Why Can't We Just Believe What We Want?

Abstract: Epistemologists have sought to explain doxastic involuntarism, typically, by appeal to constitutive features of belief. In this paper, I show that doxastic involuntarism is part of a broader class of control discrepancies between committal states, like belief, and imaginative states, like the simulation of counterfactual events in which we aren't committed to the truth or accuracy of the state's content. I explain the broad class of control discrepancies by appealing to metacognitive monitoring and control procedures. As I argue, an explanation that connects belief with other committal states is both broader in scope and more parsimonious than extant accounts of doxastic involuntarism in terms of the theoretical assumptions I make about (supposed) constitutive features of belief. Thus, my explanation should be acceptable to a wider range of epistemologist who disagree about the constitutive nature of belief.

Doxastic involuntarism—the position that we cannot form beliefs directly (as a basic cognitive act) for non-evidential reasons[1]—is standardly taken to mark a defining feature of belief. In this paper, I show that doxastic involuntarism is part of a broader class of control discrepancies between committal states, like belief, and imaginative states, like the simulation of counterfactual events in which we aren't committed to the truth or accuracy of the state's content. I argue that the control discrepancies between committal and imaginative states are explained by how metacognitive monitoring and control processes address, what I call, the mental state identification problem.[2] Imaginative states share content type, format of representation, computational processes of generation, and neural vehicle with committal states. In short, committal and imaginative states are not architecturally distinguished. Thus, cognitive systems like ours that can decouple committal and imaginative states must use an active, inferential process to determine for a given representation whether it's, say, a genuine belief or the mere entertaining of a proposition, an accurate memory of the past or the simulation of a historical counterfactual, etc. As I show, the details of how metacognitive processes solve the mental state identification problem also explain the host of control disparities between committal and imaginative states.

The paper proceeds as follows. In §1, I briefly discuss how epistemologists standardly attempt to establish doxastic involuntarism. In §2, I position doxastic involuntarism as part of a broader class of control discrepancies between committal states and corresponding imaginative states. In §3, I explain the broader class of control discrepancies by appeal to metacognitive monitoring and control procedures. Finally, in §4, I enumerate the theoretical benefits of my account over extant rivals.

§1. Doxastic Involuntarism and Constitutive Features of Belief

Dominantly, epistemologists attempt to establish doxastic involuntarism by invoking Bernard William's (1973) dictum that 'belief aims at truth' (Steglich-Petersen, 2006; Velleman, 2000). However, there is no univocal analysis of belief's (supposed) truth aim. Some offer a teleological reading (McHugh, 2011, 2012; Steglich-Petersen, 2006) on which belief aims at truth in that

(1) we adopt a person-level intention to only believe true propositions in deliberation, and/or

[1] Hieronymi (2006) calls these reasons "extrinsic". They are reasons that speak in favor of possessing a belief.
[2] Metacognition being cognition about one's own cognition (Dokic, 2014).

(2) features of subpersonal psychology constrain belief updating processes in a way that (at least in appropriate circumstances) tends to result in the adoption of true beliefs.

More popularly, some offer a normative reading on which belief aims at truth in that belief is constitutively governed by a truth norm, which prescribes or permits beliefs only if true (Shah, 2003; Shah & David Velleman, 2005; Singh, forthcoming; Wedgwood, 2002). In turn, there are multiple accounts of how belief's aim explains the truth of doxastic involuntarism. Commonly, epistemologists locate the explanans in the interaction of belief's truth aim with features of intention or intentional action. It's argued that conditions on possessing an intention to form a belief or on the intentional formation of the belief make it impossible (or incoherent), given belief's truth aim, to form beliefs at will (Benbaji, 2016; Hieronymi, 2006; Steglich-Petersen, 2006).

However, there is considerable debate about whether belief's truth aim, when properly understood, could explain doxastic involuntarism (Stich, 1990; Vermaire, 2022). For instance, cognitive and social psychology have revealed a litany of non-evidential factors that affect belief updating processes through the modulation of evidential thresholds of acceptance, evidential weighting, memory storage and retrieval processes, etc. These non-evidential factors range from ubiquitous trade-offs between accuracy and speed or cognitive resource use (Forstmann, Ratcliff, & Wagenmakers, 2016) to the motivated maintenance of social identity or a positive self-concept (Kunda, 1990; Zell, Strickhouser, Sedikides, & Alicke, 2020). Arguably, the use of non-evidential factors by belief updating processes isn't a bug or failure of proper function but a positive design feature for doxastic agents operating under time constraints with limited computational resources. If belief aims at anything, it aims at solutions to a weighted function of accuracy, speed (/cognitive resource use), maintenance of a positive self-concept, and social identity maintenance amongst other factors. The weaker (but more plausible) our characterization of belief's truth aim (e.g., belief aims at truth, minimally, in that it is a contextually weighted factor in updating processes) the less likely it is that belief's truth aim will secure the strong thesis of doxastic involuntarism. But the stronger our characterization of belief's constitutive aim the more likely it is that we will be forced to deem clear cases of belief as not really belief, given the host of non-evidential factors that affect belief updating.

In contrast, I offer a cognitive-mechanistic explanation of doxastic involuntarism that makes minimal assumptions about the nature of belief. My arguments serve as a demonstration that we don't need to rely on contentious claims about belief's constitutive nature in order to explain doxastic involuntarism. Instead of further discussing the details of extant attempts to establish doxastic involuntarism, I now turn to my account.

§2. CONTROL DISCREPANCIES BETWEEN COMMITTAL AND IMAGINATIVE STATES

In this section, I show that doxastic involuntarism is part of a broader class of control discrepancies between committal states and corresponding imaginative states. Additionally, I show that our capacity to imagine introduces a problem for the mind, the mental state identification problem: committal and imaginative states are not architecturally distinguished. Thus, cognitive systems like ours that can decouple committal and imaginative states must use an active, inferential process to determine whether a given representation is, say, a genuine belief or the mere entertaining of a proposition. In the next section, I demonstrate that our mind's solution to the mental state identification problem explains the broader class of control discrepancies.

