

SYSTEMATIC REVIEW

What are delusions? Examining the typology problem

Pablo López-Silva^{1,2,3}  | Miguel Núñez de Prado-Gordillo⁴  |
Victor Fernández-Castro⁵ 

¹Escuela de Psicología, Universidad de Valparaíso, Valparaíso, Chile

²Instituto Milenio para la Investigación en Depresión y Personalidad – MIDAP, Santiago de Chile, Chile

³Institute of Complex Systems of Valparaíso, Chile, Valparaíso, Chile

⁴Department of Philosophy and Religious Studies, Utrecht University, Utrecht, Netherlands

⁵Departamento de Filosofía I, Universidad de Granada, Granada, Spain

Correspondence

Pablo López-Silva, Escuela de Psicología, Universidad de Valparaíso, Valparaíso, Chile.

Email: pablo.lopez.silva@gmail.com

Funding information

EFINTEC, Grant/Award Number: PSI2016-76551-R; Fondo Nacional de Desarrollo Científico y Tecnológico, Grant/Award Number: 1221058; Leonardo 2021 Grant for Researchers and Cultural Creators; Juan de la Cierva Research Grant, Grant/Award Number: IJC2019-040199; FACSO, Grant/Award Number: 2/2021; “Proyectos de Generación de Conocimiento 2021” Ministerio de Ciencia e Innovación, Grant/Award Number: PID2021-126826NA-I00; Chilean National Agency for Research and Development (ANID); Universidad de Valparaíso; Discretionary International Exchange Award, Grant/Award Number: 223452/Z/21/Z; Dutch Research Council (NWO); FPI-UAM Predoctoral Fellowship; Autonomous University of Madrid

Edited by: Wayne Wu, Editor

Abstract

Delusions are a heterogenous transdiagnostic phenomenon with a higher prevalence in schizophrenia. One of the most fundamental debates surrounding the philosophical understanding of delusions concerns the question about the type of mental state in which reports that we label as *delusional* are grounded, namely, the typology problem. The formulation of potential answers for this problem seems to have important repercussions for experimental research in clinical psychiatry and the development of psychotherapeutic tools for the treatment of delusions in clinical psychology. Problematically, such alternatives are scattered in the literature, making it difficult to follow the current development and state of the target discussion. This paper offers an updated critical examination of the alternatives to the typology problem currently available in the literature. After clarifying the two main philosophical views underlying the dominant formulation of the debate (interpretivism and functionalism), we follow the usual distinction between *doxastic* (the idea that delusions are a type of belief) and *anti-doxastic* views. We then introduce two new sub-distinctions; on the doxastic camp, we distinguish between revisionist and non-revisionist proposals; on the anti-doxastic camp, we distinguish between commonsensical and non-commonsensical anti-doxasticisms. After analyzing the main claims of each view, we conclude with some of the most fundamental challenges that remain open within the discussion.

This article is categorized under:

Philosophy > Foundations of Cognitive Science

Philosophy > Consciousness

Philosophy > Psychological Capacities

Neuroscience > Cognition

KEYWORDS

delusions, doxasticism, philosophy of psychiatry; psychosis, psychosis

1 | THE TYPOLOGY PROBLEM OF DELUSIONS

Delusions have been historically regarded as the hallmark of psychosis. As Jaspers (1963) claims in his *Allgemeine Psychopathologie*: “since time immemorial, delusion has been taken as the basic characteristic of madness. To be mad was to be deluded” (p. 93). Nowadays, delusions are clinically characterized as a transdiagnostic phenomenon with a higher prevalence in schizophrenia (López-Silva, Harrow, et al., 2022; Rosen et al., 2016, 2022). Delusions are heterogeneous in scope, theme, and phenomenological features (Coltheart et al., 2011); and, for the last 20 years, delusions have especially attracted attention from philosophers and psychiatrists due to the ways in which their phenomenal and representational features challenge different claims about, among many others, the nature of self-awareness (Billon, 2023; Guillot, 2017; López-Silva, 2016; Stephens & Graham, 2000), the nature of rationality (Bortolotti, 2010; Campbell, 1999, 2001), self-knowledge (Bortolotti & Broome, 2009; Rothenfluch, 2020), and the ontology (Gibbs, 2000; O'Brien & Soteriou, 2009; Strawson, 2003) and the phenomenology of thinking (Gallagher, 2015; Humpston, 2022; López-Silva, 2020; Mishara & Zaytseva, 2019).

