Prolegomena 25 (1) 2026: 291–310
doi: https://doi.org/10.26362/20260112

Belief, Agency, and the Spontaneity of Consciousness[1]

ZHIHAN LI

Department of Philosophy, Tsinghua University, Haidian District,
Beijing 100084, China
lzhphilo@gmail.com

ORIGINAL RESEARCH PAPER – RECEIVED 24. 9. 2025. ACCEPTED 7. 2. 2026.

ABSTRACT: Mental agency has recently become a prominent theme in the philosophy of mind. One of the most debated issues within this field is doxastic agency, which especially centers on whether we can believe at will. There are three typical positions on this issue. (1) Doxastic ballistic theory holds that agency is only involved in the preceding acts that lead to beliefs, while beliefs themselves lack agency. (2) Doxastic particularism argues that beliefs involve a *sui generis* form of agency distinct from the agency in everyday actions. (3) Doxastic voluntarism maintains that beliefs are subject to direct voluntary control. However, all three positions are unsatisfying in certain respects. This paper proposes an alternative approach: "doxastic spontaneism", which offers a solution to the problem of doxastic agency. At its core lies Sartre's account of the "spontaneity of consciousness". Spontaneity is not mere instinct or automatic behavior, but pre-reflective self-determination directed toward certain ends. It constitutes the foundation of the will (which is reflective self-determination), representing a more fundamental form of mental agency than the will. Although beliefs are often beyond the direct control of the will, they nonetheless involve agency in the form of spontaneity.

KEYWORDS: doxastic voluntarism, free will, Jean-Paul Sartre, Matthew Boyle, mental agency, self-knowledge, spontaneity

1. Introduction

In recent years, mental agency has emerged as a prominent theme in the philosophy of mind. Many mental phenomena are now regarded as activities rather than mere passive states, and the subject is conceived not as a spectator but as an agent of the mind. This *agentialism* of the mind has provided new perspectives for addressing numerous philosophical problems. As one of the most prominent agentialists, Matthew Boyle writes:

[1] I wish to thank Hao Tang, Asher Jiang, and the anonymous reviewers for their careful feedback on earlier versions of this paper. I am also indebted to Aining He, Dawei Wu, and Yuchen Zhu for several stimulating conversations.

> A point of persistent controversy in recent philosophical discussions of belief concerns whether we can exercise some sort of agential control over what we believe... This idea has been invoked, for instance, to [1] characterize the basic difference between rational and non-rational cognition, to [2] account for our epistemic responsibility for what we believe, and to [3] explain how we are able, normally, to say what we presently believe without relying on self-observation or inference. (Boyle 2009: 109)

Since many contemporary discussions of mental agency tend to take belief as the paradigmatic case, this paper will focus specifically on doxastic agency. Before delving deeper, let me briefly introduce the three applications of doxastic agency that Boyle mentions:

(1) Korsgaard views human rationality as "the active dimension of the mind" (Korsgaard 2009: 30). She thinks that the source of rationality lies in our capacity to reflectively interrogate and control our mind. For instance:

> I perceive, and I find myself with a powerful impulse to believe. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I believe? Is this perception really a *reason* to believe?[2] (Korsgaard 1996: 93)

This process demands that we exercise agential control over our beliefs. As she claims: "for a rational animal, believing itself is an active state, it is *doing* something, it is an activity" (Korsgaard 2009: 37).

(2) We typically hold epistemic agents responsible for their beliefs. For example, we criticize others for wrong beliefs, for forming beliefs without sufficient deliberation or evidence, or for clinging to superstitious ideas, etc. This topic is now often discussed under the label of "ethics of belief". One way to explain doxastic responsibility is to regard ourselves as *agents* of belief, just as we are agents of action. If belief is essentially something we *do*, it is natural that we bear responsibility for it.

(3) Normally, we do not need to rely on self-observation or inference to know what we believe. Gareth Evans famously suggests that, to know whether one believes P, we do not look inward at our mind. On the contrary, we look outward to the world to see whether P is true (Evans 1982). Richard Moran calls this feature of self-knowledge "transparency" (Moran 2001). Why is it that, merely by considering whether P is true, we can also know whether we believe P? Moran's answer is that, in assessing the truth of P, we "make up our mind" about whether to believe it. In other words, belief is formed by an active self-determination. Just as our self-knowledge of action

[2] All *italics* in quoted passages in this paper are original. Her passage manifests an ambiguity in doxastic agency: is doxastic agency consist in the preceding acts that lead to belief formation, or is belief itself an act of agency? This issue will be discussed in the following section.

is non-observational,[3] our self-knowledge of belief is also non-observational.

Thus, by appealing to doxastic agency, these three issues receive compelling explanations. However, the core premise of this approach—that belief is agential—is highly contentious. Those who are skeptical of doxastic agency highlight a crucial difference between belief and action: unlike action, belief does not appear to be directly subject to voluntary control; so, the idea of doxastic agency may lack sufficient justification. This brings us to the heart of the issue: the debate over *doxastic voluntarism*. In two influential articles, Bernard Williams (1973) and William Alston (1988) argued—respectively from conceptual and psychological perspectives—against the possibility of voluntary control over beliefs. As a result, doxastic *in*voluntarism has become the mainstream position among philosophers. If belief cannot be controlled by the will, how can it be agential?

This paper offers a defense of doxastic agency by drawing on Sartre's account of the "spontaneity of consciousness". It is a more fundamental form of mental agency than the *will*. Belief's agency lies precisely in its *spontaneity*. In the next section, I will introduce three intuitively attractive claims about doxastic agency that appear to form an inconsistent triad, which leads to three typical theoretical positions on belief. My goal is to develop an account that reconciles these seemingly contradictory claims. The third section will present Boyle's Sartrean account of the transparency of self-knowledge, which provides important inspiration for my account. The fourth section will elaborate Sartre's account of the spontaneity of consciousness, and argue—through two key questions—that belief is agential in the form of spontaneity. Finally, the fifth section will offer two additional arguments to reinforce my view.