In introducing doxastic involuntarism, it's common to contrast the types of control we possess over belief and propositional imagination (e.g., Shah & David Velleman, 2005; Singh, forthcoming),

that is, imagination that involves a language-like or discursive format of representation that is shared with belief (Nichols, 2004). We can exert liberal, top-down control over conceptual composition in generating propositional representations of arbitrary states of affairs. For instance, we can propositionally imagine that we are ten feet tall directly (Mele, 2017) as a basic cognitive act at the person level for non-evidential reasons (there are no subactions at the person level we must perform to generate the imaginative state[3]). In other words, imagination is 'subject to the will,' as Wittgenstein (1982) claims. However, we can't believe that we are ten feet tall as a basic cognitive act for non-evidential reasons. More precisely, we can't *occurrently* believe (or judge) that we are ten feet tall as a basic cognitive act for non-evidential reasons. (Standing or dispositional beliefs are states of agents as opposed to events that could correspond to action types. Occurrently believing, on the other hand, is an event.)

Doxastic involuntarism is also commonly characterized using what Christman (2008) calls the *no rewards principle*: no matter the reward we cannot, as a basic cognitive act, occurrently believe a proposition for the sole reason that we desire the reward. If I were to offer you $1,000,000 dollars to occurrently believe (if only for a moment) that you are ten feet tall, you couldn't, thereby, believe that you possess a giant's stature, even if you judged that you ought all-things-considered occurrently believe the proposition. However, if I offered you $1,000,000 dollars to propositionally imagine that you are ten feet tall, I'd be down a substantial sum of money.

Imagination, more broadly, is

> the capacity to mimic…non-imaginative kinds of mental states…Imagination presents itself as a complex mental activity that re-creates non-imaginative mental states, producing mental states similar, from a phenomenological and/or functional point of view, to the non-imaginative counterparts re-created. (Arcangeli, 2018, pp. 2-3)[4]

For non-imaginative, committal states—e.g., beliefs, intentions, seeming episodic memories and perceptions (I discuss these two notions below), etc.—there are non-committal, imaginative counterparts that share content type, format of representation, phenomenological features (if the corresponding committal states possess any), and, to an extent, can mimic the functional role of the committal mental states in the scope of an imaginative context.

I am using the expression 'committal' as a functional designation. For instance, occurrent belief (a committal state) and propositional imagination differ in terms of their characteristic connections to other cognitive states and to action planning and execution. Beliefs characteristically exhibit inferential promiscuity (as Stitch 1978 puts it); they are apt to enter into inferential relations with other propositional attitudes resulting in the adoption of (further) beliefs and intentions. Propositional imaginings, on the other hand, only mimic inferential promiscuity in the scope of the assumptive or imaginative context. Akin to how assumption operates in a natural deduction system, in order to infer an occurrent belief from a propositional imagining, one would need to use a 'scope changing' rule. For instance, propositionally imagining that you are ten feet tall will not directly lead to adopting a belief that you have what it takes to be an NBA star. At best, you will infer that *if you were ten feet tall*, you'd have what it takes to be an NBA star. Similarly, beliefs, along with desires and intentions, serve as direct guides for action planning and execution (Mylopoulos & Pacherie, 2019). However, propositional imaginings do not guide action planning and execution in a manner

[3] Cognitive control can be decomposed into subpersonal components: updating and maintaining information in working memory, shifting task set, inhibiting prepotent responses, and allocating processing resources—but, again, these are subpersonal components of a basic person level capacity (Buehler, 2023, 2025).

[4] See (Currie & Ravenscroft, 2002) for similar remarks.

unmediated by a belief that results from a scope changing rule. For instance, propositionally imagining that you are ten feet tall will not immediately cause you to duck under a door frame; the ducking behavior will only be initiated if you desire to act as if you were ten feet tall and occurrently believe that if you were ten feet tall, then you'd need to duck.[5]

Importantly, the control discrepancy between belief and propositional imagination is exhibited by a host of other committal mental state types and their imaginative counterparts. For instance, take episodic memory and sensory imagination. Following Endel Tulving (1972), cognitive psychologists typically partition declarative long-term memory[6] into two distinct systems: an episodic store that contains memories *of* events linked to a spatiotemporal context, and a semantic store that subserves conceptual knowledge and standing or dispositional belief. In terms of content and format of representation, episodic memories constitute (what I call) event representations,

> **Event Representations:** Mental representations of events from a perspective, typically, the field perspective or (as if) 'through one's own eyes.' Event representations are (multi)modal in format but may also include the emotional resonance of the represented event as well as conceptual information regarding, e.g., spatiotemporal context, the gist of the event (that is, the central features or story line of a particular event), the event type, and the types of entities composing the event. (Mahr, 2020; Mitchell & Johnson, 2009)

Correspondingly, sensory imagination involves a (multi)modal format shared with episodic memory such that both sensory imagination and episodic memory can constitute event representations (Byrne, 2010).

Just as we can exert liberal, top-down control over propositional imagination, we can exert liberal, top-down control over the construction of multimodal event representations by binding memory traces and other bits of information stored across cortex under the guidance of our schematic understandings of event types. Memory traces are dispositional properties possessed by brain regions to re-activate at a time, t_1, when triggered by the right cue in, roughly, the same pattern as they did in response to a stimulus or cognitive event at an earlier time, t_0 (De Brigard, 2014; Robins, 2016). For instance, you can, as a basic cognitive act, sensorily imagine playing billiards with David Hume in 1733 by binding memory traces of previous billiard playings and traces of depictions of Hume under the guidance of a more abstract, schematic understanding of billiards. However, you can't, as a basic cognitive act, make it seem as if you are successfully remembering the event such that the state is treated as a genuine episodic memory in cognitive architecture. Event representations our mind *takes* to be episodic memories are committal states[7]. If an event representation seems to be an episodic memory, it will play a more direct functional role in driving action and other cognitive activity than if the event representation were taken to be the simulation of a historical counterfactual. Simulations of

[5] This functional difference is resisted by David Velleman (2000). However, as has been convincingly argued (O'Brien, 2005; Sinhababu, 2016; Van Leeuwen, 2009), there is a clear functional difference between believing and imagining in terms of, e.g., their characteristic connection to action that separates even immersive pretense from belief.