One of the most fundamental debates surrounding the conceptual understanding of delusions concerns the type of mental state they are, if any, grounds the reports that we label as delusional. When Geraldine says she is *watching* a movie, we can conclude without further discussion that her statement is grounded in a *perceptual experience*. When she says she was thinking about what it would be like to be a caveman, we can unproblematically conclude that she was *imagining* that something was the case; she is reporting a *cognitive experience*. However, the answer to this question becomes puzzling when a person claims to be dead (Cotard delusion; Berrios & Luque, 1995), that her bodily movements are under external control (Delusions of alien control; Frith, 1992), or that external agents are inserting thoughts into her mind (López-Silva, 2018; Mullins & Spence, 2003), among others. Although all these cases will be regarded as *delusional*, the question about the type of mental state grounding these reports is subject of discussion. In the literature, this debate has been called the *typology* problem of delusions (López-Silva, 2016, 2020; López-Silva & Cavieres, 2022; López-Silva, Harrow, et al., 2022).

The formulation of alternatives for the typological problem has important practical repercussions. Good scientific explanations first need to precisely define and operationalize their *explanandum*. In this sense, typological specification could be beneficial for experimental research and the development of precise therapeutic methods (Bayne & Pacherie, 2005; López-Silva, 2020). Let's say that delusions instantiate an altered version of the type of mental state M, namely, Mx. According to a vast research tradition in cognitive neuropsychiatry, the development of precise models of delusions could greatly benefit from specifying the way in which the M-type of mental state is produced and how alterations of this process would lead to the production of Mx (see Coltheart, 2015; Corlett & Fletcher, 2021; López-Silva, 2020). For example, if delusions are cases of *imaginings* (e.g., Currie, 2000), taking a close look at the causal mechanisms underpinning imagination and what happens when such mechanisms break down would provide a roadmap for neuropsychiatric research on delusions. Alternatively, if delusions are beliefs (e.g., Bayne & Pacherie, 2005; Bortolotti, 2010; López-Silva, 2018), one should have a look at processes typically thought to be involved in belief production, revision, and so on. Likewise, the development of effective psychotherapeutic tools could profit from the specification of the type of mental state that is causing struggle to a subject. Specifying the nature, features, and behavior of paradigmatic cases of M, can illuminate new effective ways of dealing with Mx types, leading to the development of specific intervention tools targeting the exact mechanisms behind their origin and maintenance.¹

Discussion of typological specifications could also inform the kind of attitude that practitioners and other people should adopt regarding a person's delusional statements. For instance, if delusions are beliefs, it could make sense to discuss with the person possible arguments for and against them; by contrast, if they are more like perceptual illusions (e.g., Hohwy, 2013), argumentative strategies would make little sense. Finally, the discussion seems to carry significant ethical consequences. As Bortolotti (2010) argues, there is a strong association between our ability to understand a person's behavior in usual “commonsensical” or “folk-psychological” terms (e.g., in terms of their beliefs, desires, intentions, and so on) and our ascriptions of autonomy and responsibility. Therefore, whether we treat delusions as a subtype of more regular mental states or not-even if pathological-can have enormous repercussions for core ethical issues in psychiatry concerning the person's agency, for example, their ability to consent to psychiatric treatment.