2. The "Seeming Inconsistent Triad" of Doxastic Agency

The literature on doxastic agency is vast, making a comprehensive summary impossible. Nonetheless, we can find a way to reconstruct the various positions on this topic, which I call the "seeming inconsistent triad" of doxastic agency. It consists of three intuitively attractive propositions, yet it appears that any two of these being true entails the falsity of the third. This yields three typical positions on doxastic agency. The three propositions are:

A. We cannot believe at will;[4]

---

[3] Moran draws important inspiration from Anscombe's notion of "practical knowledge" (Anscombe 2000), which is non-observational self-knowledge of one's action.

[4] In contemporary discussions of doxastic voluntarism, "voluntary" belief and belief "at will" are normally used interchangeably (see Boespflug & Jackson 2024).

B. Agency is uniform. There is no *sui generis* doxastic agency which is different from the agency in everyday action.
C. Doxastic agency is not merely involved in the preceding acts that lead to belief; belief itself is an act of agency.

Let me offer some explanations. (A) expresses the involuntarist view of belief. We usually think of belief as formed "automatically", not subject to direct voluntary control. No one denies that belief can be subject to *indirect* voluntary control—for instance, we can deliberately expose ourselves to certain evidence, which will then lead us to form certain beliefs. But involuntarists deny that belief itself can be subject to direct voluntary control. We can raise our hand at will, but we cannot believe at will a proposition we know to be false. (B) posits that although agency may manifest differently in various activities, it is essentially uniform. As we shall see, some scholars, in defending the agency of belief, propose that it is a *sui generis* form of agency which is involuntary. Yet such a proposal risks appearing *ad hoc*. (C) holds that agency is a feature of belief itself. If it is merely the preceding acts which lead to belief that are agential, it seems equivalent to the claim that belief itself is not agential. As we shall see, these three propositions appear to be incompatible, and giving up one of them constitutes adopting a specific position toward belief.

(1) (A) true, (B) true, (C) false—Doxastic Ballistic Theory

If we accept that we cannot believe at will, and reject any *sui generis* form of doxastic agency, it seems that we must deny that belief itself is agential. Instead, agency resides only in preceding acts that lead to belief formation. Scholars in this camp[5] identify certain preceding acts that can be controlled by the will and that lead to belief, yet insist that belief itself is simply its product. For instance, Robert Audi writes: "We perhaps picture ourselves as agents of belief formation when what we have really done is to create (or enter) circumstances in which it occurs as a nonvoluntary response to a pattern of evidence" (Audi, 2001: 98). Compared to actions that involve manipulating evidence, belief is more closely tied to mental action: focusing attention, imagining, recollecting, etc. These are mental actions that can be directly controlled by the will, but scholars in this camp think belief is not one of them, but only a consequence of such mental actions. These mental actions (focusing, imagining, etc.) can lead to belief formation, yet "doxastic agency" is limited only to these preceding mental actions. The most radical

[5] For example, Galen Strawson (2003), Cottingham (2002), Audi (2001), Feldman (2001), and Parrott (2015).

articulation of this view is Galen Strawson's "mental ballistics," which posits that only the initiation of mental events is genuinely agential:

> ... the role of genuine action in thought is at best indirect. It is entirely *prefatory*, it is essentially—merely—*catalytic*... No doubt there are other such preparatory, ground-setting, tuning, retuning, shepherding, active moves or intentional initiations. But action, in thinking, really goes no further than this. The rest is waiting, seeing if anything happens, waiting for content to come to mind, for the 'natural causality of reason' to operate in one. This operation is indeed spontaneous, but in the sense of 'involuntary, not due to conscious volition'. (Strawson 2003: 231–2)

If we find this position hard to accept, this intuitive discomfort reveals the problem with relying solely on "the will" to understand mental agency. I will return to this issue later. Moreover, this view faces a difficulty in explaining doxastic responsibility. When we blame someone for a wrong belief, it is typically the belief itself that is at issue, not merely the preceding acts that lead to it. Even if a person's belief-forming acts are entirely reasonable and impeccable—he is merely misled by bad information source that cannot be avoided—we may still hold that they ought not to maintain that belief. But if belief itself is not agential, how can we be responsible for it?

(2) (A) true, (C) true, (B) false—Doxastic Particularism

This kind of view accepts that we cannot believe at will, yet maintains that belief itself involves agency. How can belief be agential if it is not under voluntary control? Proponents of this view[6] reject the uniformity of agency, and suggest a *sui generis* form of doxastic agency distinct from agency in everyday voluntary action.[7] For example, Moran admits that belief is passive in one sense—it cannot be willed. But he thinks "that restriction is just the reflection of the dominance of another kind of agency, that is, my responsiveness to epistemic reasons for my beliefs" (Moran 2012: 233). When we cannot change our belief in P at will, it is precisely because the reasons for P's truth outweigh the reasons against it, and our belief appropriately responds to these reasons. This responsiveness is a kind of "agency as responsiveness to reason" (Moran 2012: 235), but distinct from "agency in production" (Moran 2012: 235), i.e., the everyday voluntary agency. In Moran's view, the two forms of agency are essentially different: doxastic agency demands the absence of voluntary control—otherwise we could believe at will regardless of evidence. Pamela Hieronymi similarly distinguishes between "manipulative control" and "evaluative control" (Hieronymi 2009: 140); the former is

[6] For example, Moran (2012), Hieronymi (2009), and Boyle (2011a).

