# Emotional Self-Knowledge

**Edited by Alba Montes Sánchez and Alessandro Salice**

**First published 2023**

## Chapter 1

## Affective Forecasting and Substantial Self-Knowledge

*Uku Tooming and Kengo Miyazono*

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

# 1 Affective Forecasting and Substantial Self-Knowledge

## Uku Tooming and Kengo Miyazono

### 1.1 Introduction

Self-knowledge comes in different varieties which seem to have little in common with one another. For instance, knowing that one is generous and knowing that one's legs are crossed have objects of different ontological categories, which differ in terms of difficulty and one has arguably much greater practical and moral importance than the other. Because of these differences, there is a reason to think that self-knowledge cannot and should not receive a uniform theoretical treatment. However, this also raises the question: along what lines should we distinguish between different kinds of self-knowledge?

In his book *Self-Knowledge for Humans*, Quassim Cassam has drawn a useful distinction between two kinds of self-knowledge, trivial and substantial. According to Cassam, while trivial self-knowledge (TSK) is easy to achieve but usually lacks practical importance for one's identity, substantial self-knowledge (SSK) requires cognitive effort and is considerably more important. If we accept this distinction, then knowing that one's legs are crossed would fall under TSK and knowing that one is generous would be a form of SSK.

Cassam has not fully answered the question as to why achieving SSK is more difficult than achieving TSK, however. Although he provides some useful gestures in that direction, there is still an explanatory gap. In this chapter, we are going to tackle that question. We will argue that the reason why cases of SSK are more difficult is that the evidential demands on SSK are such that a person can easily fail to have the relevant evidence. In particular, what is often needed is evidence about one's affective reactions and this is difficult to come by, given that people are easily mistaken about those reactions. They are easily mistaken because, as empirical evidence shows, they are prone to fail at affective forecasting.

The plan is as follows. In Section 1.2, we are going to describe the distinction between SSK and TSK in more detail. Then, in Section 1.3, we will argue that the substantiality of a case of self-knowledge correlates with the evidential demands of the case and that especially substantial cases

require evidence that one has the relevant affective reactions. Section 1.4 looks at empirical data on affective forecasting errors that people tend to make and suggests that the proneness to such errors explains why cases of SSK that require evidence for the relevant affective responses are especially difficult for them. Then, in Section 1.5, we will consider some paradigmatic cases of SSK—knowledge of one's character, values, and aptitudes—and show how our explanation of the difficulty of SSK applies to them. We will examine possible objections to our proposal in Section 1.6.

## 1.2  Trivial and Substantial Self-Knowledge

On the one hand, examples of TSK can easily be found in mainstream philosophical discussion of self-knowledge, e.g., knowing that you believe that it is raining outside, or knowing that you want to have another cup of chocolate ice cream. On the other hand, examples of SSK are rarely discussed in mainstream philosophical literature on this topic. Cassam (2014, 29) provides us with the following examples of SSK:

- Knowing that you are generous (knowledge of one's character).
- Knowing that you are not a racist (knowledge of one's values).
- Knowing that you can speak Spanish (knowledge of one's abilities).
- Knowing that you are a good administrator (knowledge of one's aptitudes).
- Knowing why you believe a controlled demolition brought down the World Trade Center on 9/11 (knowledge of one's attitudes in the 'knowing why' rather than in the 'knowing what' sense).
- Knowing that you are in love (knowledge of one's emotions).
- Knowing that a change of career would make you happy (knowledge of what makes one happy).

Cassam (2014, 30ff) distinguishes SSK from TSK by several conditions: the Fallibility Condition (it is always possible to be mistaken in one's self-ascription of SSK), the Obstacle Condition (there are obstacles to arriving at SSK), the Self-Conception Condition (acquiring SSK is entangled with one's self-understanding), the Challenge Condition (the claim to SSK can be challenged in ordinary contexts), the Corrigibility Condition (other people may be in a better position to know about substantive issues than the agent herself),[1] the Non-Transparency Condition (SSK can't be acquired through the transparency method), the Evidence Condition (SSK is based on evidence), the Cognitive Effort Condition (acquiring SSK is a matter of reflective work), the Indirectness Condition (SSK is psychologically and epistemically mediate), and the Value Condition (SSK has practical and often also moral importance).

Cassam admits that the distinction between TSK and SSK is a matter of degree, not of kind (Ibid., 29). All forms of self-knowledge are located

somewhere on a continuum between highly trivial cases (e.g., the case of knowing that you believe it is raining outside) on the one end and highly substantive cases (e.g., the case of knowing that you are generous) on the other end. Cassam also admits that the distinction between TSK and SSK does not neatly match up with the distinction between psychological categories, e.g., beliefs and desires on the one hand and character traits and values on the other. There might be somewhat substantive cases of knowing one's own beliefs or desires (see the discussion of the self-knowledge of desire in Section 1.3). And, possibly, there might be somewhat trivial cases of knowing one's own character traits and aptitudes (see the discussion of the inductive knowledge of one's character traits in Section 1.6).

Why is acquiring SSK more difficult than acquiring TSK, exactly? Cassam says that there is no single answer to this question, but his account focuses on his inferentialist theory of self-knowledge according to which "the knowledge of own beliefs, desires, hopes, and other 'intentional' states is first and foremost a form of inferential knowledge" (Cassam 2014, 137). But why is acquiring SSK more difficult than acquiring TSK according to inferentialism? Cassam's inferentialism is supposed to be an account of self-knowledge that includes both SSK and TSK; it is not the case that only TSK is inferentially acquired. Appealing to inferentialism itself does not explain the asymmetry between acquiring SSK and acquiring TSK. What we need to know is what those features of SSK are, due to which (inferentially) acquiring SSK is more difficult than (inferentially) acquiring TSK.

Cassam (2014, 194f) discusses several ways in which you can fail to (inferentially) acquire self-knowledge of attitude A:

a    You haven't performed the necessary inference from the evidence you have.
b    You lack the necessary evidence.
c    You have all the evidence you need but draw the wrong conclusion about whether you have A because (i) you reason poorly, (ii) you misinterpret the evidence, and (iii) you have a defective theory about the relationship between your evidence and your attitude.