Despite the relevance of the typology problem, alternatives to it are scattered in the literature, making it difficult to follow the current development and state of the discussion. This paper offers an updated classification and critical examination of the main alternatives to the typology problem. This mapping exercise will mostly focus on those approaches that are most pertinent to the cognitive science of delusions, and we intend it as a necessary preliminary step before the in-depth analysis of their practical implications for research and clinical practice. After clarifying the

two main philosophical views underlying the dominant formulation of the debate (interpretivism and functionalism), we follow the usual distinction between *doxastic* (the idea that delusions are a type of belief) and *anti-doxastic* views. We then introduce two new sub-distinctions: on the doxastic camp, we distinguish between *revisionist* and *non-revisionist* proposals; on the anti-doxastic camp, we distinguish between *commonsensical* and *non-commonsensical* anti-doxasticisms. After analyzing the main claims of each view, we conclude with some of the most fundamental challenges that remain open within the target discussion.

2 | DELUSIONS AND BELIEF ASCRIPTION: FRAMING THE DEBATE

Current approaches to the typology debate mostly differ on three key inter-related issues: (a) whether delusions can be properly understood as beliefs; (b) whether traditional theories of belief (see below) can accommodate delusions; and (c) whether “commonsense” or “folk” psychology provides useful resources to understand them. To facilitate the understanding of the different positions at stake, Figure 1 offers a flowchart of the different perspectives to be discussed below, depending on their differing answers to these three issues. Table 1 provides further details about each approach.

The dominant view within the typology debate conceptualizes delusions as a type of belief (Bayne & Pacherie, 2005; Bortolotti, 2010, 2012; Bortolotti, 2020; Bayne, 2010; Bayne & Hattiangadi, 2013; Bongiorno, 2022; Clutton, 2018; López-Silva, 2018; López-Silva, Harrow, et al., 2022; McKay & Dennett, 2009; Núñez de Prado-Gordillo, 2022).² This so-called *doxastic* approach has gained traction due to its conceptual and practical advantages. For example, it seems to make sense of some of the most basic issues related to the diagnosis of delusions. Usually, delusions are reported with a considerable degree of *subjective certainty*. This issue can be plausibly explained if we conclude that delusions are beliefs because high degrees of subjective certainty seem to be characteristic of beliefs (Langdon & Bayne, 2010).³ Similarly, it has been noted that delusions are usually reported *as* beliefs by patients (Green et al., 2018; Miyazono & Bortolotti, 2014) and interpreted as such by others (Rose et al., 2014). Generally, when asked whether they really believe what they report in psychotherapy, delusional patients claim that they really do so (Bisiach & Geminiani, 1991). When asked if he really believed what he was reporting, a patient exclaimed “what do you mean? Of course! I’m not inventing

