.[280]

Based on this, an internalist would contend that both Noor's and Grace's self-beliefs that they are genuinely highly agreeable are justified. Overall, this is because internalists would evaluate Noor's and Grace's self-beliefs based on things within their cognitive resources. Even though these reasons reflect a complex interaction between many psychological, social, and political factors in a way that is less prevailing in Noor's and Grace's self-beliefs about their current mental state, a strict or less strict accessibilist would still highlight the process of reflection that both Grace and Noor engaged with in evaluating their self-beliefs. A strict or less strict accessibilist would state that Noor's and Grace's self-beliefs are justified because whenever they engage in critical reflection, they can point to numerous instances where their behaviour aligns with agreeable traits and can cite several testimonies where others comment on their agreeableness. Furthermore, a mentalist

---

[278] Earl Conee and Richard Feldman "Internalism Defended," American Philosophical Quarterly (Oxford) 38, no. 1 (2001): 1–18.

[279] Earl Conee and Richard Feldman, Evidentialism, (Oxford: Oxford University Press, 2004) 57.

[280] Conee and Feldman, Evidentialism, 57-9.

would also argue that Noor's and Grace's self-beliefs are justified; however, a mentalist would not require the same level of awareness as that of the strict or less strict accessibilist, but would, rather, point out that Noor's and Grace's self-beliefs are justified as long as their beliefs are somehow linked to a relevant mental state of some kind (e.g., a memorial belief of being agreeable or a sense of feeling highly agreeable).[281] Thus, from the internalist point of view, the overall accumulation of Noor's and Grace's personal experiences makes their self-beliefs justified.

However, I argue that the internalists' position on the justification of Noor's and Grace's self-beliefs is not satisfactory. This is because the internalist position fails to capture the distinction we usually make when we perceive the justificatory status of Noor's and Grace's self-beliefs. Overall, Grace had a variety of options in which her disposition could get manifested, and therefore, her agreeableness is less affected by the previously discussed biases than Noor's. As a result, we do in fact perceive Grace's self-belief to be, relatively speaking, in a better position than Noor's self-belief. The force of the difference in the way we perceive the justificatory state of each case lies beyond the mental resources of Noor and Grace because they both have identical self-beliefs. The force of the difference in the way we perceive the justificatory state of the two cases lies in the subtle distinction

---

[281] The main point for mentalists is that facts about justification must have their basis in one's mental state. Some mentalists even specify this further by stating that such a mental state must be non-factive (e.g., Feldman & Conee, 2001). A mental state is considered non-factive if it does not necessarily imply that its propositional content is true. In contrast, a factive mental state, such as 'knowing P', entails that the propositional content is indeed true. For example, Feldman & Conee argue that having a memorial belief that P (instead of remembering that P) is an example of a non-factive mental state that can satisfy the justification requirement. This becomes easier to digest when one considers that internalists sharply distinguish between being justified and being infallible.

in the environment in which their belief-forming processes operate, and, therefore, the difference cannot be accounted for by the internalists' standard.

Noor and Grace have identical self-beliefs but different self belief-forming processes. The differences in their belief-forming processes can be attributed to the consequences of the subtle distinction in their environment. One way to illustrate this is in terms of the motivational and informational biases discussed in Chapter Three. Given the rigid expectations of her society, Noor is at higher risk of being systematically affected by the desire to fit in to her society, and this can influence the way she forms her self-beliefs. For example, she might have experienced the 'self-enhancement phenomenon', in which certain aspects of her self-view got inflated due to the desire to prove, gain, protect, and maintain a feeling of self-worth as defined in her society. She also might have experience the 'above-average effect', in which she might have misused available information and viewed herself as above average with respect to desirable traits in her society. Thus, while both Noor and Grace have identical self-belief regarding their level of agreeableness, they employ different self belief-forming processes as a consequence of the difference in their environment. However, because this difference lies outside their cognitive resources, there is no way that internalism can account for why it makes sense to think of one self-belief as somehow in a better position than the other.

One way an internalist may respond is by explaining why we have the tendency to perceive Grace's self-belief as somehow in a better position than Noor's self-belief. An internalist may argue that this is because we confuse justification with the fallibility/infallibility of Noor's and Grace's self-beliefs. An internalist may argue that we

tend to perceive Grace's self-belief as relatively in a better position than Noor's because we think Grace's self-belief is more likely to reflect her authentic self than Noor's belief. In this case, an internalist would remind us that it is almost a consensus among contemporary internalists that a false belief can still be justified.[282] Thus, even if Grace's self-belief is more likely to be an accurate description of her authentic self, an internalist would argue that, as long as justification is concerned, both self-beliefs enjoy the same level of justification.

In reply, I state that the tendency to perceive Grace's self-belief as somehow in a better position than Noor's self-belief stands independently with the fallibility/infallibility of Noor's or Grace's self-beliefs. In fact, while it is possible that Noor and Grace would have been far less or more agreeable than they currently are if they had been raised in a different environment, it is also possible that they would have expressed the same level of agreeableness even if they had lived elsewhere. This is because, given the complexity related to the psychology of personality traits, it is almost impossible to tell if the manifestation of a certain trait is more an expression of the person's biological makeup or a reflection of societal habituation.[283] However, even if we entertain the thought that both self-beliefs are (or might be) false, this still does not change the observation that because Noor's self-belief-forming process is more likely to be affected by biases provoked by her

---

[282] See Chapter One for more details.

[283] For more on the developmental complexity of our personality traits see the amazing review of Conger, Katherine Jewsbury, Tricia K. Neppl, and Laura Scaramella, "Special Section on Personality Development: Testing Environmental and Genetic Associations Across Generations," In *Developmental Psychology 57*, no. 2 (2021): 139–46.

environment, we do tend to perceive Grace's self-belief as somehow in a better epistemic position than Noor's self-belief.

This is to say that we tend to give some weight to the context in which a certain self-belief-forming process operates. Because some contexts can provoke more biases in a self-belief-forming process than others, we tend to think that a certain self-belief is less justified when it is produced by a biased self-belief-forming process. Thus, the best explanation of the asymmetry in the way we perceive the justification of the two self-beliefs is that we think that justification is not entirely internal to the individual as some external factors can also play a role. This is why I believe that, at least in the context of our beliefs about our personality traits, internalism is unsatisfying.

Part 2: Why Process Reliabilism Is Unsatisfying Too

As discussed in Chapter 1, externalism includes a variety of theories that share the assumption that justification does not depend solely on the person's mental resources.[284] Externalists argue that many factors that are external to the person's mental state can play a role in justification. However, different externalist theories provide different views as to which external factors are particularly relevant to epistemic justification.

According to process reliabilism, a belief is justified if and only if it is produced by a reliable belief-forming process, and this is because a reliable belief-forming process tends to produce more true beliefs than unreliable processes.[285] For example, according to process

---

[284] Alston, *Epistemic Justification*, 20-54.
[285] Ibid

reliabilism, a belief that the Earth revolves around the Sun would be justified if and only if it is formed through a reliable belief-forming process, such as relying on scientific observations and evidence gathered by credible astronomers. For process reliabilists, the reason why these processes are more reliable than other processes (e.g., mere speculation) is that they have a higher chance of producing a higher proportion of true beliefs overall.