Accounting for the asymmetry between TSK and SSK, (a) and (c) do not seem to be particularly informative. About (a), it is not clear why a person is more likely to fail to perform the necessary inference from the evidence in the context of acquiring SSK than in the context of acquiring TSK. For example, it is not clear why she is more likely to fail to perform the necessary inference from the evidence when she is figuring out whether she is generous than when she is figuring out whether she believes that it is raining outside. About (c), it is not clear why, when a person has all the evidence, she is more likely to draw wrong conclusions in the context of acquiring SSK than in the context of acquiring TSK. For example, it is not

clear why, when a person has all the evidence, she is more likely to draw wrong conclusions when she is figuring out whether she is generous than when she is figuring out whether she believes that it is raining outside.

The common issue with (a) and (c) as explanations of the asymmetry between SSK and TSK is that they are rather superficial. According to both (a) and (c), the difficulty of attaining knowledge pertains to the agent's deficient exercise of reasoning or inference in a given case. However, it is not very informative to say that one case is epistemically more difficult than the other just because agents more easily tend not to draw the relevant inference. What we want to know is *why* the agents tend not to draw the relevant inference.

For this reason, we think that an informative explanation of the asymmetry between SSK and TSK should be based on (b). Such an explanation should specify the kinds of evidence which are particularly difficult to gather, compared to the evidence that figures in the cases of TSK. This is not to say that (a) and (c) cannot contribute to explaining the asymmetry; (b) just provides a deeper explanation as to why (a) and (c) apply.

## 1.3  Failures of SSK as Failures of Evidence

In this section, we will flesh out the explanation of the type (b) and argue that how substantial a case is depends on the kind of evidence that is needed to arrive at self-knowledge. As a first approximation, case X is more substantial than case Y when an agent in X needs to possess the kind of evidence that an agent in Y does not need to make a knowledgeable self-attribution. In addition, we will argue that it is evidence about one's future and counterfactual affective reactions that makes a difference to the substantiality of a case.

Here it is useful to consider Krista Lawlor's (2009) conception of the self-knowledge of desire because Cassam makes extensive use of it in his analysis of SSK (Cassam 2014, 142). By looking at the Lawlor/Cassam account, we can clarify what it is about the evidential demands on a self-ascription that makes it a case of SSK.

According to Lawlor's model of causal self-interpretation, the typical way of coming to know what one wants is through causal inference from sensations, mental images, and natural language sentences. "Internal prompting" is a covering term for all these mental occurrences. An agent comes to the self-ascription of desire when, by experiencing internal promptings and actively rehearsing them, she postulates the desire as the hypothetical cause of those occurrences. Upon experiencing further internal promptings, the agent might reject the preliminary hypothesis, revise it, or confirm it, depending on whether she takes the further promptings to be good evidence for the hypothesis or not. The process is fallible and can go on for a long time, but sometimes the agent arrives at a self-ascription of desire that sticks, after which the agent counts

as knowing that he has the desire in question. According to Lawlor, the (fallible) markers of having identified one's desire include a sense that the self-ascription in question is unavoidable, given the pattern in one's internal promptings, and the self-ascription in question bringing about a change to the quality of one's internal promptings (Lawlor 2009, 60).

Lawlor's example of the internal promptings model in action concerns the case of Katherine who, upon looking at her son in his crib, hears in her head "Have another". This spontaneous inner speech utterance prompts a bout of self-scrutiny regarding other similar internal promptings that Katherine recently experienced and regarding the cause of those promptings. The hypothesis that Katherine makes is that the cause is her desire to have another child. The initial hypothesis need not settle the question of what she wants, however. She gathers further data by actively rehearsing and considering new imaginings about having children. After a while, she may experience a sense of ease with respect to the hypothesis and its explanatory strengths upon which she counts as knowing that she wants another child (2009, 59).

Katherine's case seems to satisfy sufficiently many conditions to count as SSK. Her self-knowledge is fallible (the Fallibility Condition), having a desire for another child bears on her self-conception (the Self-Conception Condition), the question of whether she wants to have another child is not transparent to the question of whether having another child is desirable (the Non-Transparency Condition), the self-ascription of a desire to have another child is based on evidence (the Evidence Condition), Katherine exercises cognitive effort to arrive at the right answer (the Cognitive Effort Condition), Katherine's self-knowledge is neither psychologically nor epistemically immediate but is based on inference (the Indirectness Condition) and knowing that she wants to have another child matters both practically and morally (the Value Condition).

If the SSK and TSK are distinguished by evidential circumstances, what is it about Katherine's evidential circumstances that make her case a case of SSK? We propose that it is a case of SSK because Katherine faces a challenge of getting enough good evidence and avoiding misleading evidence when inferring her desire from her internal promptings. Suppose that she does want to have another child, but when she imaginatively rehearses possible scenarios involving children, those imagined scenarios do not constitute unambiguous evidence that she has the desire to have another child. This situation is not unrealistic, given that one's imaginings can be elaborated in different directions and can carry redundant information. In addition, the evidence that imagination provides is not only evidence about one's mental states, but also evidence for facts about the actual world (Kind 2016) and about possibilities (Yablo 1993), and an agent might mix up different kinds of evidence. Furthermore, the feelings that those imaginings generate can often be ambiguous and indicate attitudes other than the desire to have another child. The imaginative

rehearsal that Katherine engages in does not guarantee getting access to sufficient evidence for the belief that she wants to have another child. As a result, she may fail to arrive at a correct self-ascription of desire.[2] It thus makes sense to think of Katherine's case as a case of SSK and it is reasonable to think that it is the challenge of getting enough good evidence and avoiding misleading evidence for one's self-ascription that makes it SSK. This supports the idea that the difficulty of SSK in contrast with TSK derives from the evidential situation that the agent is in.

So far so good. What we now want to suggest, though, is that there is a further kind of evidence bearing on SSK that one needs to consider. Notice that all the relevant evidence that Lawlor (and Cassam) considers is evidence for the mental causes of one's present mental occurrences. There are cases, however, in which such evidence is not sufficient for one to attain self-knowledge. There are targets of SSK, knowledge of which requires evidence about one's affective reactions to possible future events. Consequently, self-knowledge of these targets is more substantial than self-knowledge that does not require evidence about one's affective reactions.