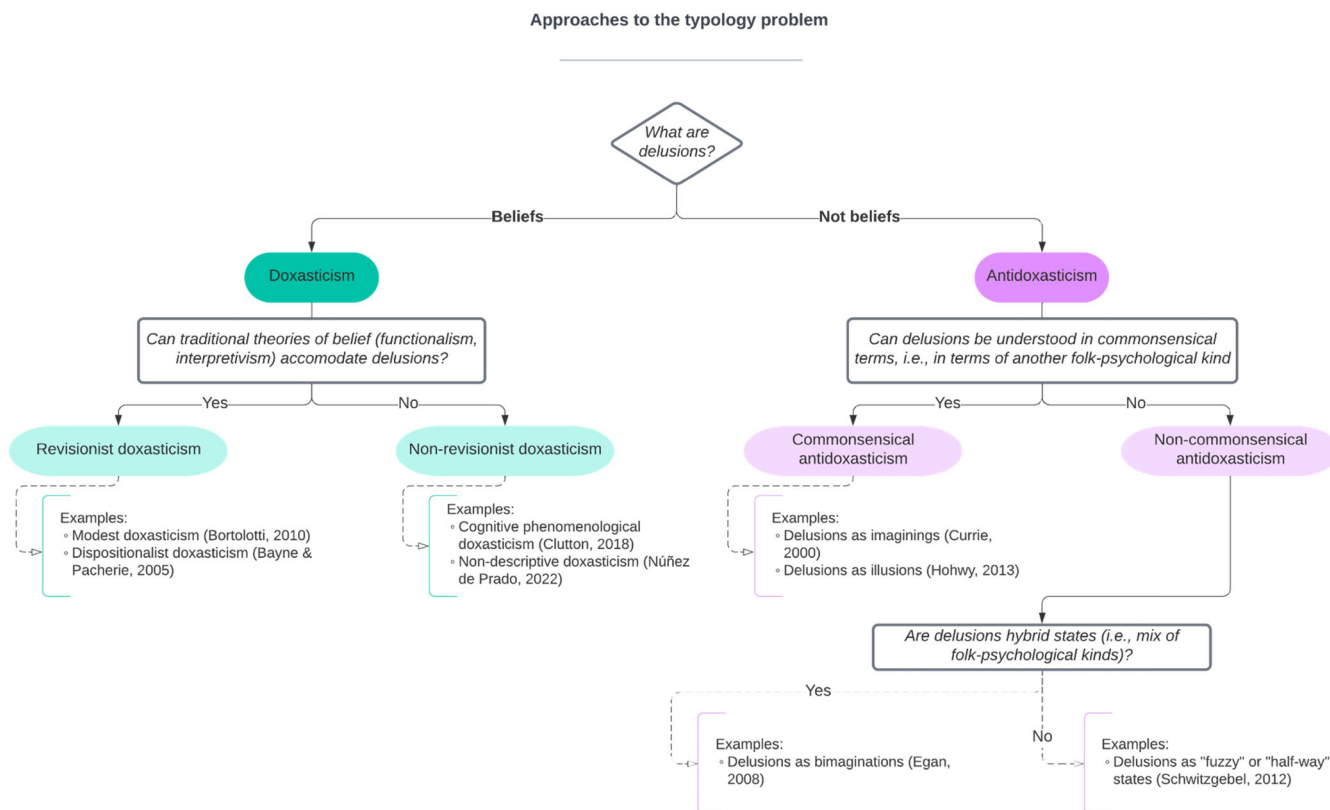


FIGURE 1 Flowchart of approaches to the typology problem.

TABLE 1 Summary definitions of main theoretical perspectives.

Theoretical account	Definition
Relevant theories of belief	
Interpretivism	A given state is a belief if it meets at least some rationality constraints: it must be truth-evaluable (contentful), responsive to evidence (epistemically rational), well-integrated with other mental states (procedurally rational), and action-guiding (agentially rational).
Functionalism	A given state is a belief if it has the typical causal roles of beliefs, i.e., if its causes (inputs) and/or effects (outputs) are belief-like.
Cognitive-phenomenological	Beliefs are dispositions to “mentally assent” to a content whenever its possibility is being considered.
Approaches to delusions	
Doxasticism	Delusions are beliefs.
Revisionist doxasticism	Traditional theories of belief (e.g., functionalism, interpretivism) can accommodate delusions once certain revisions are introduced. <i>E.g., Modest doxasticism</i> (Bortolotti, 2010), Dispositionalist doxasticism (Bayne & Pacherie, 2005).
Non-revisionist doxasticism	Traditional theories of belief cannot accommodate delusions; instead, a different theory is needed to explain why delusions are beliefs. <i>E.g., Cognitive Phenomenological Doxasticism</i> (Clutton, 2018), Non-descriptivist Doxasticism (Núñez de Prado-Gordillo, 2022).
Anti-doxasticism	Delusions are not beliefs (because traditional theories of belief cannot accommodate them).
Commonsensical anti-doxasticism	Delusions are not beliefs, but they can be still understood in terms of another folk-psychological kind (e.g., imaginings). <i>E.g., Metacognitive account</i> (Currie, 2000), Illusion account (Hohwy, 2013).
Non-commonsensical anti-doxasticism	Delusions are not beliefs nor can be understood in terms of any other folk-psychological kind. Instead, they must be understood as hybrid states (i.e., a mix of folk-psychological attitudes, like imaginings and beliefs) or in-between, “fuzzy,” or “half-way” states that do not amount to any folk-psychological state. <i>E.g., Bimagination account</i> (Egan, 2008), Dispositionalist anti-doxasticism (Schwitzgebel, 2012).

it!” (López-Silva, 2015). By defining delusions in clear conceptual terms, the doxastic approach provides a robust conceptual framework to guide empirical research on delusions (Coltheart, 2002, 2015; Davies, 2000). Once we view delusions as beliefs, researchers could focus on the way human beings come to form beliefs and understand the different alterations of these mechanisms that could give raise to delusions (Coltheart, 2015). Naturally, here the challenge is to explore and comprehend such mechanisms in adequate ways.