This means that, for a process reliabilist, even if two different belief-forming processes produce the exact same belief (i.e., that the Earth revolves around the Sun), the belief that was produced by the more reliable process is more justified. Thus, a process reliabilist would argue that, first, because both Noor's and Grace's self-belief-forming processes were at a higher risk of being affected by self-verification and because this risk affects the propensity that such process produces more true beliefs overall, their self-beliefs are generally unjustified. Second, a process reliabilist would share our tendency to differentiate between the justificatory status of Grace's and Noor's self-belief based on the differences in their self-belief-forming processes. This is to say that a process reliabilist would argue that because Noor's self-belief-forming processes was at a higher risk of being affected by informational and motivational biases than Grace's and because this risk affects the propensity that such process produces more true beliefs overall, Noor's self-belief-forming process is, relatively speaking, less reliable than Grace's self-belief-forming process. Therefore, a process reliabilist would conclude that, regardless of the fallibility or infallibility of that single self-belief, Grace's self-belief is in a better position than Noor's self-belief because Grace's self-belief is produced by a self-belief-forming process that is relatively more reliable than Noor's.

Although process reliabilism does an overall better job in explaining why we have the tendency to perceive Grace's self-belief as somehow in a better position than Noor's self-belief, the fact that process reliabilism cashes out its notion of epistemic reliability solely in a truth-conducive sense makes process reliabilism less satisfying as a theory of justification of our beliefs about our personality traits. The biases and the self-verifying nature of some of our self-beliefs makes the self-belief-forming process reliable in a *pragmatic* sense, but not necessarily in the robust epistemic sense. However, process reliabilism would not give up their adherence to epistemic reliability per se, and this makes process reliabilism less satisfying as a theory of justification, at least, in the context of our beliefs about our personality traits.

The fact that some self-belief-forming process can be biased and self-verifying is not trivial: it impacts the kind of reliability that can be obtained in this sense. Before I make this point, I need to clarify the distinction between epistemic and pragmatic reliability.[286] Epistemic reliability primarily pertains to the accuracy and the truth-tracking capabilities of the belief-forming processes in question. It focuses on the extent to which a belief-forming process reliably produces true beliefs or beliefs that are more likely to be true.[287] However, pragmatic reliability is more concerned with the practical effectiveness and the utility of belief-forming processes with respect to achieving desirable outcomes or goals.[288] This is to say that pragmatic reliability emphasizes the degree to which a belief-forming

---

[286] Robert G. Hudson, "Reliability, Pragmatic and Epistemic," *Erkenntnis* (1975-) 40, no. 1 (1994): 73-8.
[287] Hudson, "Reliability, Pragmatic and Epistemic," 73-8.
[288] Hudson, 74.

process reliably produces beliefs that are beneficial, lead to success, or facilitate effective decision-making instead of beliefs that are objectively true or accurate.[289]

A belief-forming process that is affected by biases and self-verification is necessarily pragmatically reliable but may or may not be epistemically reliable. There are two important points here. The first point is that if epistemic reliability refers to the tendency of a belief-forming process itself to aim at truth instead of utility, then, in principle, a psychological belief-forming process cannot be reliable in this sense. This is because, as psychologists suggest, the main goal of most of our psychological biases and self-verification is not necessarily to maintain the genuine expression of the person, but, rather, to simply promote survival. Thus, self-belief-forming processes are pragmatically reliable because the process itself is directed toward utility rather than truth per se.

The second point is that there is no reason to assume that the person's behaviour in a specific context exhausts all the potential expressions of the person's disposition with respect to a certain personality trait. It is true that Noor has behaved more agreeably. However, it is also conceivably possible that she would be able to express a different level of agreeableness had she lived in a different context. Thus, there is no reason to assume that what Noor has manifested in a certain context exhausts the list of how her disposition can potentially get manifested. Noor behaved more agreeably because this is the best move, pragmatically speaking, given the rigid expectations of her environment. Even if Noor had

---

[289] For more on the distinction between epistemic and pragmatic reliability see, for example, Isaac Davis, "A Framework for Pragmatic Reliability," Philosophy of Science 87, no. 4 (2020): 704–26; and Jeremy Fantl and Matthew Mcgrath, "On Pragmatic Encroachment in Epistemology," In *Philosophy and Phenomenological Research* 75, no. 3 (2007): 558–89.

the disposition not to be as agreeable as she is, her strict environment would provoke her biases and self-verification to work in a way that makes her behave more agreeably. However, this does not mean that 'behaving more agreeably in context X' is all that is there for Noor's level of agreeableness. Thus, the fact that Noor has behaved more agreeably in a particular environmental condition (especially one where she has no option to do otherwise) does not by itself establish that Noor is a highly agreeable person.[290]

One might argue that Noor's self-belief that she is highly agreeable is true in the sense that it accurately reflects her current level of agreeableness. The objection is that despite the pragmatic function of her belief, Noor's self-belief is still epistemically true simply by virtue of being a true reflection of her current state. In reply, I do not deny that having a true belief is an epistemic achievement of some kind. However, as I will illustrate in detail later, Noor's self-belief about her level of agreeableness would have more epistemic value if she had her belief while having the opportunity and the motivation to do otherwise with respect to the manifestation of her personality trait. If these conditions are met, then her belief would be less pragmatically driven, and, thus, it would make sense to think that Noor's self-belief about her level of agreeableness is more accurate in an epistemic sense.

---

[290] It is important to clarify that this thesis acknowledges that personality traits are not entirely independent of the environment. It recognizes that any personality trait must be expressed and manifested within a specific environment. However, as will be illustrated later, the primary emphasis lies in highlighting an epistemological distinction between traits that emerge within an environment that allows for various possibilities and those formed in an environment that restricts other options.

However, what is at stake in the distinction between epistemic and pragmatic reliability is what the belief-forming process as a whole aims at. A belief-forming process that is affected by biases and self-verification does not pertain to the accuracy or to the fair expression of someone's dispositions, but, rather, it aims at achieving utility. Thus, even though Noor's self-belief has some epistemic value by virtue of producing beliefs that fit her current reality, this does not undermine the fact that her self-belief-forming process as a whole is directed at achieving pragmatical outcomes.

Process reliabilists would not give up their strong adherence to epistemic reliability in the robust sense. Process reliabilists state that epistemic reliability is to be understood as a propensity in which a certain belief-forming process produces beliefs that are objectively true.[291] For example, in a recent article, Goldman distinguishes between two kinds of belief-forming processes: token and type. For Goldman, a token belief-forming process has no tracking record because it produces beliefs that occur only once at a specific time and place while a type belief-forming process has a tracking record because it produces beliefs that reoccur multiple occasions and in multiple places.[292] Goldman argues that "[a type belief-forming process] is a kind of thing that can acquire statistical properties (or propensities) such as producing true beliefs 80 percent of the time, or 20 percent of the time, etc. According to process reliabilism, it is precisely such statistical properties that determine the reliability or unreliability (or some degree of reliability) of specific process types."[293]

---

[291] Alvin Goldman. "A Different Solution to the Generality Problem for Process Reliabilism," *Philosophical Topics 49,* no. 2 (2022): 105–11.
[292] Goldman, "A Different Solution to the Generality Problem," 106.
[293] Goldman, 105.