To see what targets we have in mind, consider the following elaboration of Katherine's case. Suppose that Katherine also asks herself whether she *values* having another child and, after careful consideration (say, by analyzing the possible causes of her internal promptings as in the case of desire), concludes that she does. Then, at some point in the future, that prospect finally materializes and Katherine gives birth to a baby girl. To her unpleasant surprise, however, she discovers that she feels disappointed and torn about the decision that she made. The new child does not bring any of the contentment that she expected her to bring. Oftentimes, she even finds herself mumbling to herself: "That's not what I wanted". She is embarrassed about those feelings, but they are unambiguous in suggesting that having another child was nothing like she imagined it to be, insofar as her affective responses are concerned.[3]

In such a situation, we think that it is reasonable to think that when Katherine turns out to be disappointed in having another child, this indicates that her evaluative attitudes toward that prospect are fickle: before giving birth, she seemed to value having another child, while Katherine's subsequent affective responses indicate that her valuing is at most only surface-level or that she did not value it in the first place. It makes sense to say that the disappointment defeats her earlier belief that she valued having another child. Our precursory explanation of this is that valuing involves having appropriate affective dispositions that express one's evaluative outlook.[4] If Katherine takes herself to value having another child, her affective reactions give a reason to think that she is mistaken because they indicate that she does not have the affective disposition that is required for valuing. The analysis of the (elaborated) Katherine's case lets us see that there are domains of self-knowledge where getting one's future affective reactions right is immediately relevant for having self-knowledge.[5]

It seems, then, that there are cases of SSK, such as knowing one's values,[6] where knowing one's affective reactions to future events is necessary.[7] Prima facie, since they require more evidence than cases of SSK for which knowing one's affective reactions is not necessary, acquisition of self-knowledge in the former cases faces even more obstacles in the latter cases and it makes sense to think that they are more substantial. To flesh out our proposal, we need to give an account of what it is about the knowledge of one's affective reactions that makes it so difficult to achieve. In Section 1.4, we are going to argue that the difficulty of arriving at self-knowledge in those cases stems from the fact that knowing one's affective reactions requires affective forecasting which people tend to be bad at. Our proposal is that the evidence about how one is disposed to affectively react is difficult to gather, given people's propensity to error in affective forecasting. Because of this, affective forecasting is not a reliable source of evidence.

## 1.4  Affective Forecasting: General Theory

Affective forecasting involves predicting one's affective reactions to future events and is arguably a typical procedure that people resort to when making decisions. For instance, take Timo, who is deciding between different holiday destinations. A natural way to make an informed decision is for him to compare how he would feel with respect to the different options. Let's say his options are Paris and Tallinn. In order to decide where to go, Timo then engages in affective forecasting. With respect to Paris, he predicts that he would feel very excited by being in a place that he has wanted to visit for years. With respect to Tallinn, he predicts that he would feel bored,[8] given that he has lived there in the past and is well familiar with its main attractions. By comparing those predictions, Timo then decides to go to Paris because it seems to him to provide more affective payoff than Tallinn would.

Unfortunately, there is now a substantial body of research in psychology which shows that people are prone to robust patterns of error in affective forecasting. While they are generally not mistaken about the valence of their affective reaction—i.e., whether it would be pleasant or unpleasant—they are less accurate about the intensity and duration of their reaction (Wilson and Gilbert 2005). These mistakes can express themselves both in overestimation and underestimation. Overestimations are known as *impact bias*. For instance, people overestimate how bad they would feel if they rejected good advice regarding a sports wager (Crawford et al. 2002); academics tend to overestimate how sad they would feel upon not getting tenure and college students overestimate the affective impact of the dissolution of a romantic relationship (Gilbert et al. 1998). Examples of underestimation include underestimating the affective impact of social pain from ostracism and shame (Nordgren, Banas, and MacDonald 2011),

underestimating emotional pain caused by ambivalent sexism (Bosson, Pinel, and Vandello 2010), and underestimating emotional reaction to an outcome of a foregone gamble (Andrade and Van Boven 2010).

The most prominent causes of impact bias are focalism and immune neglect. Focalism is the tendency not to consider the impact of events other than the predicted event on one's affective condition (Schkade and Kahneman 1998). As a result, one's forecast exaggerates the affective impact of the focal event. Immune neglect concerns the tendency to not consider the effect of emotional coping mechanisms on one's circumstances where initial strength of the affective reaction diminishes over time (Hoerger et al. 2009). Since the coping mechanisms function mostly unconsciously (Wilson and Gilbert 2005, 133), it is no wonder that people do not consider them when they engage in affective forecasting.

Take Timo again. Given an all-too-human proneness to errors in affective forecasting, Timo might easily overestimate not only the feelings of excitement with respect to visiting Paris, but also the feelings of boredom with respect to Tallinn. As a result, the forecasts get skewed and Timo's comparison between Paris and Tallinn turns out to be mistaken. Had Timo's forecast been accurate, he might have even decided to visit Tallinn instead of Paris.

Some data indicate that inaccuracies in predicting the intensity of affect are not pervasive: people tend to misinterpret the question of how they would feel as a question of how intensely they would feel about a particular event, whereas when the event comes about, they are asked to rate the intensity of their feelings in general (Levine et al. 2012).[9] As a result, the earlier prediction can understandably be mistaken because it answers a different question than the evaluation at the time when the event has occurred. On the other hand, people's predictions are more accurate when both at the time of the prediction and at the time when the predicted event has occurred, they answer the question about their feelings about that particular event (Doré et al. 2016). A more nuanced take on people's capacity for affective forecasting was also suggested by a recent study by Lench et al. (2019) where college students were asked how they would feel in terms of intensity, frequency and impact on mood if they got a grade that was higher/lower/the same as they expected. As it turned out, they were less biased in predicting the intensity of the feeling than predicting its frequency or its impact on mood. Affective forecasts regarding the results of the 2016 U.S. presidential election delivered similar results.

For our purposes, however, it suffices to take on board the idea that people's affective forecasts are often erroneous, especially with respect to the duration of the predicted emotion. The arguments in favor of the view that people tend to accurately predict the intensity of the emotional reaction upon the time of the predicted event have not challenged the claim that people are easily mistaken about their affective responses across wider timescales.

The difficulty of affective forecasting might have something to do with some general features of imagination. In forecasting our affective response to a future scenario, we tend to imagine some specific aspects of the scenario, abstracting away some details that can make a difference. Maibom writes; "we often do not imagine in enough detail. Our imagination is usually constrained in a variety of ways. If, for instance, our mother instructs us to imagine how we would feel if someone were to hit us over the head with a stick and take our toy, this is exactly what we will imagine: being hit over the head and deprived of our toy. Quite likely we'll imagine little else, such as where we were being hit, who the other child was, why they might have been induced to hit us over the head, and so on. The event will not be situated the way that events in our lives are" (Maibom 2016, 192).[10]

Let's now return to SSK. Why is self-knowledge that concerns knowing one's affective dispositions especially difficult? This is because in order to know one's affective dispositions one has to know what affective responses one would have across a large variety of possible circumstances. One therefore needs to predict those responses, i.e., one needs to resort to affective forecasting in which, as the data indicate, people tend to fail. If people easily fail at accurately predicting how they would feel about a future or counterfactual event and SSK requires knowing one's affective reactions across a variety of possible circumstances, then it follows that people easily fail to arrive at SSK.