In clinical terms, the DSM-V definition of delusions is an example of the influence of the doxastic approach in current clinical psychiatry, where they are defined as:

[...] fixed beliefs that are not amenable to change in light of conflicting evidence. [...] Delusions are deemed bizarre if they are clearly implausible and not understandable to same-culture peers and do not derive from ordinary life experiences. [...] The distinction between a delusion and a strongly held idea is sometimes difficult to determine and depends in part on the degree of conviction with which the belief is held despite clear or reasonable contradictory evidence regarding its veracity. (APA, 2022, p. 102)

However, the *standard doxastic approach* exemplified in this definition has been challenged by certain *theories of belief* within analytic philosophy. These theories emphasize the analysis of belief-like phenomena from a third-person perspective, which focuses on the criteria that *external observers* typically apply when ascribing beliefs to another person. In this context, the doxastic view of delusions has been challenged under the claim that delusions fail to meet the criteria that these theories of belief establish for appropriate belief ascription (see Frankish, 2009, p. 269). These *anti-doxastic* views claim that many people with delusions often fail to reason, act, or react on their delusions (see Berrios, 1991; Currie, 2000; Currie & Jureidini, 2001; Egan, 2008; Frankish, 2009, 2012; Graham, 2010; Hamilton, 2007; Hohwy & Rajan, 2012; Murphy, 2012; Radden, 2011; Sass, 1994; Schwitzgebel, 2012; Tumulty, 2011, 2012; Young, 1999), and that this poses a threat to their interpretability as beliefs. For example, people with Capgras delusion

(the delusion that a partner has been replaced by an identically looking impostor) presumably fail to reason and act upon what they claim to believe; they continue to live with the impostor as they did with the replaced partner, they fail to offer any “excuse” as to why the impostor knows every single detail of their relationship with the replaced partner, and so on (see Coltheart et al., 2011). This way of approaching the typology problem has been mainly framed by two interrelated theoretical frameworks: *interpretivism* and *functionalism*. Before delving further into the alternatives to the typology problem, we will examine the main features of these two theories of belief.

2.1 | Interpretivism and functionalism about belief

According to interpretivist approaches, the concept of belief is primarily individuated in terms of rationality criteria (Davidson, 1986; Dennett, 1979; Dennett, 1987; see also Byrne, 1998). For Bortolotti (2010), this view entails that a belief can be ascribed to an agent only when the agent's intentional state is: (a) *contentful*, that is, evaluable in terms of its truth or falsity; (b) *epistemically rational*, that is, grounded on sufficient evidence and responsive to counter-evidence; (c) *procedurally rational*, that is, well integrated in the agent's belief system and appropriately connected with other mental states; and (d) *agentially rational*, that is, action-guiding and endorsed on the grounds of intersubjective good reasons (see also Bayne & Pacherie, 2005; Miyazono & Bortolotti, 2014). For example, considering the first criterion, this would mean that an agent cannot really *believe* things like Lewis Carroll's senseless *Jabberwocky* verses (e.g., “All mimsy were the borogoves”), because those ideas cannot be assessed in terms of their truth or falsity. On the other hand, some cases of self-deception would fail to meet the other three criteria. In those cases, a person claims to believe something (e.g., that their boss is kind and caring) despite there being clear evidence to the contrary (e.g., clear signs of psychological abuse, others' testimonies of similar experiences, etc.) and even when the person's own implicit cognition and behavior (e.g., avoidance patterns) contradicts such belief. According to interpretivism, these violations of the principles of epistemic, procedural, and agential rationality would preclude the analysis of these cases as beliefs.