[6] Declarative memory is contrasted with non-declarative memory, or memory that can't be consciously accessed. For instance, I can't consciously access the motor sequences that compose my procedural memory of how to ride a bike, but I can consciously access my declarative memory that Chelsea won the UEFA Champions League in 2012.

[7] I contrast the control we possess over sensory imagination with what our mind *takes* (or what seems) to be genuine episodic memory as opposed to episodic memory, simpliciter, because philosophers typically take memory to be factive. If memory is factive, we clearly can't remember some arbitrary event just by intending to do so. However, *seeming* to remember isn't factive, but I also take it to be clear that we can't make it seem as if we are remembering some arbitrary event just by intending to do so. (It should be noted that the position that memory is factive is somewhat of a philosophical idiosyncrasy that isn't held in the psychological study of memory (De Brigard, 2017).)

historical counterfactuals affect cognitive activity or action outside of the imaginative context only through the operations of scope changing rules.

As a final example, take (seeming) perception and sensory imagination.[8] We can exert liberal, top-down control over the operations of various sensory areas in producing sensory imagery projected onto our visual field. For instance, as a basic cognitive act, you can sensorily imagine the words of this sentence wiggling and moving. However, you can't, as a basic cognitive act, make it seem as if you are, in fact, perceiving the event such that the sensory imagery is treated as genuine perception in cognitive architecture. If a sensory representation is taken to be genuine perception, it will play a more direct role in driving action and other cognitive activity than if the representation is taken to be mere imagery. A sensory representation taken to be mere imagery will require scope changing rules to affect cognitive activity or action outside of the imaginative context.

Importantly, the overlap between committal states and their imaginative counterparts extends to neurocognitive realization. Imagination involves the 'neural reuse' (Anderson, 2010) of cognitive, perceptual, and motor systems, decoupling the operations of the systems from their typical connections to sensory receptors, motor effectors, and other cognitive, perceptual, and motor activity. For instance, perception and sensory imagination activate overlapping neural codes in sensory (and high-level) brain areas (Dijkstra, Bosch, & van Gerven, 2019; Dijkstra, Mazor, Kok, & Fleming, 2021). The ability to knowingly simulate a counterfactual event in your visual field involves decoupling the activity of sensory areas from downstream causal connections with sensory receptors and the upstream operations of attitude updating and action planning processes.

Similarly, episodic memory and sensory imagination are subserved by much the same neurocognitive activity; both constitute event representations dynamically constructed by binding memory traces and other bits of information under the guidance of schematic understandings of event types (Addis, 2018; De Brigard, 2014; Michaelian, 2011; Schacter, Addis, & Buckner, 2007). The ability to knowingly simulate a historical counterfactual involves decoupling the activity of event simulation networks from downstream causal connections to previously experienced events and the upstream operations of action planning and attitude updating processes.

Finally, one and the same capacity to utilize and combine the basic representational ingredients of semantic long-term memory (i.e., concepts) subserves both our abilities to generate committal beliefs and propositional imaginings (Abraham & Bubic, 2015). Propositional imagination involves cognitively quarantining (to use a popular expression in the imagination literature) propositional representations so that they function in belief-like ways only in the pretensive context and don't directly guide action, become stored as standing beliefs, etc.

Broadly speaking, the capacity to imagine is the capacity to decouple the activity of certain cognitive, perceptual, and motor systems from their typical downstream and upstream connections. Using this refined notion of imagination, we can further clarify doxastic involuntarism: we can exert liberal, top-down control over conceptual combination in semantic long-term memory to produce propositional representations of arbitrary states of affairs, but we can't directly (as a basic cognitive act) force propositional representations to take 'wide cognitive scope'. In other words, we can't force the representations to exhibit inferential promiscuity or directly guide action planning and execution in tandem with our desires and intentions. But analogous control discrepancies hold between numerous committal states and their imaginative counterparts, like (seeming) episodic memory, (seeming) perception, and sensory imagination. In addition, analogous *no rewards principles* hold. No matter the reward we cannot, as a basic cognitive act, force an event representation to take wide cognitive scope, thus functioning as if it were an episodic memory, nor can we force sensory

[8] Again, I contrast the control we possess over sensory imagination and *seeming* perception as opposed to perception, simpliciter, because perception is typically taking to be factive.

imagination projected onto the visual field to function as if it were genuine, bottom-up perception. Doxastic involuntarism is, thus, part of a broader class of control discrepancies between committal and imaginative states.

Importantly, our capacity to imagine—to decouple committal and imaginative states—along with the fact that committal and imaginative states are largely indistinguishable neurocognitively creates the mental state identification problem. For instance, an occurrent event representation could (minimally) be the causal result of reinstating a previously encoded perception of an event (Cheng & Werning, 2016), thus constituting an episodic memory. Or the event representation could be the causal result of binding memory traces into a novel event representation never before experienced by the agent, thus constituting the simulation of a historical counterfactual. Similarly, an occurrent propositional representation could (minimally) be the causal result of (i) reinstating a standing belief, (ii) tokening a novel propositional representation via belief updating rules (tacitly) accepted by our cognitive system in generating a new occurrent belief, or (iii) combining concepts in a manner that bears little evidential connection to our standing beliefs. In other words, a given sensory or propositional representation could have various sources, where the source of a representation is a (type of) causal story for the tokening of the representation. But sensory and propositional representations do not come with an 'abstract tag or label' (Johnson, Hashtroudi, & Lindsay, 1993, p. 3) that indicates mental-state type, e.g., whether the state is an episodic memory, a perceptual state, sensory imagery, the reinstatement of a standing belief, the inferred entailment of a set of standing beliefs, etc. Instead, source tags must be actively assigned metacognitively and used to determine cognitive scope, that is, determine whether the state will function as committal or imaginative.