It could be objected that affective forecasting research only shows that we are bad at predicting the specific features of our affective responses (e.g., duration and intensity) but it does not show that we easily fail at predicting the kinds of emotions that we are disposed to have. We admit that it is consistent with the available evidence in that field that errors in affective forecasting rarely extend to the identification of types of emotion and that our ability to predict what kinds of emotion we would feel is quite robust. However, in the present context, it is exactly the difficulty of predicting the specific features of affective responses that is at issue. SSK concerns targets that involve dispositions to have affective responses of appropriate duration and intensity. In what follows, we will consider such targets more closely and show how the proneness to affective forecasting errors regarding features of our affective responses explains their difficulty.

## 1.5 Applications

Affective forecasting is a key factor in acquiring SSK, but we do not claim that it explains all cases of SSK. We agree with Cassam that there is probably no single explanation of the difficulty of acquiring SSK. There are multiple factors, and they work differently in different cases of SSK. For instance, perhaps affective forecasting has little to do with

the SSK of the ability to speak Spanish. We predict, however, that SSK that does not involve much affective forecasting is less substantial than SSK that does involve affective forecasting. In other words, the difference between strongly substantial SSK (of character, values, etc.) and weakly substantial SSK (of linguistic abilities, etc.) can be explained by the relevance of affective forecasting.

### 1.5.1  Character

One of the central kinds of SSK that Cassam considers is the knowledge of one's character.[11] Taking that kind of knowledge to constitute SSK is *prima facie* very plausible, given that answering questions such as whether I am lazy or industrious, courageous or cowardly, generous or stingy doesn't look like a trivial matter. An agent is not immediately justified in thinking that she has a particular character trait; instead, knowing whether she has it seems to require considerable self-reflection.

Why is knowing one's character difficult? In the case of many character traits, the explanation in terms of the difficulty of affective forecasting looks very plausible. This is because one of the central dispositions that make up a character trait is a disposition to have affective reactions across a variety of circumstances and, due to people's proneness to errors in affective forecasting, knowing whether one has the relevant affective disposition is challenging.

Think about empathy as a character trait.[12] Since being empathic is partly constituted by feeling the right kind of emotions toward people in need, it is a crucial part of your figuring out whether you are empathic or not that you forecast your affective reactions to people in need. As a matter of fact, however, you are not very good at affective forecasting; you mistakenly predict that you will feel the right kind of emotions toward people in need and thus attribute yourself the character trait of empathy. This is one way in which you mistakenly conclude that you have that trait.

Cassam's own example of knowing one's character as a form of SSK concerns fastidiousness, which is a trait that involves caring about tidiness and things being in order. As an example of a person being fastidious, he considers a character called Woody who always keeps his surroundings clean and tidy (Cassam 2014, 176). How can Woody come to know that he is fastidious? According to Cassam:

> Here is how Woody might come to know that he cares about such things as tidiness and attention to detail, and that he is bothered by their absence: when he imagines the state of his teenagers' bedroom he is conscious of feeling a mixture of dismay and irritation. The same mixture of dismay and irritation is prompted by the recollection that he didn't have time to tidy his desk when he finished work

yesterday, and he is conscious of a desire to put things right as soon as possible. When he thinks about what needs to be done tomorrow, he focuses on what he sees as the need to restore order. He knows that his work colleagues aren't nearly as meticulous as he is, and is conscious of thinking thoughts along the lines of "if you want something done right, do it yourself". On the basis of his thoughts, imaginings, and emotions Woody is in a position to conclude that he cares about cleanliness and attention to detail.

<div align="right">(Ibid., 177)</div>

To know whether he is fastidious, Woody needs to know to what extent he cares about tidiness and cleanliness and this in turn requires knowledge about his affective reactions with respect to the relevant situations where such cares are manifested. Our explanation of the difficulty of SSK is thus easily applicable to Cassam's central example of SSK. The difficulty of knowing whether one is fastidious is due to the difficulty of knowing one's affective dispositions.

What about character traits that do not manifest themselves in affective reactions? We think that their number is probably quite limited, given that some kind of affective disposition seems to be a necessary condition for most character traits. That being said, there are some candidates. For instance, the traits of diligence and conscientiousness are plausibly less tied to affective dispositions than many other traits. In a Korean TV series *Stranger* a.k.a. *Secret Forest*, the main character, prosecutor Hwang Si-mok is incapable of feeling emotions but outshines his colleagues by not letting himself be corrupted by the influence of money and power. He wholeheartedly commits himself to solving the criminal cases and it makes sense to take him to be diligent and conscientious. If the series is psychologically realistic, then it provides us with some examples of character traits that do not require affective reactions.

The acknowledgement of "affectless" character traits doesn't falsify our explanation of SSK, however. It is important to stress here that we allowed that the difficulty of SSK comes in degrees and that our explanation is meant to apply to more substantive types of self-knowledge. In the case of character traits, our claim is that what makes a difference to the difficulty of knowing a type of a character trait is whether knowing that one has it requires relying on affective forecasting. If diligence, or at least some kind of diligence, does not involve affect, our prediction is that self-knowledge of that kind of diligence is less difficult and less substantive than self-knowledge of those character traits that involve affect.

A subset of knowing one's character traits that are especially difficult to know involves self-knowledge of virtue. Think about the virtue of courage, for example. Since the trait of being courageous is partly constituted by not being too afraid of risks and dangers, it is a crucial part of your figuring out whether you are courageous or not that you

forecast your affective reactions to risks and dangers. As a matter of fact, however, you are not very good at affective forecasting; you mistakenly predict that you will not be too afraid of risks and dangers and thus attribute yourself the virtue of courage. This is one way in which you mistakenly conclude that you have the virtue of courage.