According to functionalist approaches, the concept of belief is primarily individuated in terms of its stereotypical *causal* roles, that is, by their typical causes or *inputs* (e.g., environmental information) and, especially, their typical effects or *outputs* (behavior, other mental states, etc.; see Block & Fodor, 1972; 1980; Putnam, 1967; Putnam, 1975). Beliefs and other mental phenomena thus aim to fill the alleged “explanatory gaps” between perception, cognition, and action. So, for functionalists, a certain mental state is a belief if and only if it displays belief-like functional roles. For instance, the belief “the cat is hungry” would be individuated by its typical functional roles, for example, its being caused by hearing the cat meowing near the feeder, its causing one to believe that the cat has not been fed in a while, its action-orienting role in searching for cat food, and so on.

Following the classic distinction, the functionalist stance can be fleshed out either in *occurrentist* (e.g., Carruthers, 2013) or *dispositionalist*⁴ terms (e.g., Schwitzgebel, 2013). According to occurrentism—or *standard representationalism*—“beliefs are occurrences (e.g., phenomenological states or distinctive activations of the cognitive system)” (Nottelman, 2013, p. 23); thus, “to believe is to have a representation that plays belief roles” (Miyazono & Bortolotti, 2014, p. 32). On the other hand, according to *dispositionalist* accounts of belief, beliefs are dispositions to behave, cognize, or experience in certain ways (Nottelman, 2013). Schwitzgebel (2002, 2012, 2013) is one of the most relevant ones for the typology problem, as described below. According to the Schwitzgebel, beliefs are very much like personality traits: just like introverts are characterized by certain behavioral, cognitive, and experiential tendencies or dispositions (e.g., avoiding social stimuli, enjoying lonely tasks, etc.), different beliefs would also be characterized by specific dispositional profiles. For instance, the belief that all races are equal would be characterized by the behavioral, cognitive, and phenomenological patterns that we typically associate with such belief (e.g., calling out bigotry, scrutinizing, and trying to correct one's internalized racism). For Schwitzgebel, a person's dispositional profile is the most important thing for belief ascription, regardless of the exact causal mechanisms responsible for them; when the relevant dispositions are absent, we cannot really ascribe the associated belief to the person. For instance, cases of implicit bias, where a person's explicit non-prejudicial statements about a group do not cohere with their implicit prejudicial cognitive, behavioral, or experiential tendencies, could not be properly understood as beliefs, precisely because of the observed mismatch between the person's actual tendencies and the ones typically expected from someone who really had non-prejudicial attitudes (Schwitzgebel, 2013).

Interpretivism and functionalism (at least some variants of it) share some important features. For instance, both interpretivism and dispositionalist functionalism avoid deep ontological commitments about beliefs by focusing on the conditions that an agent's behavior, reasoning, and experiential tendencies have to meet for the agent to be successfully

described as having a certain belief, irrespective of the underlying mechanisms that cause such tendencies -an issue that is left for cognitive scientists to discover (see Byrne, 1998).⁵ Furthermore, functionalism could be seen as adding a further conceptual commitment to the interpretivist proposal: the idea that an agent's intentional state only counts as a belief if it is in an appropriate or rationally understandable *causal* relation with the available evidence, with other beliefs and mental states, and with the agent's subsequent actions and reason-giving (see Bayne & Hattiangadi, 2013). In turn, some interpretivist approaches conceive of mental explanations as a particular kind of causal-predictive tool, one reserved to rational creatures like human beings. According to Dennett, 1987, 1979, adding to the *physical stance* and the *design stance* adopted when we explain an agent's behavior in purely physical or biological terms, we can adopt an *intentional stance* toward the explanation of the behavior of rational creatures (see also McKay & Dennett, 2009). When we causally explain and predict an agent's behavior in terms of their mental states (e.g., when we explain our partner's going to the kitchen in terms of their *desire* to eat and their *belief* that there is food on the fridge), we are taking such an intentional stance. Anti-doxastic argumentation, as we will now see, has built up on these two theories of belief.