Indeed, Goldman is explicit in excluding psychological factors (such as wishful thinking) from his notion of epistemic reliability. He writes: "According to process reliabilism, a person's belief is justified just in case it is generated by a reliable belief-forming process. If somebody merely wishes that P is true, and proceeds to believe it for that reason only, the belief is not justified." Thus, pragmatical considerations, such as the ones that are posed by our biases and self-verification, do not satisfy the standard of process reliabilism.

Thus, while process reliabilism does account for external factors in justification and, therefore, explains why we perceive Grace's self-belief as relatively in a better position than Noor's self-beliefs, the fact that self-belief-forming processes are ultimately driven by pragmatical considerations makes all beliefs produced by such processes unjustified according to process reliabilism. However, given that it is almost impossible to entirely get rid of such biases and self-verification, a process reliabilist would insist that no self-belief about personality traits can *ever* be ultimately justified. At this point, we can either accept the process reliabilist conclusion or look for another theory of justification that can account for the complexities of our self-beliefs about our personality traits. In the next section, I will take the second option and argue why virtue reliabilism seems a promising candidate.

Why Virtue Reliabilism Is More Suitable for Our Self-Beliefs About Our Personality Traits

As discussed in Chapter 1, virtue reliabilism shares many assumptions with process reliabilism. As an externalist theory, virtue reliabilism shares the idea that justification does not depend only on the person's internal resources. As a reliabilist theory, virtue reliabilism shares the emphasis on the role that a reliable belief-forming process plays in justification.

However, virtue reliabilism differs from process reliabilism in, at least, two respects. After illustrating these differences, I argue that these two features of virtue reliabilism make it more promising as a theory of epistemic justification in the context of our beliefs about our personality traits.

The first respect in which virtue reliabilism differs from process reliabilism is in its way of cashing out the notion of *reliability*. While reliability for process reliabilism is to be cashed out in a robust statistical sense, for virtue reliabilism, the reliability of a belief-forming process is not solely determined in a true-conducive sense. As we saw earlier, Sosa explicitly criticizes how process reliabilism understands reliabilism solely in terms of the mere likelihood of a belief being true.[294] Instead, virtue reliabilism emphasizes that the notion of reliability is to be extended beyond producing true beliefs to the exercise of intellectual virtues.[295] According to virtue reliabilism, a belief-forming process is reliable if it is epistemically competent.[296] For virtue reliabilism, the epistemic competencies include a range of cognitive virtues that have to be exhibited both by the believer and the belief-forming process. Examples of these virtues include open-mindedness, intellectual honesty, careful deliberation, and any other feature that might contribute to the overall epistemic excellence of an individual's belief-forming processes.[297] For virtue reliabilism,

---

[294] Sosa, Epistemology, 40.
[295] Ibid, 47
[296] Ibid, 44-9.
[297] Ibid, 40-2.

the mere manifestation of epistemic competence holds value, irrespective of the ultimate truth value of the beliefs formed through such processes.[298]

In this sense, virtue reliabilism places importance not only on the truth-conduciveness aspect of the belief-forming process but more on the manifestation and the exercise of epistemic virtues.[299] This makes virtue reliabilism more suitable in the context of our self-beliefs about our personality traits because it allows us to say that by exercising epistemic virtues, some of our self-beliefs about our personality traits can become more justified in some sense even if they are linked to pragmatical considerations. Again, this is because virtue reliabilism prioritizes the engagement and the manifestation of intellectual virtues over the mere production of true beliefs in the statistical sense. Thus, the fact that virtue reliabilism goes beyond process reliabilism by emphasizing the significance of epistemic competence makes virtue reliabilism more suitable in the context of our self-beliefs about our personality traits.

The second respect in which virtue reliabilism differs from process reliabilism, making it more suitable in the context of our beliefs about our personality traits, is its perspective on the role of *internal resources* in justification. In contrast to process reliabilism where the person's internal resources are neither necessary nor sufficient for justification, virtue reliabilism holds that the person's internal resources are necessary but insufficient. As discussed in Chapter One, for process reliabilism, a belief can be justified if and only if external factors are obtained; however, for virtue reliabilism, internal factors

---

[298] Ibid
[299] Ibid

are still needed for the exercise of intellectual virtue even if epistemic justification must not solely rely on such internal factors.[300]

But why would giving some weight to the internal factors matter in the context of our beliefs about our personality traits? The significance of considering internal factors in the context of our beliefs about our personality traits can be better understood through the following thought experiment. Imagine a scenario where scientists have developed a machine capable of generating highly accurate beliefs. While it might seem acceptable for scientists to utilize such a machine for acquiring quick and accurate beliefs about the world, it raises a degree of unease when Noor and Grace rely solely on such a machine to form self-beliefs about their own personality traits, without engaging their cognitive resources to some extent. We might have no concerns about the justification of a scientist's belief about the world generated by the machine. However, it does not seem weird to have some reservations about the justification of a self-belief generated by the machine.

One potential reason for this asymmetry in our attitude lies in the fact that, unlike other beliefs about the world, self-beliefs are somehow deeply intertwined with personal experiences, values, and self-identity. This deep integration of self-beliefs with our individuality makes it challenging to comprehend how a belief imported by such an external machine could genuinely and meaningfully capture one's authentic self. In other words, it is very hard to conceive how a belief imported by such an external machine is a belief about *oneself* in an authentic and meaningful sense. This is because the authenticity and

---

[300] Sosa, "Goldman's Reliabilism and Virtue Epistemology," 388.

meaningfulness of self-beliefs depend on the rich inputs from one's personal history, emotions, and subjective interpretation, which may be difficult for an external machine to fully grasp. As a result, the importation of beliefs by such a machine may lack the nuanced understanding and depth necessary to align with what we would think of a self-belief. Thus, although we already established that justification needs not to rely solely on internal resources, the mere reliance on external factors, on the other hand, also does not do justice to what we perceive to be justification in the context of self-beliefs. Thus, by giving some weight to internal factors, virtue reliabilism becomes a more suitable theory of epistemic justification in the context of our beliefs about our personality traits.

At this point, I have established that, in addition to internal sources, the process by which the self-beliefs are formed matters in their justification. I compared and contrasted two forms of reliabilism and argued in favour of virtue reliabilism over process reliabilism. The whole discussion makes use of several facts from the empirical psychology of personality traits. These facts evolve around the biases that combine some of our beliefs about our personality traits. These facts undermine the ability of a self-belief-forming process to maintain epistemic reliability in the ultimate sense. I used this as one of the reasons that make virtue reliabilism more plausible in the context of our self-beliefs about our personal traits. In the next section, I will highlight some elements in line with the virtue reliabilist framework. I will suggest some conditions that if considered, will boost the epistemic competency of a self-belief-forming process in the context of our beliefs about our personality traits.

A Proposal: What a Competent Self-belief-forming Process Has to Consider in the Context of our Self-beliefs about our Personality Traits

Sosa argues that a justified belief is a "manifestation of an epistemic competence," [301] where he argues that "[a] belief or judgment is epistemically competent iff it is formed well enough in epistemic respects."[302] Based on this, it is not unreasonable to think that a belief-forming process that tries to avoid biases is more epistemically competent than one that does not. However, since it is almost unattainable to get rid of all biases that accompany our self-beliefs, I propose that, within the virtue reliabilist framework, the epistemic competency of self-belief-forming processes can be enhanced, to some extent, if certain conditions are met. More specifically, I argue that in forming a self-belief about a certain personality trait, a competent self-belief-forming process should rely on cases where the person manifests the trait in question despite having the opportunity and the motivation *not* to do so. This is to say that a self-belief-forming process would be more epistemically competent in the context of personality traits if the self-belief-forming process considers cases where the person had the opportunity and the motivation to do otherwise but still could not help but to manifest the trait in question.