The fact that source isn't luminous is evidenced by mistakes in metacognitive source monitoring. For instance, as is well known, episodic memory is not the perfect reproduction of a previous experience that is stored, unsullied, in some mental file cabinet until it's recalled. Because episodic memory and counterfactual event simulation both result from a constructive process that generates coherent event representations from memory traces utilizing event schemas, source monitoring for event representations is a predictably error-prone process (De Brigard, 2014). For example, autobiographical conjunction errors are quite common in which aspects of distinct memories become unwittingly incorporated into one and the same event representation, R, and treated like genuine episodic memory (Devitt & Schacter, 2016). In treating R as a genuine episodic memory, we are committed to the accuracy of R (functionally speaking), despite R depicting a combination of events that didn't co-occur (unbeknownst to us).

Similarly, the processing of information in sensory hierarchies doesn't come with a tag that indicates whether the information was ultimately caused by the stimulation of sensory receptors (thus constituting perception) or some other process (thus constituting sensory imagination). Source must be inferred. Although hallucination in neuroatypical populations is a prime example of failures of source monitoring, similar failures to distinguish sensory imagination from genuine perception can be reliably induced by various experimental tasks in neurotypical populations (Dijkstra & Fleming, 2021; Dijkstra et al., 2021). Finally, as I argue in the next section, we can even error in determining whether an occurrent propositional representation is reflective of our standing or dispositional beliefs, mistakenly taking the representation to be the reinstatement of a previous event of judging a proposition to be true.

In sum, doxastic voluntarism is part of a broader class of control discrepancies between committal states and their imaginative counterparts. As I establish in the next section, the means by which metacognitive monitoring processes solve the mental state identification problem explains this broader class of control discrepancies.

§3. METACOGNITIVE MONITORING AND CONTROL

In the following, I explain the broader class of control discrepancies between committal and imaginative states by appealing to metacognitive monitoring and control procedures. As I demonstrate, the role of processing fluency in metacognitive monitoring—that is, the perceived ease or efficiency with which a representation is processed (Alter & Oppenheimer, 2009)—is central to explaining the control discrepancies.

So, how does the mind solve the mental state identification problem? Stepping back, as Peter Geach rightly notes, "thought is by nature assertoric, and is inhibited from being so [as in imagination] only by standing in some special relation to a context of other thoughts" (Geach, 1971, p. xi). I argue that this 'context of other thoughts' involves metacognitive monitoring procedures that assign source tags to representational states. Committal mental states are cognitively basic and only come to be treated as imaginative (and, thus, restricted in cognitive scope) through the operations of metacognitive source monitoring. Simpler cognitive systems that lack metacognitive monitoring exhibit a 'naive realism' (as psychologists put it) in which cognitive representations are treated as true architecturally in virtue of the fact that the systems "would simply use the information found inside…to regulate behavior and to carry out further computations" (Tooby & Cosmides, 2000, p. 60). In other words, without metacognitive monitoring all representations take wide cognitive scope and constitute commitments.

Source tags operate over committal representations and function either to decouple the representations from their characteristic connections to downstream and upstream cognitive, sensory, and motor processes or to allow those default connections to remain intact. On the dominant source monitoring framework, metacognitive monitoring processes use myriad features or cues of a representation and its process of generation to assign the representation a source tag (Johnson et al., 1993; Johnson & Raye, 1981). For example, an event representation

(1) that has relatively more spatiotemporal contextual attributes coded in the representation,

(2) that has a relatively high-degree of sensory (e.g., visual-perceptual) detail,

(3) that initiates a strong affective response (Sharot, Delgado, & Phelps, 2004), and/or

(4) that is experienced as fluently processed, where, again, processing fluency is the perceived ease or efficiency with which a representation is processed (Alter & Oppenheimer, 2009)

is more likely to be tagged as an episodic memory and function as such in cognitive architecture (regardless of whether it is). The exact details of the function from metacognitive cues to source tag are not important for explaining control discrepancies between committal and imaginative states (Mazzoni, 2008; Mazzoni & Kirsch, 2002). Instead, what is important is the role that processing fluency, in particular, plays as a cue for source tag assignment.

Why might processing fluency be used as a cue in determining whether an event representation should be assigned an episodic memory or simulated historical counterfactual tag? The various memory traces and information stored across cortex used to generate an event representation that constitutes a genuine episodic memory will tend to be more strongly associatively connected, relative to simulated historical counterfactuals, given that the information was encoded together in virtue of a previous experience. In other words, the memory traces that compose an occurrent event representation that constitutes an episodic memory will tend to be 'wired' together and, thus, activate one another. The ability to combine multimodal memory traces and information of other types that

aren't strongly associatively connected into a representation of a novel event requires more use of frontoparietal control networks that undergird the executive functions of domain-general cognitive flexibility and the inhibition of automatically activated associations (Roberts et al., 2017). Put simply, episodic memories are, on balance, easier to generate (although, processing fluency, like any of the cues used in source tagging, is a fallible guide for source (Finke, Johnson, & Shyi, 1988)).

The fact that metacognition uses processing fluency as a cue for source plays an important role in explaining why we cannot generate *apparent* episodic memories at will. Say you generate (whether intentionally or not) an event representation, R, that isn't causally grounded in a previous experience. The increased demands on executive resources required to flexibly combine disparate memory traces and other types of information (while inhibiting automatic associations) will result in metacognitive source monitoring processes tagging R as sensory imagination and, thus, quarantining R. Clearly, we don't have direct control over the amount of executive resources required to generate a given event representation—this will be a function of the associative strength between the various elements of the representation—nor do we have direct control over metacognitive monitoring processes that issue source tags. In fact, trying to make it seem as if we are remembering—e.g., expending attentional resources to generate a plausible event representation complete with spatiotemporal context, visuo-perceptual detail, emotional poignancy, etc.—is antithetical to it actually seeming as if we are remembering. The more we attempt to will a seeming memory, the less it will seem like one.