[7] They do not regard this *sui generis* form of agency as unique to belief, but as characteristic of a broader class of intentional attitudes, such as desire, intention, etc.

voluntary control, the latter is attitudinal control. She takes the will as direct responsiveness to *practical* reasons (Hieronymi 2006: 58). Because belief only indirectly responds to practical reasons, it requires a non-voluntary evaluative control. For example, someone who wants to believe her child is safe might make a call to provide herself with convincing evidence and thereby form the belief (Hieronymi 2009: 141). But Hieronymi contends that this is merely extrinsic control over beliefs. The intrinsic control over beliefs is direct responsiveness to *epistemic* reasons, which must be involuntary. However, this distinction of two kinds of agency also faces challenges. Typically, we regard voluntary control in actions as the paradigm of agency. Introducing a *sui generis* form of agency to explain doxastic responsibility and non-observational knowledge might seem *ad hoc*. As Campbell and Greenberg (2024) argue, we could simply treat belief's responsiveness to reasons as an explanation for doxastic responsibility and non-observational knowledge, without regarding this responsiveness as a kind of agency. This suggests that a robust defense of agentialism requires explaining why responsiveness to reasons amounts to agency, rather than merely positing that it does.

(3) (B) true, (C) true, (A) false—Doxastic Voluntarism

If belief is itself an act of agency, and this agency is essentially uniform with that of everyday action, it seems natural to assimilate belief and action, and regard belief as voluntary. One form of doxastic voluntarism is "epistemic permissivism",[8] which is the view that "the evidence may underdetermine what is rational to believe, such that there can be multiple rational doxastic attitudes to a proposition, given some evidence" (Boespflug & Jackson 2024). Permissivists acknowledge that in many cases of overwhelming evidence, belief cannot be chosen at will (e.g., "1+1=3" cannot be believed). But in situations of evidential underdetermination, we can choose to believe. For instance, Carl Ginet describes a case where, having driven fifty miles from home and unsure whether the door is locked, the inconvenience of returning prompts him to decide to believe that it is locked (Ginet 2001: 64). This view, however, only accounts for responsibility for beliefs that are formed in evidentially underdetermined contexts (for they can be chosen by will), and it struggles to explain responsibility for beliefs formed under decisive evidence. Another form of doxastic voluntarism is doxastic compatibilism,[9] which is the view that freedom and determinism are compatible: beliefs may be determined by evidence, yet still be free if the determination process is

[8] Proponents of this view, for example, are Ginet (2001), Nickel (2010), and Roeber (2020).
[9] For example, Steup (2008) and Ryan (2003).

appropriate. For example, Matthias Steup thinks the view that "whereas responsiveness to practical reasons grounds freedom, responsiveness to epistemic reasons does not" is a kind of "chauvinism" of practical reason (Steup 2008: 388). He maintains that, just as acting in response to *practical* reasons is an expression of free will, so too is forming a belief in response to *epistemic* reasons. If compatibilism about action is right, there is no reason to reject that it also holds for belief. While some beliefs cannot be changed arbitrarily and are determined by evidence, compatibilism holds that they are still chosen by free will. This position is a bit counter-intuitive in cases where we apparently cannot hold a belief "at will"; so the meaning of "the will" needs clarification, which is what I will undertake in the following discussions. As we'll see, my own account shares with doxastic voluntarism the idea that all beliefs involve self-determination. But I will show that beliefs remain involuntary in an important sense.

In sum, these three intuitively attractive propositions seem to form an inconsistent triad. My task is to develop an account that can reconcile them and show that they can be true together. As we will see, Sartre's views on the spontaneity of consciousness provide an especially illuminating framework. Building on this, I will advance the "doxastic spontaneism". The next section will discuss Boyle's application of Sartre's account to the transparency of self-knowledge—an account that also suggests resources for addressing doxastic agency.

3. Boyle on Transparency and Reflection

As stated in the first section, one major motivation for introducing doxastic agency is to explain the transparency of self-knowledge. Moran links transparency directly to doxastic agency: our ability to know that we believe P simply by judging P to be true arises because we are agents who determine what we believe, and thus possess agential knowledge of our beliefs. However, Moran does not offer a detailed explanation of how this self-knowledge comes about. Boyle's recent book (Boyle 2024) presents the most comprehensive account to date of the mechanism of transparent self-knowledge, which offers inspiration and basis for my own account. So, I will first briefly introduce Boyle's account.

Boyle draws important resources from Sartre's theory of consciousness. Following Husserl, Sartre thinks the essence of consciousness is intentionality: "All consciousness is consciousness *of* something" (Sartre 2021: 9); or in his words, consciousness always "posits" an object. Thus all consciousness is "positional consciousness". Sartre claims that, in order for a positional consciousness to become conscious of itself, it does not require a second-

order positional consciousness (i.e., reflective consciousness). Every positional consciousness is immediately accompanied by a pre-reflective self-consciousness, which he calls "non-positional consciousness": "Any positional consciousness of an object is at the same time a non-positional consciousness of itself" (Sartre 2021: 11). Non-positional consciousness is not intentional, objectifying, or reflective; it is not a higher-order consciousness but an integral part of every consciousness, making every consciousness directly presented to the subject. Sartre holds that pre-reflective consciousness (i.e., non-positional consciousness) is the fundamental form of self-consciousness, and is the condition for reflective self-consciousness: "nonreflective consciousness is what makes reflection possible: there is a prereflective *cogito*, which is the condition of the Cartesian *cogito*"[10] (Sartre 2021: 12).