The following example illustrates my point. Consider Noah, who holds the self-belief that he is an introvert. Noah's self-belief-forming process could have employed many ways to arrive at his self-belief. For example, his self-belief-forming process could generate such a belief from instances where Noah acted like an introvert, without a consideration

---

[301] Sosa, Epistemology, 143.
[302] Sosa, 106.

whatsoever as to whether he had the opportunity or the motivation to do otherwise. An example of such instances would be the times when Noah preferred solitude, felt drained after social interactions, or exhibited any typical introverted behaviour regardless of the context. However, Noah's self-belief-forming process could also generate such a belief from instances where Noah could not help but be an introvert, despite having the opportunity and the encouragement to be otherwise. An example of such instances would be the times when Noah promised himself to talk to more people, but, yet ended up naturally gravitating towards a quieter corner in a large social gathering, engaging in one-on-one conversations, or feeling more comfortable observing from a distance. I propose that the epistemic competency of Noah's self-belief-forming process would be enhanced, within the virtue reliabilist framework, if it generates his self-belief from cases that resemble the one in the social gathering.

This is simply because the self-belief-forming process that considers the cases where the person had the opportunity and the motivation to do otherwise but still exhibits the trait in question is less influenced by the biases that combine self-beliefs about our personality traits. My argument unfolds in the answers to the following two questions: what do I mean by the opportunity and the motivation to do otherwise in the context of our self-beliefs about personality traits? Why would the reliance on such cases enhance the epistemic competency of our self-belief-forming process in the context of our beliefs about our personality traits?

In terms of the first question, the *opportunity* to do otherwise in the context of our self-belief about our personality traits refers to the amount of environmental pressure that

is involved in the unfolding of the personality trait in question.[303] A person who undergoes less social and environmental pressure has the opportunity to do otherwise with respect to the manifestation of a certain personality trait compared to someone who had to meet specific environmental needs. In other words, the person who experiences lower levels of social and environmental pressure enjoys a greater range of options in terms of expressing a specific personality trait, whereas the person who faces higher levels of social and environmental pressure is compelled to conform or adapt their behaviours based on the societal influences and expectations. Thus, the opportunity to do otherwise reflects the extent to which a person's environmental circumstances allow for alternative expressions of the person's personality traits.[304] An example of having the opportunity to do otherwise was illustrated in the case of Noor and Grace: Noor had to fit her personality with the strict expectations put by her society while Grace lacked such a constraint and therefore, had a wider range of variety with respect to the expression of her personality traits.

---

[303] By adopting this perspective on the notion of having the opportunity to do otherwise, I intend to bypass the current debates surrounding doxastic voluntarism. Doxastic voluntarism posits that individuals possess direct control over their beliefs and can alter them at their discretion (see, Michael J. Shaffer, "Doxastic Voluntarism, Epistemic Deontology, And Belief-Contravening Commitments," In American Philosophical Quarterly (Oxford) 50, no. 1 (2013): 73–81). However, I have not taken a definitive stance on the matter. Therefore, I would rather explain the ability to do otherwise concerning the magnitude of environmental influence as this can be compatible with accepting or rejecting doxastic voluntarism.

[304] Psychologists agree that the manifestation of any personality trait is a complex interaction between the person's environment and biological and genetic makeup (see for example, Conger et al, "Special Section on Personality Development: Testing Environmental and Genetic Associations Across Generations," In Developmental Psychology 57, no. 2 (2021): 139–46). However, there are studies that suggest that certain levels of environmental pressure can intervene with the developmental trajectories of a certain personality trait. This has been particularly studied in the context of twins who lived together compared to those who lived apart (see, for example, John C. Loehlin, and Robert C. Nichols, Heredity, Environment, and Personality: a Study of 850 Sets of Twins. (Austin, Texas: University of Texas Press, 1976) 25-142.

However, the *motivation* to do otherwise in the context of our self-belief about our personality traits refers to the individual's general openness or willingness to exhibit behaviours or characteristics that deviate from their existing traits. The motivation to do otherwise can include a potential or an actual desire for change: a person has the motivation to do otherwise in respect of their personality traits if she would want to modify her personality trait or if she possesses an actual desire to exhibit different behaviours or characteristics. A person who has the motivation to do otherwise does not necessarily dismiss or reject her current traits. A person has the motivation to do otherwise as long as the person is willing somehow or at least is open to exploring different dimensions of their selves. An example of having the motivation to do otherwise was illustrated in the case of Noah, who views himself as an introvert. Noah has the motivation to do otherwise simply because he promised himself to talk to more people at the party, which means that he was willing and open to engaging in more extroverted behaviours.

In terms of the second question, the self-belief-forming process that generates self-beliefs from cases in which the person exhibits the personality trait despite having the opportunity and the motivation to do otherwise has more epistemic competency than other self-belief-forming processes. This is simply because the self-belief-forming process that generates self-beliefs from such cases is less vulnerable to the biases discussed earlier. This point can be divided into two sub-arguments. First, I argue that the self-belief-forming process that generates self-beliefs from cases in which the person has the *opportunity* to do otherwise is less likely to be subject to motivational and informational *biases*. Second, I

argue that the self-belief-forming process that generates self-beliefs from cases in which the person has the *motivation* to do otherwise is less likely to be subject to self-verification.

First, the self-belief-forming process that generates self-beliefs from cases in which the person has the *opportunity* to do otherwise is less likely to be subject to motivational and informational *biases* discussed earlier. The very fact that someone can do otherwise with respect to a certain personality trait means that less environmental pressure is involved in the unfolding of that trait. Indeed, a trait that is expressed in less environmental pressure may get manifested in various ways without risking the person's survival or sense of belonging. Concerns related to the person's survival or sense of belonging may affect how the trait is manifested.[305] Thus, a self-belief-forming process that generates self-beliefs from cases where the manifestation of the trait was less affected by survival concerns is more likely to capture the more authentic expressions of that trait than other self-belief-forming processes.

This has been illustrated in the case of Noor and the case of Grace. Grace faced less environmental pressure and, therefore, had the opportunity to express various levels of agreeableness without risking her survival or her sense of belonging. She had the opportunity to be highly, somewhat, or even a little bit agreeable while maintaining a secure sense of belonging. In contrast, Noor faced rigid expectations and had to maintain a narrower range of acceptable expressions of agreeableness to secure her sense of belonging. They both ended up expressing a high level of agreeableness. However, because Grace had

---

[305] Taylor and Brown, "Illusion and Well-Being: a Social Psychological Perspective on Mental Health," Psychological Bulletin 103, no. Mar 88 (1988): 193–210.

the opportunity to do otherwise, we feel she is less affected by biases such as the self-enhancement. Indeed, because we feel that her self-belief-forming process is less affected by the self-enhancement, we feel that her self-belief-forming process is more competent. Further, because we feel that her self-belief-forming process is more competent, we feel that her self-belief that she is highly agreeable is somehow more justified. In short, the self-belief-forming process is more competent when it generates self-beliefs from cases in which the person has the *opportunity* to do otherwise.