Processing fluency is also an important cue for distinguishing perception from sensory imagination and belief from propositional imagination. Because belief is our focus, I center further discussion on belief and propositional imagination. As previously mentioned, an occurrent propositional representation could (minimally) be the causal result of (i) reinstating a standing belief, (ii) tokening a novel propositional representation via belief updating rules (tacitly) accepted by our cognitive system in generating a new occurrent belief, or (iii) combining concepts in a manner that bears little evidential connection to our standing beliefs. In the following, we examine how we distinguish standing beliefs from propositional imagination (i from iii) and how we distinguish new occurrent beliefs from propositional imagination (ii from iii) in turn.

Semantic memory, like episodic memory, is constructive. Standing beliefs are not static information structures stored at some discrete memory address until recalled (Aronowitz, 2023). Because of semantic memory's constructive nature, propositional representations are not pulled from long-term memory complete with an 'abstract tag or label' that indicates their mental state type. Metacognitive monitoring processes utilize processing fluency, amongst other cues, in determining whether a given propositional representation is a genuine semantic memory and, thus, the reinstatement of a standing belief.

Let's look at one representative study that demonstrates the role of processing fluency in how metacognitive processes distinguish standing belief from propositional imagination. Fazio, Brashier, Payne, and Marsh (2015) tested whether our standing beliefs modulate the illusory truth effect—the phenomenon in which repeated statements receive higher confidence ratings than new statements (Dechêne, Stahl, Hansen, & Wänke, 2010). Study participants had to rate 88 statements on a scale of subjective interest under the assumption that their ratings would guide stimulus development for future experiments. The experimenters truthfully informed participants that some statements were true and others false. In a subsequent task, participants rated their confidence in 176 statements using a six-point scale ranging from 'definitely false' to 'definitely true'. The experimenters told participants that they would encounter both true and false statements and that some statements appeared earlier in the experiment while others were new. Finally, participants performed a knowledge check in which they answered 176 multiple choice questions.

Surprisingly, reading a statement in the subjective interest phase increased confidence ratings for the statement in the second phase, even if study participants dispositionally believed otherwise as demonstrated by their performance on the knowledge check. For instance, if study participants read the statement, "A sari is the name of the short pleated skirt worn by Scots," this (tended to) increase their later professed confidence that the statement was true, even if they could correctly answer the question, "What is the name of the short pleated skirt worn by Scots?" on the knowledge check (ibid. p. 996). Because the sentence, "A sari is the name of the short pleated skirt worn by Scots," was recently parsed in the subjective interest task, the sentence was experienced as more fluently processed when read in the confidence rating task. The processing fluency experienced while reading the sentence was taken as a cue that the proposition expressed must be something one remembers being the case and, thus, one dispositionally believes. However, if participants had engaged in a more labored search of semantic memory, they would have realized that they know that "kilt" is the name of the short pleated skirt worn by Scots and, thus, that the sentence, "A sari is the name of the short pleated skirt worn by Scots," should be rated as definitely false.

Fazio et al.'s result is closely related to a host of other findings that demonstrate we frequently use processing fluency as a cue for what we dispositionally believe. To briefly describe one such finding—famously called the Moses Illusion (Erickson & Mattson, 1981)—when people are asked how many animals of each kind Moses took aboard the ark in biblical lore, the vast majority incorrectly respond with "two," despite knowing that it was Noah (not Moses) who built and sailed the ark. Again, because "two" springs to mind easily, "two" is experienced as fluently processed and, thus, incorrectly treated as if it were representative of our dispositional beliefs.

Let's now turn to how we distinguish new occurrent beliefs from propositional imagination. Rakafet Ackerman and Valerie Thompson (2015, 2017a, 2017b) have generated an important model of the metacognitive processes that assist conscious reasoning ('meta-reasoning' for short) that will be crucial for my explanation. Importantly, on Ackerman's and Thompson's model, meta-reasoning monitoring processes give rise to feelings of (un)certainty throughout deliberation, where feelings of (un)certainty constitute non-conceptual assessments of the epistemic quality of the attitudinal transitions a reasoner makes. As Jérôme Dokic puts it, feelings of (un)certainty constitute evaluations of "the non-perceptual method [i.e., inference] we have used to reach [our] conclusion" (Dokic, 2014, p. 136).

Importantly, feelings of (un)certainty determine, in part, how active propositional representations function. If a propositional representation, P, is tagged with a sufficiently strong feeling of certainty by meta-reasoning processes, P will, by default, play the functional role of an occurrent belief. For instance, P will terminate inquiry or solution search and cause overt actions, like assenting to the content of P. Conversely, if P is tagged with a feeling of uncertainty, P will be treated as if it were the non-committal entertaining of a proposition. Thus, solution search will continue, and one won't assent to the contents of P.

It's well established that feelings of (un)certainty are affected by (amongst other cues) processing fluency (Unkelbach & Greifeneder, 2013). In conscious deliberation, a propositional representation that is fluently processed typically engenders a strong feeling of certainty. For instance, take the famous Cognitive Reflection Test (hereafter, CRT; Frederick, 2005). The CRT is composed of a set of computationally simple problems that require one to inhibit/reject an initially intuitive answer to determine the correct solution. For example,

> A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

To determine the answer to the above question, i.e., that the ball costs five cents, one has to inhibit/reject the intuitive yet misleading response that the ball costs ten cents. Many tend to answer the bat-and-ball problem incorrectly, despite possessing ample ability to find the right solution, because a propositional representation, P, of the intuitive yet misleading response is generated rapidly and easily. Because P is fluidly processed, it causes a strong feeling of certainty that terminates solution search. Thereby, P functions as an occurrent belief in virtue of terminating reasoning and driving further behavior, like asserting that the ball costs ten cents. Conversely, experimentally decreasing the processing fluency one experiences in response to the bat-and-ball problem—e.g., by printing the question in a hard to read font that requires increased attention to decipher—will tend to reduce error on the CRT (Alter, 2013). Because of increased cognitive demands, P is not experienced as fluently processed. Thus, P is not assigned a strong feeling of certainty and doesn't function as an occurrent judgment.