Our account can provide an explanation, admittedly a very speculative one, as to why self-knowledge of virtue stands out in difficulty among traits that are not virtues if it is conjoined with the Aristotelian theory of virtue, according to which being virtuous involves being disposed to act in accordance with the golden mean, by avoiding extremes that constitute vices (*Nicomachean Ethics* II.6). For instance, courage is the intermediate between the vices of cowardice and recklessness. Achieving the golden mean in one's conduct requires a properly balanced affective response to the situation at hand: the response should neither be too strong nor too weak and it should be of appropriate duration. Since people are generally bad at forecasting the intensity and duration of their affective response, they are easily mistaken in their estimations of whether they would act in accordance with the golden mean. As a result, they are also easily mistaken in evaluating their virtuousness.[13]

A similar explanation can also be applied to epistemic virtues. For instance, being epistemically virtuous arguably requires sufficient motivation for truth (Montmarquet 2019; Zagzebski 2003). For instance, when a person cares little about whether her beliefs are true or not, her motivation for truth is too weak for her to count as epistemically virtuous, but there is also another extreme where a person is so obsessed about getting things right that she checks her beliefs even when there is no need to check any further.[14] If knowing whether one is sufficiently motivated for truth requires knowing the intensity and duration of one's affective reactions in relevant situations, then failures in affective forecasting can also hinder an agent from knowing whether she is epistemically virtuous.

There are also more specific epistemic virtues in which case it is plausible that self-knowledge of them requires knowing that one has a disposition to have affective responses within appropriate limits. These include: intellectual perseverance (requires a disposition not to be too easily discouraged by intellectual obstacles), open-mindedness (requires a disposition not to be too easily offended by different viewpoints), curiosity (requires a disposition to get properly emotionally invested in inquiry), and carefulness (requires a disposition not to be carried away by preliminary evidence in favor of one's own positions). In all those cases, one has to exercise moderation in one's affective responses.

The literature on virtue is expansive and the disagreements over its nature run deep. Here we do not pretend to have presented a full account of the self-knowledge of virtue. All we can claim is that our explanation of the difficulty of knowing one's virtues as a form of SSK goes smoothly

with one, historically prominent and up to this day popular conception of virtues, both moral and epistemic.

### 1.5.2 *Values*

We already saw above how knowing one's future affective reactions with respect to X was relevant for knowing whether one *valued* X, where valuing of X was taken to require consistency in affective reactions with respect to X over time. In this section, we will expand on this idea.

To get a better grip on what we talk about when we talk about values, Anderson's definition is helpful here:

> To value something is to have a complex of positive attitudes toward it, governed by distinct standards for perception, emotion, deliberation, desire, and conduct. People who care about something are emotionally involved in what concerns the object of care. Parents, who love their children will normally be happy when their children are successful and alarmed when they are injured. They will be alert to their needs, take their welfare seriously in their deliberations, and want to take actions that express their care.
>
> (Anderson 1993, 2)

Also, Scheffler's:

> valuing involves a distinctive fusion of reason and emotion. It comprises a complex syndrome of interrelated dispositions and attitudes, including, at least, certain characteristic types of belief, dispositions to treat certain kinds of considerations as reasons for action, and susceptibility to a wide range of emotions.
>
> (Scheffler 2010, 29)

Although valuing cannot be reduced to affective dispositions,[15] valuing involves a disposition to experience a variety of emotions (see Helm 2001, this volume). What emotional responses a type of valuing is disposed to manifest depends on the domain of things that are valued. Valuing one's children manifests itself in different emotional responses than valuing basketball, for instance.[16]

Valuing is also closely related to caring:

> Typical components of caring include joy and satisfaction when the object of one's care is doing well and advancing and frustration over its misfortunes or setbacks, anger at agents who heedlessly cause such misfortunes or setbacks, pride in the successes for the object and disappointment over its defeats or failures, the desire to help ensure those successes and to help avoid the setbacks, fear when the

object is in jeopardy and relief when it escapes untouched, and grief
at the loss of the object and the subsequent nostalgia.

<div align="right">(Jaworska 2007, 483)</div>

What valuing also requires is that, depending on a situation, the affec-
tive response must have the appropriate valence, the appropriate level of
arousal that is within the appropriate range and appropriate duration.
For instance, for A to value another person, B, as a friend, A should
manifest sadness upon B's death (valence condition) that is of appro-
priate intensity (arousal condition) and that does not immediately fade
(duration condition). If A didn't satisfy one or more of those conditions,
it would put into question whether she actually values B as a friend. Of
course, we can think of possible cases in which one or more of those
conditions are not met but then it is plausible to think that there are
some excuses available for A. For instance, A can be excused from not
grieving B's death and still count as valuing B as a friend, if she is under
extreme duress and needs to harness all her emotional responses for
dealing with it. As another example, if A values B as a friend, we also
expect her to enjoy B's company (at least for most of the time), where the
enjoyment is of appropriate intensity and duration. Admittedly, these
conditions on enjoyment are relatively lax, but if A's enjoyment were
very superficial and fleeting or if A would not enjoy B's company at all,
she would not count as someone who values B as a friend.

In the case of valuing, then, it seems that valuing something, X,
involves, among other things, a disposition to emotionally respond to
changes that affect X across a variety of circumstances. Putting this in
the context of self-knowledge, knowing what one values seems to require
knowing that one is disposed to emotionally respond in those ways that
are relevant for valuing and, given the issues with affective forecasting,
that kind of knowledge is difficult to come by.

In his book, Cassam also considers how people come to know their
values. His example considers knowing that one is not a racist (Cassam
2014, 178). In particular, one needs to rule out instinctive racism, which
is a matter of feeling moral solidarity with members of one's own race
at the expense of other races (Taylor 1985, 61). According to Cassam,
internal promptings play a key role here exactly because racism is a mat-
ter of feeling, among other things. What Cassam does not sufficiently
acknowledge, however, is that harboring a value also requires the dispo-
sition to continue having certain feelings that are characteristic to that
value in the future. An agent does not know that he is an instinctive rac-
ist unless he is justified in thinking that the feelings of partiality toward
one's own race won't seep into his affective and evaluative outlook over
time. Having a particular value is demanding in the sense of requir-
ing consistency over time, including in the future. Knowledge of values
therefore requires the ability to predict one's affective reactions.

### 1.5.3 Aptitudes

On the face of it, you might think, self-knowledge of aptitudes is beyond the scope of our account, e.g., it is not clear how knowing that you are a good administrator is related to affective forecasting. To know whether you are a good administrator, you can just think about whether you have some non-affective skills and dispositions, such as PC skill or efficiency. And knowing these relevant skills and dispositions does not require affective forecasting.