Second, the self-belief-forming process that generates self-beliefs from cases in which the person has the *motivation* to do otherwise is less likely to be subject to the self-verification. The very fact that the person in those cases expresses an openness or willingness to change or modify a certain personality trait indicates that she does not view this trait as essential to her identity. 'Not viewing a certain trait as essential to one's identity' is crucial to reduce the effect of the self-verifying phenomenon. If a person does not view a certain trait as essential to his identity, then it is less likely that such a person would feel the urge to seek behaviours and make decisions that verify his beliefs about that trait. In other words, a person who has the motivation to do otherwise does not view a certain trait as essential to his identity and, therefore, does not experience the fundamental need to verify their self-belief about this trait through the continuous reinforcement of it. Thus, his self-belief-forming process in this case is more epistemically competent.

This is not to say that the motivation to do otherwise would eliminate all forms of biases. However, this is to say that, at least in the context of our self-beliefs about our personality traits, the motivation to do otherwise can to some extent decrease the impact of

such biases. To see this point, compare the case of Noor to the case of Noah. Noor had always thought of herself as a highly agreeable person. She seems to have developed a special attachment to her self-view and seems not to be motivated to do otherwise. This can be seen when Noor kept accepting her colleagues' requests to substitute on their behalf, despite the advice she got, merely because she recalled her preexisting self-belief. She was not motivated to give up or to even modify her level of agreeableness, and, therefore, she ended up making the decisions that verify her preexisting self-beliefs.

However, Noah had the motivation to do otherwise with respect to his level of introversion. He promised himself that he would try to talk to more people at the party. This motivation to do otherwise means that Noah did not view 'being an introvert' as an essential part of his identity. This entails that Noah would have no reason to seek the behaviours and make the decisions that verify his preexisting self-belief as an introvert. If Noah failed not to be an introvert, despite having the desire to do so, then his failure would probably be reflected by the state of affairs itself (i.e., his actual dispositional level of introversion) rather than the accumulation of his previous beliefs in the matter. Because Noah had the motivation to do otherwise, we feel he is less affected by the self-verifying phenomenon. Indeed, because we feel that his self-belief-forming process is less affected by the self-verifying phenomenon, we feel that his self-belief-forming process is more epistemically competent. Further, because we feel that his self-belief-forming process is more epistemically competent, we feel that his self-belief that he is an introvert is somehow more justified. In short, the self-belief-forming process is more epistemically competent

when it generates self-beliefs from cases in which the person has the *motivation* to do otherwise.

In conclusion, as Sosa puts it, "epistemic competence must include the proper conduct of inquiry."[306] Based on this, it is reasonable to assume that a belief-forming process that avoids (or tries to avoid) biases and self-verification has a better conduct of inquiry than the ones that do not. I showed that, in the context of our self-beliefs about our personality traits, a self-belief-forming process can be more competent within the virtue reliabilism framework, when it generates its beliefs from cases in which the person manifests a certain trait while having the opportunity and the motivation to do otherwise. I already discussed one important reason for this, which is that such cases are less susceptible to biases related to self-beliefs, which undermine the intellectual virtue of the process. I showed that it is through these cases that the epistemic competency of our self-belief-forming process can be enhanced to some extent. Thus, if the person had the opportunity and the motivation to do otherwise with respect to personality trait x but still fails not to manifest the trait x, then it is more likely that x is genuine to the person. In other words, if the person fails every attempt not to manifest x despite having the opportunity and the motivation to do so, then x is probably an authentic expression of the person's traits. Thus, a self-belief that is formed based on such cases is more likely to reflect the epistemic competency of the process and, therefore, be more epistemically justified.

---

[306] Sosa, Epistemology, 160.

Bibliography

Alicke, Mark D., "Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives." *Journal of Personality and Social Psychology* 49, no. 6 (1985): 1621–30. https://doi.org/10.1037/0022-3514.49.6.1621.

Alston, William P. *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca: Cornell University Press, 1989.

Armstrong, D. M. (David Malet). *A Materialist Theory of the Mind*. Rev. ed. London: Routledge, 1993.

Armstrong, D. M. *Belief, Truth and Knowledge.* Cambridge University Press, 1973. https://doi.org/10.1017/CBO9780511570827.

Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. Vol. 2. London: Routledge, 1998. https://doi.org/10.4324/9780203982105.

Back, Mitja D, Stefan C Schmukle, and Boris Egloff. "Predicting Actual Behavior From the Explicit and Implicit Self-Concept of Personality." *Journal of Personality and Social Psychology* 97, no. 3 (2009): 533–48. https://doi.org/10.1037/a0016229.

Baumeister, Roy F., Laura Smart, and Joseph M. Boden. "Relation of Threatened Egotism to Violence and Aggression: The Dark Side of High Self-Esteem." *Psychological Review 103,* no. 1 (01, 1996): 5-33. doi:https://doi.org/10.1037/0033-295X.103.1.5.

Bénabou, Roland, and Jean Tirole. "Intrinsic and Extrinsic Motivation." *The Review of Economic Studies* 70, no. 3 (2003): 489–520. https://doi.org/10.1111/1467-937X.00253.

Bergmann, Michael. *Justification Without Awareness: a Defense of Epistemic Externalism.* Oxford: Clarendon, 2006.

Bermúdez, José Luis. "The Sources of Self-Consciousness." In *The Bodily Self*. The MIT Press, 2018. https://doi.org/10.7551/mitpress/9780262037501.003.0004.

Biesanz, Jeremy C, Stephen G West, and William G Graziano. "Moderators of Self-Other Agreement: Reconsidering Temporal Stability in Personality." *Journal of Personality and Social Psychology 75,* no. 2 (1998): 467–77. https://doi.org/10.1037/0022-3514.75.2.467.Bonjour, Laurence. "Externalist Theories of Empirical Knowledge." *Midwest Studies in Philosophy 5*, no. 1 (1980): 53–74. https://doi.org/10.1111/j.1475-4975.1980.tb00396.x.

Bonjour, Laurence. "The Coherence Theory of Empirical Knowledge*." Philosophical Studies 30*, no. 5 (1976): 281–312. https://doi.org/10.1007/BF00357928.

BonJour, Laurence. "The Dialectic of Foundationalism and Coherentism." In *The Blackwell Guide to Epistemology,* 117–42. Oxford, UK: Blackwell Publishing Ltd, 2017. https://doi.org/10.1002/9781405164863.ch4.

BonJour, Laurence. *The Structure of Empirical Knowledge.* Cambridge, Mass: Harvard University Press, 1985.

Burge, Tyler, and Peacocke, Christopher. "Our Entitlement to Self-Knowledge." *Proceedings of the Aristotelian Society* 96, no. 1 (1996): 91–116. https://doi.org/10.1093/aristotelian/96.1.91.

Carlson, Erika N., and R. Michael Furr. "Evidence of Differential Meta-Accuracy: People Understand the Different Impressions They Make." *Psychological Science 20*, no. 8 (2009): 1033–39. https://doi.org/10.1111/j.1467-9280.2009.02409.x.

Carruthers, Peter. "Inner Sense Theories." In *The Opacity of Mind.,* 192–222*,* United Kingdom: Oxford University Press, 2011. https://doi.org/10.1093/acprof:oso/9780199596195.003.0007.