Why might processing fluency be used as a cue for feelings of (un)certainty and, ultimately, for how propositional representations function? The fact that a propositional representation, P, is fluently processed is defeasible evidence that P was reached by the iterative application of belief updating rules (tacitly) accepted by the cognitive system. Analogously, grammatical sentences in a natural language with which one is competent—or in an artificial grammar one learns as part of a study (Buchner, 1994; Topolinski & Strack, 2009)—tend to be experienced as fluently processed, as compared to nongrammatical sentences. The processing fluency one experiences when encountering a novel, grammatical sentence isn't the result of repeated exposure to the sentence—the sentence is novel and, thus, not previously encountered. Instead, the processing fluency experienced when encountering a novel, grammatical sentence is the result of how one's tacit knowledge of the generative rules of grammar is stored, and the fact that the sentence can be produced by an iterative application of the rules.

More generally, it's commonly acknowledged that if the content of a representation is 'consistent' or 'coherent' (as psychologists put it) with background information stored in semantic memory, it will be easier to token the representation and, thus, the representation will be experienced as fluently processed (Winkielman, Huber, Kavanagh, & Schwarz, 2012). This is nicely illustrated by the repeated finding that lying tends to be more cognitively demanding than telling the truth (Verschuere, Köbis, Bereby-Meyer, Rand, & Shalvi, 2018; Walczyk, Roper, Seemann, & Humphrey, 2003).[9] Propositional imagination requires flexibly combining concepts into representations of states of affairs, while inhibiting the automatic activation of what one takes to be true, which tends to require relatively more cognitive resources.

Of course, fluency is a defeasible guide to whether a representation was actually reached—or is reachable—by the iterative application of belief updating rules (tacitly) accepted by a cognitive system (just as fluency is a defeasible guide for grammaticality). However, the fact that meta-reasoning processes use fluency as a cue provides us with the beginnings of an explanation for why we can't believe at will—an explanation that mirrors the one offered for why we can't make it seem as if we are remembering at will. Although we can use our background conceptual resources to generate an arbitrary propositional representation, we don't have direct control over metacognitive monitoring

[9] It should be noted that, by 'consistent' psychologists don't mean strict logical consistency. It's beyond the scope of this paper to engage in an exegetical discussion of what psychologists mean by 'consistent,' but, suffice it to say, that my use of the expression "reached by the iterative application of belief updating rules (tacitly) accepted by the cognitive system" is commensurate with what psychologists mean. Just as representations of events one has actually experienced in the past tend to be easier to generate in virtue of architectural features of how information is stored in long-term episodic memory, propositional representations, the contents of which are reachable by belief updating rules (tacitly) accepted by the cognitive system, are easier to generate in virtue of architectural features of how information is stored in long-term semantic memory. See (Winkielman et al., 2012) for a more detailed computational picture of how this works.

processes or the features of propositional representations that make them fluently processed. In sum, try as we might, we can't occurrently believe or make it seem as if we are semantically recalling a proposition at will. In fact, the mental effort employed to combine disparate conceptual resources into a propositional representation that (i) isn't the reinstatement of a dispositional belief or (ii) isn't reachable by (tacitly) accepted updating rules will cause the representation to be accompanied by a low feeling of certainty. The more we attempt to will an occurrent belief or a seeming manifestation of a dispositional belief, the less confident we will feel in the resulting propositional representation.

At this point, one might object that, surely, we are not hostage to the outputs of metacognitive monitoring and control procedures. For instance, some people are able to resist the siren's song of initially intuitive answers to the questions on the CRT. This requires inhibiting the strong feeling of certainty tagged to intuitive responses and engaging in a deliberative review of the evidence. Thus, I've yet to show that we can't believe at will—or so the objection goes—as I've yet to show that we can't, at will, inhibit a feeling of uncertainty directed at a propositional representation, P, by adopting an explicit metacognitive belief that, say, we in fact believe the contents of P.

Presumably, the deliberative processes involved in inhibiting the initial feelings of (un)certainty will, themselves, be composed of a series of propositional representations. In turn, our cognitive system will have to solve a mental state identification problem for these propositional representations (are they occurrent beliefs or mere propositional imaginings?). Thus, we are back where we started. Propositional representations will only function as occurrent beliefs if they are accompanied by a sufficiently strong feeling of certainty. So, as a direct result of aiming to do so, we can generate an arbitrary propositional representation, P; however, we don't have direct control over the feeling of (un)certainty tagged to P, which determines whether P will function, by default, as an occurrent belief or as the mere entertaining of a propositional content. Of course, we could, as a direct result of aiming to do so, generate a metacognitive propositional representation, P^1, with the content that P is an occurrent belief (or that P expresses a truth). However, we don't have direct control over the feeling of (un)certainty tagged to P^1 that determines whether P^1 functions as an occurrent belief that could inhibit the functioning of the feeling of (un)certainty assigned to P. We can continue to attempt to climb the metacognitive ladder (as it were), but, ultimately, metacognitive processes will have the final say as to how a propositional representation functions.

Thus, we have a common explanation for control discrepancies between committal states and their imaginative counterparts: As a basic cognitive act, we can compose the elements of sensory and propositional representations (e.g., memory traces, concepts) into representations of events or states of affairs. However, features of cognitive architecture make the generation of genuine episodic memories, the reinstatements of standing beliefs, and the tokening of novel propositional representations reachable by the iterative application of belief updating rules easier (on balance) to produce. The metacognitive processes responsible for decoupling committal and imaginative states use processing fluency as an important cue to determine cognitive scope. The more cognitive effort we expend in trying to force our mind to treat an event representation as an episodic memory or a propositional representation as an occurrent belief, the less likely we will succeed in doing so.