According to Boyle, the problem of transparency is essentially about the relationship between a world-directed consciousness and a self-directed consciousness. Drawing on Sartre's account, Boyle views the world-directed consciousness as *positional* consciousness, which can be distinguished into two dimensions: "content" and "mode" (Boyle 2024: Ch. 3). For example, the same house (content) can be represented as perceived, imagined, or remembered (mode). In the case of intention, it represents the content P in the mode of "to be done"; it is not a prediction or hypothetical attitude toward P but a *practical resolution* that P must be done. Similarly, belief represents P in the mode of "to be held true". It is not an imagination or a hypothetical attitude toward P, but a *cognitive resolution* that P is true. The question "Is P true?" is closed rather than open for me. According to Sartre's theory, when a *positional* consciousness represents an object in a certain mode, we simultaneously have a *non-positional* consciousness of our own mental states. When P is positionally represented as "to be done", we have a non-positional consciousness that "I intend P"; when P is positionally represented as "to be held true", we have a non-positional consciousness that "I believe P". In other words, world-directed consciousness already includes a pre-reflective consciousness of one's own mind. From Boyle's "reflectivism"[11] of self-knowledge, the way we acquire transparent self-knowledge is through reflection: when we turn our attention from the external world to ourselves,

[10] In Sartre's terminology, "nonreflective consciousness" is the same as pre-reflective consciousness; and "Cartesian *cogito*" means reflective *cogito*.

[11] Boyle advanced "reflectivism" in an earlier paper (see Boyle 2011b). What differs is that in his earlier work he characterized our pre-reflective awareness of the mind as a form of "tacit knowledge", whereas he now prefers the Sartrean "non-positional consciousness". The reasons for this shift will not be discussed here (see Boyle 2024: 64–5), but what remains unchanged in his view is that transparent self-knowledge is acquired through reflection.

we make the pre-reflective consciousness of our mind explicit, thus achieving explicit self-knowledge.

As an advocate of agentialism, Boyle had previously argued for doxastic agency in two influential early papers (Boyle 2009, 2011a). His view is doxastic particularism, maintaining that belief involves a form of agency different from that of everyday action. However, in his recent turn to the Sartrean theory, he does not apply the Sartrean framework to the problem of doxastic agency. In the following sections, I will argue that Sartre's theory of consciousness does indeed provide theoretical resources for understanding doxastic agency, and it can preserve the three intuitively attractive propositions of doxastic agency discussed in the previous section. My approach also fills a significant gap in Boyle's book: whereas he applies Sartre's theory only to the transparency of self-knowledge, I will demonstrate how it can also illuminate the problem of doxastic agency.

4. Belief and the Spontaneity of Consciousness

In my view, a major shortcoming of contemporary debates on doxastic voluntarism is the insufficient attention given to the very nature of the *will*. Many scholars engaged in this discussion seem to operate under a naïve assumption: that free acts simply amount to acts at will. Sartre is sharply critical of this volitionalist assumption. For him, the fundamental form of self-determination is not by the will but *spontaneity*. Sartre takes spontaneity to be the essential feature of consciousness. This resonates with contemporary agentialist views of the mind, which see mental phenomena as active rather than passive. I will show that the spontaneity of consciousness offers a promising path to account for doxastic agency. However, the notion of "spontaneity" seems to contain some tensions. It faces the two following questions.

4.1 Spontaneity and the Will

Question (1): According to the tradition of German idealism, "spontaneity" is the fundamental form of freedom—it means self-determination rather than being determined by anything else. If belief is spontaneous in this sense, it is of course agential. However, as Galen Strawson points out, "spontaneity" also has an ordinary meaning of instinctive, automatic occurrence—it is "'involuntary, not due to conscious volition', not actions at all" (Strawson 2003: 229). He writes: "spontaneity has nothing particularly to do with action or will, and nothing at all to do with freedom of choice" (Strawson 2003: 233). We are thus faced with two different senses of "spontaneity"—the German idealist sense and the ordinary sense. Is it merely an

ambiguity of the term? We find Sartre's answer to it in his account of spontaneity and the will.

Sartre believes that consciousness is spontaneous in that:

> It determines itself to exist at every instant, without us being able to conceive of anything *before* it. Thus every instant of our conscious lives reveals to us a creation *ex nihilo*. Not a new *arrangement* but a new existence. (Sartre 2004: 27)

For Sartre, consciousness must be understood from the first-person perspective. He is here describing a phenomenological fact of consciousness:[12] from the perspective of the subject, consciousness is always experienced as arising from itself, never as affected by something more basic. In this sense, consciousness is creation *ex nihilo*, we cannot experience anything before consciousness. For Sartre, spontaneity is the first principle of consciousness and the most fundamental form of mental agency. We cannot conceive of a more basic form of mental agency than spontaneity. According to this view, belief is also a spontaneous mental attitude: the subject exercises doxastic self-determination to believe something. Galen Strawson also acknowledges the spontaneity of belief, but only in the sense of belief as an instinctive, automatic occurrence: beliefs "occur, spring up—the process is largely automatic" (Strawson 2003: 229). Such spontaneous beliefs cannot be changed by the will and thus cannot truly be called self-determined, as illustrated by Alston's famous example: we simply cannot will ourselves to believe "the U.S. is still a colony of Great Britain" (Alston 1988: 263). According to Strawson, if there is any role for the will in belief, it is only an initiating one: we can will ourselves to focus our minds on certain topics, thus allowing the "natural rational powers" to help beliefs form spontaneously (Strawson 2003: 233). For Strawson, the relationship between *spontaneity* and the *will* is one of necessary separation: spontaneity is a kind of involuntary, automatic occurrence, it has nothing to do with the will, and is not a form of agency.[13]

By contrast, Sartre thinks that spontaneity and the will are intimately connected. He writes:

---

[12] What complicates matters is that Sartre, following Heidegger, sees phenomenology as a pathway to ontology. For Sartre and Heidegger, phenomenology is by no means merely a description of subjective experience. This touches on profoundly deep issues that lie beyond the scope of this paper.