Cassam, Quassim. *Self-Knowledge for Humans*. Oxford: Oxford University Press,2014.

Chisholm, Roderick M. *The Foundations of Knowing*. NED - New edition. Minneapolis: University of Minnesota Press, 1982. https://doi.org/10.5749/j.ctttskfs.

Cohen, Stewart. "Justification and Truth." *Philosophical Studies* 46, no. 3 (1984): 279–95. https://doi.org/10.1007/BF00372907.

Coliva, Annalisa. *The Self and Self-Knowledge*. Oxford: Oxford University Press, 2012.

Conee, Earl, Feldman, Richard. *Evidentialism*. Oxford: Oxford University Press, 2004. https://doi.org/10.1093/0199253722.003.0005.

Conger, Katherine Jewsbury, Tricia K. Neppl, and Laura Scaramella. "Special Section on Personality Development: Testing Environmental and Genetic Associations Across Generations." *Developmental Psychology 57*, no. 2 (2021): 139–46. https://doi.org/10.1037/dev0001049.

Davis, Isaac. "A Framework for Pragmatic Reliability." *Philosophy of Science* 87, no. 4 (2020): 704–26. https://doi.org/10.1086/709755

De Cruz, Helen, Maarten Boudry, Johan De Smedt, and Stefaan Blancke. "Evolutionary Approaches to Epistemic Justification." *Dialectica 65*, no. 4 (2011): 517–35. https://doi.org/10.1111/j.1746-8361.2011.01283.x.

Dennett, Daniel C. "Real Patterns." *The Journal of Philosophy 88, no*. 1 (1991): 27–51. https://doi.org/10.2307/2027085.

Descartes, René. *Discourse on Method*. New York: Liberal Arts Press, 1950.

Dunn, Jeff. "Reliability for Degrees of Belief." *Philosophical Studies* 172, no. 7 (2015): 1929–52. https://doi.org/10.1007/s11098-014-0380-2.

Emundts, Dina. "*Kant's Ideal of Self-Knowledge."* In Self-Knowledge. New York: Oxford University Press, 2017. https://doi.org/10.1093/acprof:oso/9780190226411.003.0013.

Fantl, Jeremy, and Matthew McGrath. "Evidence, Pragmatics, and Justification." The *Philosophical Review* 111, no. 1 (2002): 67–94. https://doi.org/10.1215/00318108-111-1-67.

Fantl, Jeremy, and Mcgrath, Matthew., "On Pragmatic Encroachment in Epistemology," In *Philosophy and Phenomenological Research* 75, no. 3 (2007): 558–89. https://doi.org/10.1111/j.1933-1592.2007.00093.x.

Feldman, Richard, and Conee, Earl. "Internalism Defended." *American Philosophical Quarterly (Oxford)* 38, no. 1 (2001): 1–18.

Feldman, Richard, and Conee, Earl. "Evidence," In *Epistemology*. Oxford: Oxford University Press, 2008, 83-104. https://doi.org/10.1093/acprof:oso/9780199264933.003.0005.

Feldman, Richard. "Fallibilism and Knowing That One Knows." *The Philosophical Review* 90, no. 2 (1981): 266–82. https://doi.org/10.2307/2184442.

Foley, Richard. "Rationality and Perspective." *Analysis* 53, no. 2 (1993): 65–68. https://doi.org/10.2307/3328835.

Fumerton, Richard, "Acquaintance: The Foundation of Knowledge and Thought", in *Acquaintance: New Essays*, 245-459, Jonathan Knowles, and Thomas Raleigh (eds), Oxford, 2019; Oxford Academic, https://doi-org.libaccess.lib.mcmaster.ca/10.1093/oso/9780198803461.003.0012.

Fumerton, Richard. "Epistemic Internalism, Philosophical Assurance and the Skeptical Predicament." In *Knowledge and Reality,* 179–91. Dordrecht: Springer Netherlands, 2006. https://doi.org/10.1007/1-4020-4733-9_8.

Fumerton, Richard A. *Metaepistemology and Skepticism*. Lanham, Md: Rowman & Littlefield, 1995.

Gertler, Brie. *Privileged Access : Philosophical Accounts of Self-Knowledge*. Edited by Brie Gertler. London: Routledge, 2016.

Gertler, Brie. *Self-Knowledge*. New York: Routledge, 2011.

Gertler, Brie. "Self-Knowledge and Rational Agency: A Defense of Empiricism." *Philosophy and Phenomenological Research* 96, no. 1 (2018): 91–109. https://doi.org/10.1111/phpr.12288.

Gertler, Brie. "The Acquaintance Theory Of Self-Knowledge." In *Self-Knowledge*, 97–138. United Kingdom: Routledge, 2011. https://doi.org/10.4324/9780203835678-8.

Gettier, E. L. "Is Justified True Belief Knowledge?" *Analysis (Oxford)* 23, no. 6 (1963): 121–23. https://doi.org/10.1093/analys/23.6.121.

Goldman, Alvin I. "A Causal Theory of Knowing." *The Journal of Philosophy 64*, no. 12 (1967): 357–72. https://doi.org/10.2307/2024268

Goldman, Alvin. "A Different Solution to the Generality Problem for Process Reliabilism." *Philosophical Topics 49,* no. 2 (2022): 105–11.

Goldman, Alvin I. "Strong and Weak Justification." *Philosophical Perspectives 2* (1988): 51–69. https://doi.org/10.2307/2214068.

Goldman, Alvin I. "What Is Justified Belief?" In *Reliabilism and Contemporary Epistemology*. New York: Oxford University Press, 2012. https://doi.org/10.1093/acprof:oso/9780199812875.003.0002.

Goldberg, Sanford C. "The Psychology and Epistemology of Self-Knowledge." *Synthese (Dordrecht)* 118, no. 2 (1999): 165–99. https://doi.org/10.1023/A:1005110319056.

Goldberg, Sanford C. *To the Best of Our Knowledge: Social Expectations and Epistemic Normativity.* Oxford: Oxford University Press, 2018, 13–47. https://doi.org/10.1093/oso/9780198793670.001.0001.

Goldberg, Sanford C. "Epistemically Proper Belief." In *To the Best of Our Knowledge*. Oxford: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780198793670.003.0002.

Greco, John, and Christoph Kelp. *Virtue-Theoretic Epistemology: New Methods and Approaches*. Edited by John Greco and Christoph Kelp. Cambridge, UK: Cambridge University Press, 2020.

Halliwell, Stephen. "Justifying The World As An Aesthetic Phenomenon." *The Cambridge Classical Journal* 64 (2018): 91–112. doi:10.1017/S1750270518000064.

Hasan, Ali. "Access Internalism, Mentalism, and Reliabilism, 117-46, " In *A Critical Introduction to the Epistemology of Perception*. United Kingdom: Bloomsbury Publishing Plc, 2017.

Haslett, D. W. *Equal Consideration: a Theory of Moral Justification.* Newark: University of Delaware Press, 1987.

Hatzimoysis, Anthony. *Self-Knowledge*. Oxford: Oxford University Press, 2011.

Hofmann. "A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures." *Personality & Social Psychology Bulletin*. 31, no. 10 (n.d.): 1369–85. https://doi.org/info:doi/.