This, however, is not the case. It is a crucial part of being a good administrator that you exhibit the right kind of affective responses in the right kind of conditions. And knowing about the former does require knowing about the latter. In other words, the failure of knowing about the latter leads to the failure of knowing about the former. To see this, think about the following case. Anna thinks that she has the aptitude for being a good administrator, which is why she decides to take up an administrative job after graduation from college. She is efficient, responsive, and hard working. She has strong leadership and management skills. She was a star student at high school and college and highly evaluated by teachers, professors, and fellow students. It turns out, however, that she is not a very good administrator despite her excellent skills and dispositions. Her problem is an affective and interpersonal one; she is extremely jealous, which causes serious troubles especially in the context of working collaboratively with co-workers. When somebody in her office achieves a remarkable result or is promoted (especially when the person is not as good as Anna, at least according to Anna's own standard), Anna gets extremely frustrated and becomes unable to concentrate on her work, which has a considerably negative effect on her actual performance. Anna does understand that she is relatively new in her office, and it would take some time for others to understand how good and talented she is and for her to receive the evaluation and reputation she deserves. But her strong jealousy overwhelms her every time she learns about somebody else's achievement and promotion. Anna also has troubles with her boss, who, according to Anna's standard, is not as good as she is. Anna does not respect nor trust the boss, which makes it very difficult for them to work together as a team. Outside the office, Anna always complains about the fact that the boss, who is inferior, earns much more than her, who is superior. After a year of affective and interpersonal struggling, she eventually quits her job without achieving much.

Anna's case shows that it is crucial for one to know about one's aptitude for being a good administrator that one goes through some affective forecasting. Anna's failure of self-knowledge is due to her failure of forecasting her own affective reactions to co-workers' achievements and promotions.

We think that this is not only true about the particular aptitude for being a good administrator, but also for many other aptitudes. It is crucial

for knowing about one's aptitude for being a good teacher that one correctly predicts one's affective reactions to troubling students or demanding parents; it is crucial for knowing about one's aptitude for being a good musician that one correctly predicts one's affective reactions when one needs to perform in front of a large audience; it is crucial for knowing about one's aptitude for being a good nurse that one correctly predicts one's affective and empathic reactions to suffering patients, etc.

Here is a possible objection. One might think that there is a crucial difference between the case of virtues and the case of aptitudes. It is relatively uncontroversial that having the right kind of emotional reaction to danger is a constitutive part of the virtue of courageousness. It is not clear, however, that having the right kind of emotional reaction to somebody else's achievements and promotions is a constitutive part of the aptitude for being a good administrator. The latter is constituted by non-affective skills and dispositions. If having the right kind of affective reaction is not constitutive, then one can insist, against our stipulation, that Anna actually has the aptitude for being a good administrator, and she has the accurate self-knowledge about it. After all, she has the non-affective skills and dispositions that are constitutive of aptitude for a good administrator. She fails as an administrator as a matter of fact, but the failure is due to some unfortunate causal factor. You might think, for instance, that her aptitude for being a good administrator was causally "masked" in those cases.

We do not dispute whether affective reactions are constitutive of the aptitude for being a good administrator or they are merely causal for it. Either way, our proposal can be defended. Suppose that affective reactions are merely causal for the aptitude for being a good administrator. The aptitude for being a good administrator is constituted only by non-affective skills and dispositions. If we conceive of aptitudes in this way, however, the SSK of aptitudes is similar to the SSK of ability to speak Spanish, which is only weakly substantial. It is certainly true that, in this case, Anna did have an accurate self-knowledge that she has the aptitude for a good administrator despite her failure in the real workplace, but this is simply because the self-knowledge is not very substantial. Suppose, in contrast, that affective reactions are not merely causal but rather constitutive of the aptitude for being a good administrator. The aptitude for being a good administrator is constituted not only by non-affective skills and dispositions but also affective responses. If we conceive of aptitudes in this way, however, the SSK of aptitudes is deeply substantive. In this case, Anna had inaccurate self-knowledge that she has the aptitude for a good administrator, and her failure is due to the difficulty of affective forecasting.

In short, affective reactions can be causal for or constitutive of aptitudes. SSK of aptitudes is deeply substantive in the latter case, and the substantiveness is due to the relevance of affective forecasting. SSK of aptitude is less substantive in the former case, and our proposal does not have much to say about it.

## 1.6  Objections

### 1.6.1  Objection 1: Functionalist Objection

Does our account really explain the difference between TSK and SSK? Perhaps affective forecasting is relevant not only for SSK but also for TSK, such as TSK of my belief that it is raining outside, or the TSK of my desire for another cup of chocolate ice cream. Given functionalism or dispositionalism, believing that it is raining outside is partly constituted by some affective reactions, such as being surprised when it turns out that it is sunny outside (see Schwitzgebel's [2002] phenomenal dispositionalist account of believing).

The functionalist objection clarifies an important issue; strictly speaking, what makes self-knowledge of X substantive is not the *metaphysical* fact that X is (partly) constituted by affective reactions. Even if beliefs and desires are partly constituted by affective reactions in the way functionalists or dispositionalists claim, it does not necessarily mean that self-knowledge of beliefs and desires is always substantial (although, as we already pointed out, that self-knowledge of beliefs and desires can be substantial in some cases).

Rather, what makes self-knowledge of X substantial is the *epistemological* fact that knowing about relevant affective reactions is (at least typically) indispensable for knowing about X. The metaphysical fact that X is (partially) constituted by affective reactions is certainly relevant to the substantiality of the self-knowledge of X but only in an indirect manner, via the epistemological fact about X, i.e., the epistemological fact about X obtains, when it does, often in virtue of the metaphysical fact about X. For instance, knowing about relevant affective reactions is indispensable for knowing about character traits in virtue of the fact that character traits are constituted by affective dispositions. However, the metaphysical fact about X does not necessarily imply the epistemological fact about X. According to functionalism or dispositionalism, believing that it is raining outside and desiring another cup of chocolate ice cream are constituted by some affective reactions. But it does not seem to be the case that knowing about relevant affective reactions is indispensable for knowing about beliefs about raining outside or desires for another cup of chocolate ice cream. Typically, we identify our belief about raining outside or desire for another cup of chocolate ice cream in a direct and immediate way without going through affective forecasting.