[13] Strawson's understanding of the term "spontaneity" is taken from the *Shorter Oxford English Dictionary*, which defines it as "involuntary, not due to conscious volition" (Strawson 2003: 229). But the issue is not so straightforward. The root of "spontaneity" is the Latin word "spons", which, according to the *Oxford Latin Dictionary*, carries meanings such as "will, volition", "by one's own agency", "deliberately, purposely", etc. (Glare 2012: 1995) Clearly, the Latin etymology of "spontaneity" is closely tied to the "will"—an association that Sartre also inherits and develops.

> The voluntary act is distinguished from nonvoluntary spontaneity by the fact that the latter is a purely unreflective consciousness... The structure of a voluntary act, on the contrary, requires the appearance of a reflective consciousness. (Sartre 2021: 591)

Sartre's distinctive contribution is that he creatively links the distinction between spontaneity/the will to the distinction between pre-reflective/reflective consciousness. According to Sartre, *spontaneity is pre-reflective self-determination, while the will is reflective self-determination*. Spontaneity and the will are not separate powers; rather, they have a relationship of dependence: *just as pre-reflective consciousness is the basis and precondition of reflective consciousness, spontaneity is the basis and precondition of the will*.

Let me illustrate this. We can do A spontaneously without being reflectively conscious that "I decide to do A", yet the action remains self-determined in a pre-reflective form. We can also deliberate on various options of action and resolve to do A, meanwhile explicitly aware that "I decide to do A"—this is reflective self-determination. The latter case is what we ordinarily recognize as an act of the will, for reflective choice is accompanied by a sense of "up-to-me-ness". But Sartre thinks reflection is not a necessary condition for self-determination; we can also decide in a pre-reflective, spontaneous way. For example, someone spontaneously breaks into a run to catch a tram that is about to leave. His action might look like an instinctive impulse, akin to sneezing or a knee-jerk reflex. But unlike those, his tram-catching can be clarified in terms of purpose and reason. According to Sartre, this action actually comes from a spontaneous self-determination. (I'll say more about this later.) Now imagine that, when feeling the strong impulse to run toward the tram, that guy starts to reflectively deliberate whether to act, as Korsgaard writes:

> I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is this desire really a *reason* to act? (Korsgaard 1996: 93)

Korsgaard's view helps to illuminate why the will is essentially reflective: the exercise of the will requires stepping back to critically examine one's first-order attitudes, and thereby exerting reflective control over them. That's part of the reason why Sartre thinks the will, as a special kind of mental phenomenon, is not the fundamental form of self-determination:

> The will is not a preeminent manifestation of freedom but a psychological event with a distinctive structure, constituted at the same level as the others, and which is, no more or less than the others, supported by an original and ontological freedom. (Sartre 2021: 593)

This "original and ontological freedom" is the spontaneity of consciousness, which is the basis of the will. The fact that the will cannot alter spontaneity does not mean that the person is not self-determining. Based on this, Sartre criticizes the philosophical tradition of "free will", arguing that freedom in its most fundamental form is not freedom of the *will*, but freedom of *consciousness*—that is, spontaneity. Freedom of the will is merely one manifestation of this spontaneity. Building on these considerations, Sartre reconciles the German idealist sense and the ordinary sense of "spontaneity". He appeals to the essential feature of consciousness to capture the self-determining aspect of spontaneity, while characterizing its instinctive and automatic aspect in terms of the pre-reflective. Although many actions may appear merely instinctive or automatic, Sartre thinks they are in fact freely determined by the subject pre-reflectively. This account can also be applied to belief. We cannot *will* ourselves to believe that "the U.S. is still a colony of Great Britain", because the available evidence makes us believe *spontaneously* that "the U.S. is *not* a colony of Great Britain". In such cases, reflective will is unable to override a more fundamental spontaneous decision—unless reflection provides us with reasons to reject the belief.

4.2 Spontaneity and Freedom

Question (2): Sartre's description of spontaneity seems more like "random spontaneity", which seems unlikely to be a kind of freedom. In his account, spontaneous consciousness appears as a flux of randomly, unpredictably emerging thoughts that are beyond the subject's control. How, then, can it be called free? This is the problem of "unfree spontaneity".

To answer this question, let us once again consider some examples from everyday actions. (1) Some "spontaneous" actions are purely instinctive or automatic bodily movements, such as sneezing or the knee-jerk reflex. (2) Some other "spontaneous" actions may appear to be goal-directed and self-determined, yet in fact arise from irresistible psychological drives—for instance, a person with obsessive-compulsive disorder who, after leaving home, repeatedly returns to check a dozen times whether the door is locked. (3) Finally, there are actions we would readily regard as both spontaneous and free. Consider the case of driving: one's attention is fully absorbed in the world, and one "automatically" turns the steering wheel, presses or releases the accelerator and brake, without reflecting on what to do. All three cases can be called "spontaneous" in ordinary language, but why is only the third an instance of free action? In my view, it is because the concrete movements involved in driving are governed by some broader intentions, such as "driving to the office", "making a living", etc. The subject's "automatic" acts of

driving are the realization of these broader intentions. Since these broader intentions are freely chosen by the subject, the "automatic" actions performed in driving also count as free. The hierarchical structure of human intentions (e.g., "pressing the accelerator" → "driving to the office" → "making a living") is the key to explaining actions that are both spontaneous and free. How does this relate to Sartre's account? For Sartre, the structure of human intentions can also be illuminated through phenomenological investigations of consciousness.