§4. THEORETICAL BENEFITS OF MY EXPLANATION

My explanation of doxastic involuntarism has several theoretical benefits over extant rivals. First, my explanation is broader in scope. As I demonstrated, doxastic involuntarism is one of a suite of control discrepancies between committal and imaginative states. In addition, I argued that there is a common explanation for the control discrepancies involving the use of processing fluency by metacognitive monitoring processes.

Second, my explanation is more parsimonious in terms of the number of assumptions I make about constitutive features of belief. Instead of needing to rely on contentious claims about (supposed) constitutive features of belief, I showed doxastic involuntarism follows from more general features of our ability to imagine, that is, to decouple committal and imaginative states. Thus, my explanation ought to be acceptable to a number of philosophers who disagree about the constitutive nature of belief. In response, one might object that my explanation treats doxastic involuntarism as a mere psychological contingency, whereas epistemologists have typically assumed that doxastic involuntarism holds of conceptual necessity (cf. Levy, Mandelbaum, Matheson, & Vitz, 2014). However, even if the correct account of belief were to make doxastic involuntarism (when appropriately construed) conceptually necessary, a cognitive-mechanistic explanation would still be valuable. Analogously, take the following fact,

> **Species**: Domestic pigs (*Sus domesticus)* can't produce viable, fertile offspring with domestic dogs (*Canis familiaris*).

Famously, biologists have offered a series of competing, incommensurate accounts of the nature of species (Aldhebiani, 2018). On some accounts, Species comes out as conceptually necessary and on others it doesn't. For instance, on the biological species account of Ernst Mayr (1999), species are "groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups." If species are individuated by possible breeding relationships, Species is conceptually necessary. However, even if the biological species account were correct, the following, related fact couldn't be explained by appeal to individuation conditions on species,

> **Gametes**: *Those gametes*—the gametes of pigs and dogs—can't form a zygote (where, by stipulation, I am picking out the gametes *de re* and not *de dicto* via the description "the gametes of pigs and dogs").

Similarly, even if the correct account of belief were to make doxastic involuntarism conceptually necessary, the following, related fact couldn't be explained by appeal to an analysis of belief alone,

> **Causal-Functional Involuntarism:** It's not the case that we can token a propositional representation with the causal-functional profile that we stereotypically associate with belief (e.g., inferential promiscuity) just by intending to do so.

Finally, discussions of doxastic involuntarism have implications for deontological discussions of epistemic permissions and obligations. An understanding of metacognitive source monitoring and the cues used can help us better understand the forms of indirect control that we can exert over belief updating processes. For instance, an extensive body of literature demonstrates that the repeated exposure to misinformation, even if we don't initially find it plausible, can raise our credence in the information as our memory of the source degrades over time (Fort & Shulman, 2024). With repeated exposure, information is experienced as more fluently processed and, thus, is more likely to be tagged as genuine memory or the reinstatement of dispositional belief. What is especially troubling about repeated exposure to misinformation is that it can have similar cognitive effects on the media naïve and savvy alike and isn't fully remedied by increased media literacy. These results, arguably, have important implications for our individual and collective epistemic practices (e.g., see Mandelbaum & Quilty-Dunn, 2015), implications that will be missed with a narrow focus on constitutive features of belief.

BIBLIOGRAPHY

Abraham, A., & Bubic, A. (2015). Semantic memory as the root of imagination. *Front Psychol, 6*, 325. doi:10.3389/fpsyg.2015.00325

Ackerman, R., & Thompson, V. A. (2015). Meta-reasoning. In A. Feeney & V. A. Thompson (Eds.), *Reasoning as memory* (pp. 164-182). New York: Psychology Press.

Ackerman, R., & Thompson, V. A. (2017a). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21*(8), 607-617.

Ackerman, R., & Thompson, V. A. (2017b). Meta-reasoning: Shedding Meta-cognitive Light on Reasoning Research. In L. J. Ball & V. A. Thompson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 1-15). New York: Psychology Press.

Addis, D. R. (2018). Are episodic memories special? On the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand, 48*(2-3), 64-88.

Aldhebiani, A. Y. (2018). Species concept and speciation. *Saudi journal of biological sciences, 25*(3), 437-440.

Alter, A. L. (2013). The benefits of cognitive disfluency. *Current Directions in Psychological Science, 22*(6), 437-442.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*(3), 219-235.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245-266.

Arcangeli, M. (2018). *Supposition and the imaginative realm: A philosophical inquiry*: Routledge.

Aronowitz, S. (2023). A planning theory of belief. *Philosophical Perspectives, 37*(1).

Benbaji, H. (2016). What can we not do at will and why. *Philosophical Studies, 173*(7), 1941-1961.

Buchner, A. (1994). Indirect effects of synthetic grammar learning in an identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 550.

Buehler, D. (2023). Explicating agency: The case of visual attention. *The Philosophical Quarterly, 73*(2), 379-413.

Buehler, D. (2025). What Is Cognitive Control? 1. *Wiley Interdisciplinary Reviews: Cognitive Science, 16*(2), e70004.

Byrne, A. (2010). Recollection, perception, imagination. *Philosophical Studies, 148*(1), 15-26.

Cheng, S., & Werning, M. (2016). What is episodic memory if it is a natural kind? *Synthese, 193*(5), 1345-1385. doi:10.1007/s11229-014-0628-6

Chrisman, M. (2008). Ought to believe. *The Journal of philosophy, 105*(7), 346-370.

Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*: Clarendon Press.

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191*(2), 155-185.

De Brigard, F. (2017). Memory and imagination. *The Routledge handbook of philosophy of memory*, 127-140.

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review, 14*(2), 238-257.

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346-359. doi:https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2019). Shared neural mechanisms of visual perception and imagery. *Trends in Cognitive Sciences, 23*(5), 423-434.

Dijkstra, N., & Fleming, S. (2021). Fundamental constraints on distinguishing reality from imagination.