Hudson, Robert G. "Reliability, Pragmatic and Epistemic." *Erkenntnis* 40, no. 1 (1994): 71–86. https://doi.org/10.1007/BF01128716.

Kant, Immanuel, Paul Guyer, and Allen W. Wood. *Critique of Pure Reason*. Cambridge: Cambridge University Press, 1998.

Kim, Kihyeon. "Internalism and Externalism in Epistemology." *American Philosophical Quarterly* (Oxford) 30, no. 4 (1993): 303–16.

Kornblith, Hilary. " What Does Logic Have to do With Justified Belief?" Chapter Three in *Propositional and Doxastic Justification*, 40-67, New York: Routledge, 2022.

Kvanvig, Jonathan L. " Theoretical Unity and the Priority of Propositional Justification." Chapter Two In *Propositional and Doxastic Justification*, 27-39, New York: Routledge, 2022.

Klein, Peter D. "No Final End in Sight." In Current Controversies in Epistemology, 95–115. Routledge, 2014. https://doi.org/10.4324/9780203123522-6.

Korsgaard, Christine M. "The Activity of Reason." Proceedings and Addresses of the American *Philosophical Association* 83, no. 2 (2009): 23–43.

Lennon, Thomas M. *The Battle of the Gods and Giants: the Legacies of Descartes and Gassendi*, 1655-1715. Princeton, N.J: Princeton University Press, 1993.

Letzring, Tera D., and Jana S. Spain. *The Oxford Handbook of Accurate Personality Judgment*. Edited by Tera D. Letzring and Jana S. Spain. New York: Oxford University Press, 2019.

Lewis, Clarence Irving. *An Analysis of Knowledge and Valuation*. La Salle, Ill: The Open court publishing co., 1946.

Littlejohn, Clayton. *Justification and the Truth-Connection,* 1-44. Cambridge: Cambridge University Press, 2012. https://doi.org/10.1017/CBO9781139060097.

Locke, John. *An Essay Concerning Humane Understanding*. In Four Books. Written by John Locke, Gent. The second edition, with large additions. London: printed for Thomas Dring, at the Harrow,over-against the Inner-Temple Gate in Fleet-street; and Samuel Manship, at the Ship in Cornhill, near the Royal Exchange, 1694.

Loehlin, John C., and Robert C. Nichols. *Heredity, Environment, and Personality: a Study of 850 Sets of Twin*s. Austin, Texas: University of Texas Press, 1976.

Lyons, Jack C. "Should Reliabilists Be Worried About Demon Worlds?" *Philosophy and Phenomenological Research* 86, no. 1 (2013): 1–40. https://doi.org/10.1111/j.1933-1592.2012.00614.x.

Lyons, Jack C. "What We Talk About When We Talk About Epistemic Justification." *Inquiry (Oslo) 59, no*. 7-8 (2016): 867–88. https://doi.org/10.1080/0020174X.2016.1200811.

Madison, B. J. C. "Epistemic Value and the New Evil Demon." *Pacific Philosophical Quarterly* 98, no. 1 (2017): 89–107. https://doi.org/10.1111/papq.12054.

Matthews, Gerald, Ian J. Deary, and Martha C. Whiteman. *Personality Traits*. Third edition. Cambridge: Cambridge University Press, 2009.

McCain, Kevin. *Evidentialism and Epistemic Justification*. Vol. 59. London: Routledge, 2014. https://doi.org/10.4324/9781315882390.

Merton, Robert K., "The Self-Fulfilling Prophecy." *The Antioch Review*. 8, no. 1 (1948): 193–190. https://doi.org/info:doi/.

Montmarquet, James A. *Epistemic Virtue and Doxastic Responsibility*. Lanham, Md: Rowman & Littlefield Publishers, 1993.

Moore, Don A, and Small, Deborah A., "Error and Bias in Comparative Judgment: On Being Both Better and Worse Than We Think We Are." *Journal of Personality and Social Psychology* 92, no. 6 (2007): 972–89. https://doi.org/10.1037/0022-3514.92.6.972.

Moran, Richard. *Authority and Estrangement: an Essay on Self-Knowledge*. Princeton, N.J., Princeton University Press, 2001.

Neta, Ram. " Justification Ex Ante and Ex Post" Chapter Four in *Propositional and Doxastic Justification*, 70-87, New York: Routledge, 2022.

Nguyen, A. Minh. "What Good Is Self-Knowledge?" *Journal of Philosophical Research 40 (2015)*: 137–54. https://doi.org/10.5840/jpr2015111656.

Oliveira, Luis R. G., and Paul Silva. *Propositional and Doxastic Justification: New Essays on Their Nature and Significance*. Edited by Luis R. G. Oliveira and Paul Silva. New York, NY: Routledge, 2022. https://doi.org/10.4324/9781003008101.

Parent, Ted. "Self-Knowledge and Externalism About Empty Concepts." *Analytic Philosophy* 56, no. 2 (2015): 158–68. https://doi.org/10.1111/phib.12050.

Pashler.H, "Is the Replicability Crisis Overblown? Three Arguments Examined." *Perspectives on Psychological Science : a Journal of the Association for Psychological Science. 7, no. 6* (2012): 531–36. https://doi.org/info:doi/.

Paulhus, Delroy L., and Oliver P. John. "Egoistic and Moralistic Biases in Self-Perception: The Interplay of Self-Deceptive Styles With Basic Traits and Motives." *Journal of Personality 66,* no. 6 (1998): 1025–60. https://doi.org/10.1111/1467-6494.00041.

Pettigrew, Richard. "Radical Epistemology, Structural Explanations, and Epistemic Weaponry." *Philosophical Studies* 179, no. 1 (2022): 289–304. https://doi.org/10.1007/s11098-021-01660-x.

Plato., and John. McDowell. *Theaetetus*. Oxford: Clarendon Press, 1973.

Popkin, Richard H. *History of Scepticism: From Savonarola to Bayle.* Cary: Oxford University Press, 2003, 8-34.

Pronin, Emily, Justin Kruger, Kenneth Savitsky, and Lee Ross. "You Don't Know Me, But I Know You: The Illusion of Asymmetric Insight*." Journal of Personality and Social Psychology* 81, no. 4 (2001): 639–56. https://doi.org/10.1037/0022-3514.81.4.639.

Russell, Bertrand. "Knowledge by acquaintance and knowledge by description." *In The Basic Writings of Bertrand Russell,* 223–30. Routledge, 2009. https://doi.org/10.4324/9780203875391-39.

Russell, Bertrand. "Knowledge by acquaintance and knowledge by description." *In The Basic Writings of Bertrand Russell,* 223–30. Routledge, 2009. https://doi.org/10.4324/9780203875391-39.

Russman, Thomas A. "Roderick Chisholm: Self and Others." *The Review of Metaphysics 33*, no. 1 (1979): 135–66.

Ryle, Gilbert. *The Concept of Mind*. New York, Barnes & Noble, 1950.

Savary, Jennifer, and Kelly Goldsmith. "Unobserved Altruism: How Self-Signaling Motivations and Social Benefits Shape Willingness to Donate." *Journal of Experimental Psychology.* Applied 26, no. 3 (2020): 538–50. https://doi.org/10.1037/xap0000261.

Schaffer, Jonathan. "The Irrelevance of the Subject: Against Subject-Sensitive Invariantism." *Philosophical Studies*, 127, no. 1 (2006): 87–107. https://doi.org/info:doi/.