For a similar reason, Goldman (1993, 2006) rejects the idea that we acquire self-knowledge of belief by tracking its functional role or dispositional profile; self-knowledge of belief is direct and immediate, while functional or dispositional properties (including the property of causing relevant affective responses) cannot be detected directly or immediately. In response, one might deny the idea of direct and immediate access to one's own beliefs and think that self-knowledge in general is tracking functional

or dispositional properties (e.g., Carruthers 2009, 2011). But the skepticism about direct and immediate access is empirically and theoretically controversial (see Engelbert and Carruthers [2010] for an overview), and one of us (Miyazono forthcoming) presents an argument against it; such a skeptical view fails to explain the self-ascription of belief by people with delusion; they regard their delusions to be their beliefs, while their delusions often lack functional or dispositional properties of belief.

### 1.6.2  Objection 2: Inductive Objection

Knowing about relevant affective reactions is (at least typically) indispensable for knowing about courage. Even if this is true, do we really need affective *forecasting* in order to know about relevant affective reactions? Perhaps in some cases, we have some past experience of affective reactions to dangers that are enough to support an inductive inference about a conclusion. Perhaps somebody has been constantly fearless in the face of dangers in the past, and he can inductively come to the conclusion that he is courageous.

There are, however, some reasons to think that the past experience is not sufficient for knowing about courage at least in many cases. Firstly, some of the targets of SSK, such as character traits and virtues, are supposed to be stable and constant across time and situation. For this reason, it is likely that in many cases the limited past experiences in limited situations are not sufficient for knowing about character traits and virtues; you also need to think about different possible future situations.

Secondly and relatedly, past experiences are open to different interpretations. Even if you have constantly been fearful in the face of dangers in the past, it is still possible to explain it away by saying that past situations were not quite right. You might argue that you will be different the next time, in the right kind of situation. And we need to appeal to affective forecasting in order to test these hypotheses.

That said, we do think that past experience can be sufficient at least for some people in some cases. It is possible that elderly people have sufficient past experience to inductively know about their own characters, values, aptitudes, etc., while young people do not have sufficient past experience and thus need to appeal to affective forecasts. This is consistent with the commonsensical observation that knowing about one's own characters, values, aptitudes, etc. is a more serious and difficult issue for younger people than for elderly people. It is typically young people rather than elderly people who seriously contemplate their own characters, values, aptitudes, etc. For instance, a young army soldier with no real combat experience might reasonably wonder whether he is courageous enough for being a good soldier, while a decorated veteran with rich combat experience might reasonably be confident about his courageousness. The difference between them can be explained by the

fact that the veteran can appeal to induction, while the young soldier needs to appeal to affective forecasting.

## 1.7 Conclusion

In our chapter, we discussed Cassam's useful distinction between SSK and TSK and proposed that what makes a case of self-knowledge substantive is the evidential situation of the agent. In the case of SSK, the agent needs evidence that is more difficult to gather than evidence in the case of TSK. We then argued that within the domain of SSK, there are more or less substantial forms of self-knowledge and that the more substantial cases are those in which the relevant evidence includes facts about one's affective reactions. Since knowing these facts requires engaging in affective forecasting and it is well established that people easily fail at affective forecasting, our account provides a neat explanation as to why some cases of SSK are especially difficult to acquire.

## Acknowledgments

## Notes

1. In fact, others may play a crucial role in making such self-knowledge possible in the first place (see Edward Harcourt's contribution to the present volume).
2. See Tooming (2020) for a complementary analysis of what makes Katherine's case a case of SSK. According to Tooming, it is the lack of experiential familiarity with the content of the self-ascribed desire that makes it difficult to know if one has the desire. The lack of such familiarity can contribute to the explanation as to why the evidential circumstances were poor in Katherine's case.
3. Since the feeling that one's child is a mistake carries extremely negative moral overtones, the reader can substitute this example with a more neutral one. What matters is that we have a case in which an agent took herself to value something and that self-ascription of value informed her decision, but she turns out to be disappointed in the outcome of her decision and to regret it.
4. We look at self-knowledge of values more closely below, in Section 5.2.
5. Not all affective reactions are equally relevant, of course. There are differences in the extent to which an affective response reveals something about ourselves (see Krista Thomason's contribution to this volume).
6. When we talk about values here, we take it to be a shorthand for the subjective attitude of valuing.

7. We leave open as to whether knowing one's affective dispositions is also necessary for the self-knowledge of desire because the answer to this question depends on what conception of desire one has in mind. We prefer not to take a stand on this.

8. On boredom and self-knowledge, see Antonio Gómez Ramos's contribution to this volume.

9. For some doubts about Levine et al.'s arguments, see Wilson and Gilbert (2013).

10. Another, more speculative, idea as to why affective forecasting is prone to failure is that an agent predicts her emotional reaction by imagining it, in which case she can easily misattribute her actual emotional response during the imagining to the content of the imagining. In imagining the reaction, one's perspective is doubled, in that there is the imagined affective reaction to the imagined scenario, but there is usually also the actual affective reaction to the imagining (see Goldie 2012, 80). One's emotional response to one's affective forecast regarding an imagined scenario can be either exaggerated or subdued when compared to the emotional response that one would actually have to the scenario in question. When the agent takes the former to reflect the latter, she can thus be easily mistaken.

11. Cassam acknowledges the situationist critique of the idea that there are character traits as dispositions that robustly explain people's actions. However, he points out that such a critique still lets us maintain that character traits can explain behavior at least sometimes and that the critique hasn't really disputed the existence of character-constituting dispositions other than dispositions to act, such as dispositions to think, feel, and want in certain ways (Cassam 2014, 174f).

12. Here we are talking about empathy as a character trait rather than empathy as a state, e.g., "You may have a general disposition to feel for others in need, and this trait may have amplified the empathic concern you felt for your friend" (Batson 2018, 33). For different senses of "empathy", see Batson (2018) and Maibom (2012).

13. Interestingly, among philosophers of antiquity, there was a divide between those who thought that affective states were necessary for virtue (e.g., Plato and Aristotle) and those who said that they weren't (e.g., Socrates and the Stoics) (see Homiak, 2019). Needless to say, we align ourselves with Plato and Aristotle.

14. Compare this with Kelahan's (2018) interpretation of Hume's mitigated skepticism as a golden mean.

15. Arguably, valuing X also involves the experience of grasping the value of X, and that experience has a distinctive phenomenological character. However, if a person who has that experience is not disposed in any way to affectively respond to X, it seems counterintuitive to take that person to actually value X. We thank Íngrid Vendrell Ferran for pressing us on this issue.