Dijkstra, N., Mazor, M., Kok, P., & Fleming, S. (2021). Mistaking imagination for reality: Congruent mental imagery leads to more liberal perceptual detection. *Cognition, 212*, 104719.

Dokic, J. (2014). Feelings of (un) certainty and margins for error. *Philosophical inquiries, 2*(1), 123-144.
Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of verbal learning and verbal behavior, 20*(5), 540-551.
Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General, 144*(5), 993.
Finke, R. A., Johnson, M. K., & Shyi, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition, 16*(2), 133-137.
Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual review of psychology, 67*(1), 641-666.
Fort, K. S., & Shulman, H. C. (2024). Using a signal detection approach to understand the impacts of processing fluency and efficacy on accuracy in misinformation detection. *Frontiers in psychology, 15*, 1417910.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.
Geach, P. (1971). *Mental acts* (2 ed.). London: Routledge & Kegan Paul.
Hieronymi, P. (2006). CONTROLLING ATTITUDES. *Pacific Philosophical Quarterly, 87*(1), 45-74. doi:https://doi.org/10.1111/j.1468-0114.2006.00247.x
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3.
Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological review, 88*(1), 67.
Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin, 108*(3), 480.
Levy, N., Mandelbaum, E., Matheson, J., & Vitz, R. (2014). The powers that bind: Doxastic voluntarism and epistemic obligation. *The Ethics of belief: Individual and social*, 15-32.
Mahr, J. B. (2020). The dimensions of episodic simulation. *Cognition, 196*, 104085.
Mandelbaum, E., & Quilty-Dunn, J. (2015). Believing without reason, or: Why liberals shouldn't watch Fox News. *The Harvard Review of Philosophy, 22*, 42-52.
Mayr, E. (1999). *Systematics and the origin of species, from the viewpoint of a zoologist*: Harvard University Press.
Mazzoni, G. (2008). Metacognitive processes in creating false beliefs and false memories: The role of event plausibility. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 315-331): Psychology Press.
Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 121-145): Cambridge University Press.
McHugh, C. (2011). What do we aim at when we believe? *Dialectica, 65*(3), 369-392.
McHugh, C. (2012). Belief and aims. *Philosophical Studies, 160*, 425-439.
Mele, A. R. (2017). Direct control. *Philosophical Studies, 174*(2), 275-290.
Michaelian, K. (2011). Generative memory. *Philosophical Psychology, 24*(3), 323-342. doi:10.1080/09515089.2011.559623
Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psychol Bull, 135*(4), 638-677. doi:10.1037/a0015849
Mylopoulos, M., & Pacherie, E. (2019). Intentions: The dynamic hierarchical model revisited. *Wiley Interdisciplinary Reviews: Cognitive Science, 10*(2), e1481.
Nichols, S. (2004). Imagining and believing: The promise of a single code. *The Journal of Aesthetics and Art Criticism, 62*(2), 129-139.
O'Brien, L. (2005). Imagination and the motivational view of belief. *Analysis, 65*(1), 55-62.

Roberts, R., Wiebels, K., Sumner, R., van Mulukom, V., Grady, C., Schacter, D., & Addis, D. (2017). An fMRI investigation of the relationship between future imagination and cognitive flexibility. *Neuropsychologia, 95*, 156-172.

Robins, S. (2016). Representing the past: Memory traces and the causal theory of memory. *Philosophical Studies, 173*, 2993-3013.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature reviews neuroscience, 8*(9), 657-661.

Shah, N. (2003). How Truth Governs Belief. *Philosophical Review, 112*(4), 447-482.

Shah, N., & David Velleman, J. (2005). Doxastic deliberation. *Philosophical Review, 114*(4), 497-534.

Sharot, T., Delgado, M. R., & Phelps, E. A. (2004). How emotion enhances the feeling of remembering. *Nature neuroscience, 7*(12), 1376-1380.

Singh, K. (forthcoming). Belief as Commitment to the Truth. In E. Schwitzgebel & J. Jong (Eds.), *The Nature of Belief*. Oxford: Oxford University Press.

Sinhababu, N. (2016). Imagination and belief. *The Routledge handbook of philosophy of imagination*, 111-123.

Steglich-Petersen, A. (2006). No norm needed: On the aim of belief. *The Philosophical Quarterly, 56*(225), 499-516.

Stich, S. P. (1978). Autonomous psychology and the belief-desire thesis. *The Monist*, 573-591.

Stich, S. P. (1990). *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*: The MIT Press.

Tooby, J., & Cosmides, L. (2000). Consider the source: the evolution of adaptations for decoupling and metarepresentations. *Metarepresentations: A Multidisciplinary Perspective, New York*, 53-116.

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General, 138*(1), 39.

Tulving, E. (1972). Episodic and semantic memory.

Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In C. Unkelbach & R. Greifender (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 11-32): Psychology Press.

Van Leeuwen, D. N. (2009). The motivational role of belief. *Philosophical Papers, 38*(2), 219-246.

Velleman, D. (2000). On the aim of belief. In D. Velleman (Ed.), *The Possibility of Practical Reason* (pp. 244--281): Oxford University Press.

Vermaire, M. (2022). In search of doxastic involuntarism. *Philosophical Studies, 179*(2), 615-631.

Verschuere, B., Köbis, N. C., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2018). Taxing the brain to uncover lying? Meta-analyzing the effect of imposing cognitive load on the reaction-time costs of lying. *Journal of applied research in memory and cognition, 7*(3), 462-469.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 17*(7), 755-774.

Wedgwood, R. (2002). The Aim of Belief. *Philosophical Perspectives, 36*(16), 267-297.

Williams, B. (1973). Deciding to believe. In B. Williams (Ed.), *Problems of the Self: Philosophical Papers 1956–1972* (pp. 136–151): Cambridge University Press.

Winkielman, P., Huber, D. E., Kavanagh, L., & Schwarz, N. (2012). Fluency of consistency: When thoughts fit nicely and flow smoothly. *Cognitive consistency: A fundamental principle in social cognition*, 89-111.

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological bulletin, 146*(2), 118.