Schwitzgebel, Eric. "A Phenomenal, Dispositional Account of Belief." *Noûs (Bloomington, Indiana) 36*, no. 2 (2002): 249–75. https://doi.org/10.1111/1468-0068.00370.

Shaffer, Michael J. "Doxastic Voluntarism, Epistemic Deontology, And Belief-Contravening Commitments," *American Philosophical Quarterly (Oxford)* 50, no. 1 (2013): 73–81.

Smithies, Declan. "Belief and Self-Knowledge: Lessons From Moore's Paradox: Belief and Self-Knowledge." *Philosophical Issues 26*, no. 1 (2016): 393–421. https://doi.org/10.1111/phis.12075.

Smithies, Declan. *The Epistemic Role of Consciousness.* New York, NY: Oxford University Press, 2019.

Smithies, Declan. "Why Justification Matters." *In Epistemic Evaluation*. Oxford: Oxford University Press, 2015. 224–244. https://doi.org/10.1093/acprof:oso/9780199642632.003.0010.

Smith, Nicholas J. J. "Degree of Belief Is Expected Truth Value." *In Cuts and Clouds.* Oxford: Oxford University Press, 2010. https://doi.org/10.1093/acprof:oso/9780199570386.003.0029.

Sosa, Ernest. *Epistemology.* Princeton, NJ: Princeton University Press, 2017. https://doi.org/10.1515/9781400883059.

Sosa, Ernest. "Goldman's Reliabilism and Virtue Epistemology." *Philosophical Topics* 29, no. 1/2 (2001): 383–400. https://doi.org/10.5840/philtopics2001291/214.

Sosa, Ernest. "How Competence Matters In Epistemology." *Philosophical Perspectives* 24, no. 1 (2010): 465–75. https://doi.org/10.1111/j.1520-8583.2010.00200.x.

Sosa, Ernest. *Judgment and Agency*. First edition. Oxford: Oxford University Press, 2015.

Sosa, Ernest. "Reliabilism and Intellectual Virtue." In *Knowledge in Perspective*, 131–46. Cambridge University Press, 1991. https://doi.org/10.1017/CBO9780511625299.009.

Sosa, Ernest. "The Raft and the Pyramid: Coherence Versus Foundations in the Theory of Knowledge." *Midwest Studies in Philosophy* 5, no. 1 (1980): 3–26. https://doi.org/10.1111/j.1475-4975.1980.tb00394.x.

Srinivasan, Amia. "Radical Externalism." *The Philosophical Review* 129, no. 3 (2020): 395–431. https://doi.org/10.1215/00318108-8311261.

Steup, Matthias. *An Introduction to Contemporary Epistemology*. Upper Saddle River, NJ: Prentice Hall, 1996.

Steup, Matthias., John. Turri, and Ernest. Sosa. *Contemporary Debates in Epistemology*. 2nd ed. Hoboken: Wiley-Blackwell, 2014.

Stone, Arthur A. *The Science of Self-Report: Implications for Research and Practice*. Mahwah, NJ: Psychology Press, 2000. https://search-ebscohost-com.libaccess.lib.mcmaster.ca/login.aspx?direct=true&db=nlebk&AN=24260&site=ehost-live&scope=site.

Swann, William B., and Brett Pelham. "Who Wants Out When the Going Gets Good? Psychological Investment and Preference for Self-Verifying College Roommates.*" Self and Identity* 1, no. 3 (2002): 219–33. https://doi.org/10.1080/152988602760124856.

Swann, William B. *Handbook of Theories of Social Psychology*: Volume 2. 2 vols. London: SAGE Publications Ltd, (2012): 2-33. https://doi.org/10.4135/9781446249222

Swann, William B, Richard M Wenzlaff, Douglas S Krull, and Brett W Pelham. "Allure of Negative Feedback: Self-Verification Strivings Among Depressed Persons." *Journal of Abnormal Psychology (1965)* 101, no. 2 (1992): 293–306. https://doi.org/10.1037/0021-843X.101.2.293.

Taylor, S E, and J D Brown. "Illusion and Well-Being: a Social Psychological Perspective on Mental Health." *Psychological Bulletin* 103, no. Mar 88 (1988): 193–210.

Vazire, Simine, and Erika N. Carlson. "Others Sometimes Know Us Better Than We Know Ourselves." *Current Directions in Psychological Science: a Journal of the American Psychological Society* 20, no. 2 (2011): 104–8. https://doi.org/10.1177/0963721411402478.

Vazire, Simine, and Erika N. Carlson. "Self-Knowledge of Personality: Do People Know Themselves?" *Social and Personality Psychology Compass* 4, no. 8 (2010): 605–20. https://doi.org/10.1111/j.1751-9004.2010.00280.x.

Vazire, Simine, and Mehl R., Matthias. "Knowing Me, Knowing You: The Accuracy and Unique Predictive Validity of Self-Ratings and Other-Ratings of Daily Behavior." *Journal of Personality and Social Psychology* 95, no. 5 (2008): 1202–16. https://doi.org/10.1037/a0013314.

Vazire, Simine, and Wilson D. Timothy. *Handbook of Self-Knowledge*. New York, N.Y: Guilford Press, 2012.

Vazire, Simine. "Informant Reports: A Cheap, Fast, and Easy Method for Personality Assessment." *Journal of Research in Personality* 40, no. 5 (2006): 472–81. https://doi.org/10.1016/j.jrp.2005.03.003.

Vazire, Simine., "Knowing Me, Knowing You: The Accuracy and Unique Predictive Validity of Self-Ratings and Other-Ratings of Daily Behavior." *Journal of Personality and Social Psychology*. 95, no. 5 (2008): 1202–16. https://doi.org/info:doi/.

Vazire, Simine., "Who Knows What About a Person? The Self-Other Knowledge Asymmetry (SOKA) Model." *Journal of Personality and Social Psychology*. 98, no. 2 (2010): 281–300. https://doi.org/info:doi/.

Widiger, Thomas A. *The Oxford Handbook of the Five Factor Model of Personality*. Edited by Thomas A. Widiger. New York, NY: Oxford University Press, 2015.

Wilson, Timothy D. "Know Thyself." *Perspectives on Psychological Science* 4, no. 4 (2009): 384–89. https://doi.org/10.1111/j.1745-6924.2009.01143.x.

Wright, Crispin, Barry C. Smith, and Cynthia Macdonald. "*Knowing Our Own Minds*." Oxford: Clarendon, 1998.

Wright, Crispin. "Scepticism and Dreaming: Imploding The Demon." *Noûs (Bloomington, Indiana)* 25, no. 2 (1991): 87-116. https://doi.org/10.2307/2215583.

Zagzebski, Linda. "Précis of Virtues of the Mind." *Philosophy and Phenomenological Research* 60, no. 1 (2000): 169–77. https://doi.org/10.2307/2653437.

Zagzebski, Linda. "The Inescapability of Gettier Problems." *The Philosophical Quarterly* 44, no. 174 (1994): 65–73. https://doi.org/10.2307/2220147.

Zagzebski, Linda Trinkaus. "The Methodology of Epistemology." Chapter One. In *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*, 1–76. Cambridge: Cambridge University Press, 1996. doi:10.1017/CBO9781139174763.003.