16. Cassam, in contrast, appeals to Lewis's conception of valuing, according to which it is a kind of second-order desire (Cassam 2014, 179). Under such a conception, affective disposition might not seem immediately relevant for valuing. However, we find the more complex notion of valuing more realistic.

## References

Anderson, Elizabeth C. 1993. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.

Andrade, Eduardo B., and Leaf van Boven. 2010. "Feelings not forgone: Underestimating affective reactions to what does not happen." *Psychological Science* 21(5): 706–11.

Batson, C. Daniel. 2018. *A Scientific Search for Altruism: Do We Only Care about Ourselves?* Oxford: Oxford University Press.

Bosson, Jennifer K., Elizabeth C. Pinel, and Joseph A. Vandello. 2010. "The emotional impact of ambivalent sexism: Forecasts versus real experiences." *Sex Roles* 62(7): 520–31.

Carruthers, Peter. 2009. "Mindreading underlies metacognition." *Behavioral and Brain Sciences* 32(2): 164–82.

Carruthers, Peter. 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge.* Oxford: Oxford University Press.

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

Crawford, Matthew T., Allen R. McConnell, Amy C. Lewis, and Steven J. Sherman. 2002. "Reactance, compliance, and anticipated regret." *Journal of Experimental Social Psychology* 38(1): 56–63.

Doré, Bruce P., Robert Meksin, Mara Mather, William Hirst, and Kevin N. Ochsner. 2016. "Highly accurate prediction of emotions surrounding the attacks of September 11, 2001 over 1-, 2-, and 7-year prediction intervals." *Journal of Experimental Psychology* 145(6): 788–95.

Engelbert, Mark, and Peter Carruthers. 2010. "Introspection." *Wiley Interdisciplinary Reviews: Cognitive Science* 1(2): 245–53.

Gilbert, Daniel T., Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley. 1998. "Immune neglect: A source of durability bias in affective forecasting." *Journal of Personality and Social Psychology* 75(3): 617–38.

Goldie, Peter. 2012. *The Mess Inside: Narrative, Emotion, and the Mind.* Oxford: Oxford University Press.

Goldman, Alvin I. 1993. "The psychology of folk psychology." *Behavioral and Brain Sciences* 16(1): 15–28.

Goldman, Alvin I. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading.* Oxford: Oxford University Press.

Helm, Bennett W. 2001. *Emotional Reason. Deliberation, Motivation, and the Nature of Value.* Cambridge: Cambridge University Press.

Hoerger, Michael, Stuart W. Quirk, Richard E. Lucas, and Thomas H. Carr. 2009. "Immune neglect in affective forecasting." *Journal of Research in Personality* 43(1): 91–4.

Homiak, Marcia. 2019. "Moral Character." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2019/entries/moral-character/.

Jaworska, Agnieszka. 2007. "Caring and full moral standing." *Ethics* 117(3): 460–97.

Kelahan, Emily. 2018. "Naturalness and artificiality in Humean virtue theory." *Hume Studies* 44(2): 249–73.

Kind, Amy. 2016. "Imagining under constraints." In *Knowledge through Imagination*, edited by Amy Kind and Peter Kung, 145–59. Oxford: Oxford University Press.

Lawlor, Krista. 2009. "Knowing what one wants." *Philosophy and Phenomenological Research* 79(1): 47–75.

Lench, Heather C., Linda J. Levine, Kenneth Perez, Zari Koelbel Carpenter, Coleen D. Carlson, Shane W. Bench, and Yidou Wan. 2019. "When and why people misestimate future feelings: Identifying strengths and weaknesses in affective forecasting." *Journal of Personality and Social Psychology* 116(5): 724–42.

Levine, Linda J., Heather C. Lench, Robin L. Kaplan, and Martin A. Safer. 2012. "Accuracy and artifact: Reexamining the intensity bias in affective forecasting." *Journal of Personality and Social Psychology* 103(4): 584–605.

Maibom, Heidi L. 2012. "The many faces of empathy and their relation to prosocial action and aggression inhibition." *Wiley Interdisciplinary Reviews: Cognitive Science* 3(2): 253–63.

Maibom, Heidi. 2016. "Knowing me, knowing you: Failure to forecast and the empathic imagination." In *Knowledge through Imagination*, edited by Amy Kind and Peter Kung, 185–206. Oxford: Oxford University Press.

Miyazono, Kengo. Forthcoming. "Delusion and self-knowledge." In *Belief, Imagination, and Delusion*, edited by Ema Sullivan-Bissett, Oxford: Oxford University Press.

Montmarquet, James. 2019. "Do epistemic virtues require a motivation for truth?" In *The Routledge Handbook of Virtue Epistemology*, edited by Heather Battaly, 37–46. New York: Routledge.

Nordgren, Loran F., Kasia Banas, and Geoff MacDonald. 2011. "Empathy gaps for social pain: Why people underestimate the pain of social suffering." *Journal of Personality and Social Psychology* 100(1): 120–8.

Scheffler, Samuel. 2010. "Valuing." In his *Equality and Tradition: Questions of Value in Moral and Political Theory*, 15–40. Oxford: Oxford University Press.

Schkade, David A., and Daniel Kahneman. 1998. "Does living in California make people happy? A focusing illusion in judgments of life satisfaction." *Psychological Science* 9(5): 340–6.

Schwitzgebel, Eric. 2002. "A phenomenal, dispositional account of belief." *Noûs* 36(2): 249–75.

Taylor, Charles. 1985. "Self-interpreting animals." In his *Human Agency and Language: Philosophical Papers 1*, 45–76. Cambridge: Cambridge University Press.

Tooming, Uku. 2020. "Being familiar with what one wants." *Pacific Philosophical Quarterly* 101(4): 690–710.

Wilson, Timothy D., and Daniel T. Gilbert. 2005. "Affective forecasting: Knowing what to want." *Current Directions in Psychological Science* 14(3): 131–4.

Wilson, Timothy D., and Daniel T. Gilbert. 2013. "The impact bias is alive and well." *Journal of Personality and Social Psychology* 105(5): 740–8.

Yablo, Stephen. 1993. "Is conceivability a guide to possibility?" *Philosophy and Phenomenological Research* 53(1): 1–42.

Zagzebski, Linda. 2003. "The search for the source of epistemic good." *Metaphilosophy* 34(1–2): 12–28.