For Sartre, spontaneity is not random; it always directs toward certain ends. Following Husserl, Sartre takes the essence of consciousness as *intentionality*. He does not conceive of intentionality as passive mirroring of objects, but rather as active orientation toward them. Gavin Rae's (2024) discussion is helpful here. He points out that intentionality, in addition to its technical meaning of "aboutness", also carries a more everyday meaning: "to aim to do"—that is, it denotes active goal-directedness. Rae points out that Sartre's conception of intentionality incorporates both meanings: consciousness is not only *about* its object, but also *intends* it—that is, it spontaneously and actively directs toward its object in certain ways. These "ways" are precisely what Boyle calls "modes" of representation. Attitudes like intention and belief contain distinct modes in which consciousness actively directs toward its object—intention represents its object as "to be done", while belief represents its object as "to be held true". They are both manifestations of the active intentionality of consciousness.

To fully grasp the goal-directedness of consciousness and the hierarchical structure of intention, we must understand a central concept in Sartre's philosophy: *project*, which is "the impulse through which it [i.e., consciousness] throws itself toward its end" (Sartre 2021: 723). Every consciousness commits to certain projects. That is to say, every consciousness spontaneously determines to orient toward its object in certain ways—determines to hold P as true, determines to do P, etc. What Boyle calls "modes" of representation are in fact the ways consciousness projects its objects. Projects can be very specific—like drinking milk, jogging, and playing football. Projects can also be very general—such as "live healthily"—under which the three specific projects above are subsumed. Sartre thinks that a person's projects are ultimately unified under a "fundamental project".[14] The fundamental project is the highest unity of a person's structure of intentions and it defines the person's identity. An action is free when it is undertaken in light of one's fundamental project. Based on this, let us revisit the three examples above.

[14] "The fundamental project that I am does not concern my relations with this or that particular object in the world but my being-in-the-world as a whole" (Sartre 2021: 626).

(1) Sneezing and knee-jerk reflex are not free, because they involve no project at all. (2) Compulsive door-checking is goal-directed, and thus involves projects. Yet it is still unfree, because it does not cohere with a person's deeper structure of projects, but is only a local disorder. (3) Driving to work is both spontaneous and free, because this project can be integrated into a person's deeper structure of projects.

I will not attempt to define what precisely someone's fundamental project is—Sartre himself considers it difficult to articulate. I only want to apply this account to propose that belief is structured by a fundamental project as well: namely, the project of "truth". As long as a belief is formed in light of this project, it is free—even if it is not chosen by the will but arises spontaneously. While "the project of belief is truth" may sound unfamiliar, its ground is a widely held view in analytic philosophy, namely Bernard Williams's "beliefs aim at truth" (Williams 1973). While philosophers have long debated what it means for beliefs to have an "aim", Sartre's account provides a good understanding: belief aims at truth, because it is a spontaneous project directed toward truth.

On this basis, we can distinguish *three different senses* of "spontaneity," and the flaw in Strawson's understanding of spontaneity lies precisely in his failure to fully separate them:

1. *Mere instinct*, such as sneezing and knee-jerk reflex.
2. *Unfree spontaneity*, such as obsessive–compulsive behavior. In the domain of belief, unfree spontaneity manifests as what are called *alienated* or *recalcitrant* beliefs—e.g., unconscious beliefs, implicit biases, superstitions, etc. The reason these beliefs are unfree is not that they cannot be controlled by the will, but that they are not formed in light of belief's project of truth.
3. *Free spontaneity*, such as driving to work. In the domain of belief, free spontaneity manifests as normal, non-alienated beliefs. These beliefs are formed in light of the project of truth, and even when they are not subject to the reflective control of the will, they are still free.

4.3 The Uniformity of Agency

The preceding discussions reveal that belief does not, as doxastic particularists claim, have a *sui generis* form of agency distinct from that of everyday action. It's not the case that, the agency of everyday action lies in the *will*, while the agency of belief lies in *spontaneous* responsiveness to reasons. According to my account, first, both action and belief can be spontaneous. We can focus our consciousness on a practical goal P and spontaneously represent it as "to

be done", thereby pre-reflectively committing ourselves to act. Similarly, we can perceive that P and spontaneously represent it as "to be held true", thereby pre-reflectively believing it. We might "automatically" run to catch a departing tram, or "automatically" believe that it's raining upon seeing rain outside—both are spontaneous projects toward certain ends. Second, both action and belief can be subject to reflective control of the will. We can start a practical deliberation among options A, B, C, and conclude that A is best, thereby reflectively decide to do A. Likewise, we can start an epistemic deliberation among options D, E, F, and conclude that D is best supported by evidence, thereby reflectively decide to believe D.[15] For instance, we can resist eating delicious but unhealthy food; and we can also resist being misled by a visual illusion, insisting on a correct belief despite misleading appearances. In both cases, the exercises of the will are most evident in resisting temptation—this applies equally to action and belief. Finally, both action and belief cannot be performed "at will" in a certain sense. Doxastic involuntarists often argue that voluntary action is something we can do "just like that", without having to do anything else, such as raising one's hand (see Boespflug & Jackson 2024). Whereas we cannot believe "the U.S. is still a colony of Great Britain" "just like that". But this claim is problematic for several reasons:

First, what it means to say that we can act "just like that" is unclear. We might just as well say that, upon seeing it raining outside, we believe that it's raining "just like that"—this too is very simple and immediate like raising one's hand. Second, if "just like that" means moving one's body *purely* by an act of will, then we could imagine a case where, having decided to raise our hand, we suddenly find it paralyzed—so even bodily action is not simply a matter of pure volition. Third, and most importantly, we cannot act "at will" in a certain sense too. We cannot will ourselves to believe that "the U.S. is still a colony of Great Britain" because the reasons against it decisively outweigh the reasons for it. Similarly, we can raise our hand "just like that" when no reason weighs significantly against doing so. In cases where raising one's hand has consequences—such as signaling a desire to speak in a conference—we may be too hesitant to raise our hand "at will". In Sartre's novel "The Age of Reason", there is an interesting example (Sartre 1986; see also Webber 2009). Daniel, in an effort to prove that he is a hard, masculine person rather than a gentle and sentimental one, attempts to drown his pet cats in a river. But upon reaching the riverbank, he finds himself unable to carry it out and brings his cats home. Daniel cannot drown his cats "at will",

[15] Moran refers to the beliefs that involve agency as those that occupy the "deliberative stance" of believing (Moran 2001: 163), but this is a potentially misleading formulation. As we have seen, spontaneous beliefs can also be agential, without requiring deliberation.

because doing so would involve a profound shift in his deep projects—from that of a compassionate person to a cruel one. Such a transformation is not something one can achieve by a simple decision. He cannot undertake the project "drowning the cat", because the deeper project "being compassionate" inhibits it. Similarly, we cannot believe that "the U.S. is still a colony of Great Britain" at will. That is also because the project of believing so contradicts belief's project of "truth". The inability to act or believe "at will" stems from a conflict between reflective will and deeper projects we spontaneously commit—it does not imply that belief lacks agency. To conclude, the agency of action and belief is uniform: both involve spontaneous, pre-reflective self-determination, both are subject to reflective control of the will, and for both of them, the power of the will has its limits.

5. Arguments for the Spontaneity of Consciousness

In the previous section, drawing on Sartre's phenomenology of consciousness, I conclude that the fundamental agency underlying both action and belief lies not in the will but in spontaneity. However, this phenomenology of consciousness seems quite unfamiliar, especially for analytic philosophers. To strengthen the points of this paper, this section presents two arguments in support of the claim that spontaneity is a more fundamental form of agency than the will. Both arguments can be found in Sartre's writings. I refer to them as: (1) the phenomenological argument, and (2) the regress argument.

(1) *The phenomenological argument*. Sartre observes that in typical everyday action, the subject's consciousness is wholly absorbed in the world, not in himself:

> there is no I on the unreflected level. When I run after a tram, when I look at the time, when I become absorbed in the contemplation of a portrait, there is no I. There is a consciousness of the *tram-needing-to-be-caught*, etc., and a non-positional consciousness of consciousness. In fact, I am then plunged into the world of objects, it is they which constitute the unity of my consciousnesses... but as for *me*, I have disappeared, I have annihilated myself. There is no place for *me* at this level, and this is not the result of some chance, some momentary failure of attention: it stems from the very structure of consciousness. (Sartre 2004: 7–8)

When the tram is about to depart, I am conscious that "tram-needing-to-be-caught", not that "*I* need to catch the tram"; when Peter needs aid, I am conscious that "Peter-having-to-be-aided" (Sartre 2004: 10), not that "*I* have to aid Peter". In many such everyday actions, our consciousness is entirely "plunged into the world of objects" (Sartre 2004: 8), rather than reflectively focused on the self. In these moments, we are not exercising reflective

control of the will, but rather pre-reflectively, spontaneously deciding to act. These spontaneous actions are not governed by unfree instinctual compulsion. Though we do not reflectively thematize our decisions, we nonetheless have an implicit, non-positional consciousness of our actions. Through reflection, we can render this implicit consciousness explicit and thereby come to know our actions explicitly. After reflection, we might say: "Yes, I want to catch the tram. If I miss it, I would be late for work. Running after it is my free act, undertaken for the purpose of fulfilling my responsibility for work". But it is not reflection that makes the action free; reflection merely clarifies the reasons for action. Even without reflection, the act of running for the tram is self-determined and free.

Why can we decide to act spontaneously even while our consciousness is fully absorbed in the external world? Sartre thinks this is because perception and action are intimately connected. We are not mere spectators of external objects; objects appear to us as potential targets of action—as "doables" or "affordances". According to Sartre, our perceptual field "is not defined by pure spatial coordinates but in relation to axes whose reference is practical" (Sartre 2021: 431). In this practical frame of reference, the subject occupies the zero point, and objects are represented according to relative spatial positions. More importantly, objects are not merely spatially related to the subject but are also related to the subject's practical interests: objects are to be approached, avoided, utilized, etc. Objects also interrelate according to the subject's practical interests, offering contextual information for action and forming what Sartre calls a "hodological space": "it is crisscrossed by paths and routes; it is instrumental and the *site* of tools" (Sartre 2021: 432). For Sartre, the world is not first experienced in a neutral and disinterested manner and only subsequently endowed with meaning by the subject's practical interests; rather, the world is from the outset experienced in relation to the subject's practical interests, in a way that is already non-neutral. This helps to answer why we can act spontaneously: when we see the tram about to leave, we do not first perceive the objective event "the tram is departing" and then think "I should catch it to avoid being late", and thereby decide to catch it reflectively. Instead, we perceive the tram as "needing-to-be-caught" immediately and spontaneously commit to the project of catching it, while simultaneously non-positionally conscious of this action. From this phenomenological perspective, spontaneity is indeed a more fundamental form of self-determination than the will.

(2) *The regress argument*. Which form of self-determination—spontaneity or the will—is more fundamental? Sartre writes: "If I must come to deliberate, it is simply because it is part of my original project to take account of my motives *by means of deliberation*" (Sartre 2021: 591). In Sartre's vocabulary,

"deliberation" refers to the reflective decision of the will. What he means is that, even when we choose to deliberate, that very deliberation is spontaneous. Of course, one might also deliberately decide to deliberate, but that second-order deliberation would be spontaneous. If the level of deliberation doesn't stop somewhere, we'll have to deliberate *ad infinitum*. Spontaneity blocks the regress by providing a ground-level form of self-determination that requires no reflection. Even without reflection, consciousness can spontaneously commit to a project.[16] Reflection renders this project explicit, but it is not necessary. Sartre's regress argument thus shows that spontaneity is the regress-stopper, and that free agents must possess spontaneity which is more fundamental than the will.

6. Conclusion

According to the "doxastic spontaneism" developed in this paper, the "seeming inconsistent triad" of doxastic agency can be resolved in a way that preserves the intuitive appeal of all three propositions discussed in Section 2:

(A) We cannot believe at will, because the will as reflective self-determination is not capable of overriding the spontaneous, pre-reflective self-determination in a belief, unless reflection provides us with reasons to revise the belief.

(B) Agency is uniform. There is no *sui generis* doxastic agency which is different from the agency in everyday action. This is because both action and belief can be spontaneously determined as well as reflectively controlled by the will.

(C) Doxastic agency is not merely involved in the preceding acts that lead to belief; belief itself is an act of agency. This is because belief is agential in the form of the spontaneity of consciousness.

References

Alston, W. P. 1988. "The deontological conception of epistemic justification", *Philosophical Perspectives*, 2, 257–299.

Anscombe, G. E. M. 2000. *Intention* (2nd ed.) (Cambridge, MA: Harvard University Press).

Audi, R. 2001. "Doxastic voluntarism and the ethics of belief", in Steup 2001: 93–111.

[16] We can compare this with Sartre's regress argument for non-positional consciousness. Very briefly, if self-consciousness has to be *positional* consciousness, then there will be infinite levels of higher-order positional consciousness. To stop the regress, there must be non-positional consciousness (see Sartre 2021: 10-11). These two arguments are intimately connected, both are meant to show the pre-reflectivity of consciousness.

Boespflug, M. and E. Jackson. 2024. "Doxastic voluntarism", *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman (eds.), https://plato.stanford.edu/archives/win2024/entries/doxastic-voluntarism/

Boyle, M. 2009. "Active belief", *Canadian Journal of Philosophy Supplementary Volume*, 35, 119–147.

Boyle, M. 2011a. "'Making up your mind' and the activity of reason", *Philosophers' Imprint*, 11(17), 1–24.

Boyle, M. 2011b. "Transparent self-knowledge", *Aristotelian Society Supplementary Volume*, 85(1), 223–241.

Boyle, M. 2024. *Transparency and Reflection: A Study of Self-Knowledge and the Nature of Mind* (Oxford: Oxford University Press).

Campbell, L. and A. Greenberg. 2024. "Mental agency and rational subjectivity", *European Journal of Philosophy*, 32(1), 224–245.

Cottingham, J. 2002. "Descartes and the voluntariness of belief", *The Monist*, 85(3), 343–360.

Evans, G. 1982. *The Varieties of Reference*, J. McDowell (ed.) (Oxford: Oxford University Press).

Feldman, R. 2001. "Voluntary belief and epistemic evaluation", in Steup 2001: 77–92.

Ginet, C. 2001. "Deciding to believe", in Steup 2001: 63–76.

Glare, P. G. W. (Ed.). 2012. *Oxford Latin Dictionary* (2nd ed.) (Oxford: Oxford University Press).

Hieronymi, P. 2006. "Controlling attitudes", *Pacific Philosophical Quarterly*, 87(1), 45–74.

Hieronymi, P. 2009. "Two kinds of agency", in L. O'Brien and M. Soteriou (eds.), *Mental Actions* (Oxford: Oxford University Press), 138–162.

Korsgaard, C. M. 1996. *The Sources of Normativity* (Cambridge: Cambridge University Press).

Korsgaard, C. M. 2009. "The activity of reason", *Proceedings and Addresses of the American Philosophical Association*, 83(2), 23–43.

Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton, NJ: Princeton University Press).

Moran, R. 2012. "Self-knowledge, 'transparency', and the forms of activity", in D. Stoljar and D. Smithies (eds.), *Introspection and Consciousness* (Oxford: Oxford University Press), 211–236.

Nickel, P. J. 2010. "Voluntary belief on a reasonable basis", *Philosophy and Phenomenological Research*, 81(2), 312–334.

Parrott, M. 2015. "Expressing first-person authority", *Philosophical Studies*, 172(8), 2215-2237.

Rae, G. 2024. "Sartre on action: decentring the will", *Journal of the British Society for Phenomenology*, 55(3), 201–220.

Roeber, B. 2020. "Permissive situations and direct doxastic control", *Philosophy and Phenomenological Research*, 101(2), 415–431.

Ryan, S. 2003. "Doxastic compatibilism and the ethics of belief", *Philosophical Studies*, 114(1–2), 47–79.

Sartre, J.-P. 1986. *The Age of Reason*, E. Sutton (trans.) (Harmondsworth: Penguin).

Sartre, J.-P. 2004. *The Transcendence of the Ego: A Sketch for a Phenomenological Description*, A. Brown (trans.) (London: Routledge).

Sartre, J.-P. 2021. *Being and Nothingness: An Essay in Phenomenological Ontology*, S. Richmond (trans.) (New York: Washington Square Press).

Steup, M. 2008. "Doxastic freedom", *Synthese*, 161(3), 375–392.

Steup, M. (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue* (Oxford: Oxford University Press).

Strawson, G. 2003. "Mental ballistics or the involuntariness of spontaneity", *Proceedings of the Aristotelian Society*, 103(3), 227–256.

Webber, J. 2009. *The Existentialism of Jean-Paul Sartre* (New York: Routledge).

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