2.2 | Anti-doxasticisms about delusions

The doxastic conception of delusions has been contested from interpretivist and functionalist fronts. The most discussed examples concern *monothematic delusions*, single belief-like states, or small sets of belief-like states that are held towards a single theme. Cases of Capgras, Cotard, and mirrored-self misidentification delusions (where people fail to identify themselves in the mirror, thinking instead that they are looking at a stranger) are among the most widely discussed cases in the literature. On the contrary, *polythematic delusions*—more common in schizophrenia—have not been so widely discussed. The reason is that, unlike polythematic delusions, monothematic ones are often extremely circumscribed and disconnected from the person's other mental states and action, which poses a threat to their interpretability as beliefs from a functionalist/interpretivist framework.⁶ This is precisely why antidoxasticists claim that delusions -at least monothematic ones- are not beliefs. However, it is not always clear what exact theoretical framework supports each anti-doxastic argument. For the sake of clarity, let's discuss the two main arguments made by anti-doxasticists separately.

From an interpretivist anti-doxastic viewpoint, delusions do not qualify as beliefs because people that experience delusions usually fail to meet the relevant rationality constraints. We will call this the *Rationality Constraint argument* (RC) against Doxasticism about delusions (see Bortolotti, 2010, 2012), and it runs as follows:

Rationality constraint argument

Premise 1 (Interpretivism): A given state is a belief if it meets at least some rationality constraints: it must be truth-evaluable (contentful), responsive to evidence (epistemically rational), well-integrated with other mental states (procedurally rational), and action-guiding (agentially rational).

Premise 2 (Empirical observation): Delusional cases (at least some) fail to meet either one or all of these rationality constraints.

Conclusion: Thus, delusions (or at least many of them) do not count as beliefs.

According to functionalist anti-doxasticists, many people with delusions systematically fail to display the behavioral, cognitive, or phenomenological patterns expected if they really believed the content of their reports. Delusions do not properly fit the stereotypical causal profile of paradigmatic belief, so doxasticism is misplaced. We will call this the *Stereotypical Causal Profile argument* (SCP) (see Miyazono & Bortolotti, 2014; see also Miyazono, 2019), and it runs as follows:

Stereotypical causal profile argument

Premise 1 (Functionalism): A given state is a belief if it has the typical causal roles of beliefs, that is, if its causes (inputs) and/or effects (outputs) are belief-like.

Premise 2 (Empirical observation): Delusions (at least many of them) fail to play belief-like causal roles.

Conclusion: Thus, delusions (or at least some) do not count as beliefs.

As we can see, anti-doxastic approaches are fundamentally defined in terms of a negative thesis: they deny a belief status to delusions because (a) people with delusions do not behave, cognize, or experience as it would be rational to expect of them if they really believed their delusions; or (b) delusions do not display stereotypical belief-like causal roles.⁷ Anti-doxasticists have mainly focused on “output” inconsistencies (see Schwitzgebel, 2012), that is, inconsistencies between the person’s delusion and their other mental states and behavior. This comprises what has been called the *bad integration* and *double-bookkeeping* objections to doxasticism about delusions (e.g., Bortolotti, 2010, 2018; Gallagher, 2009). Bad integration refers to “obvious mistakes in deductive reasoning, or fail[ing] to obey basic inferential rules governing the relations among beliefs and other intentional states” (Bortolotti, 2010, p. 62); that is, displaying attitude-attitude inconsistencies. It thus compromises the assumption of procedural rationality or of the appropriate causal relations among beliefs and other intentional states. Claiming to believe that some omnipotent divine force dictates one’s fate while thinking that one can “cheat on” it would be a good example of badly integrated beliefs. On the other hand, double-bookkeeping is a phenomenon where a certain agent, despite claiming to believe that *p*, behaves in ways inconsistent with what they claim to believe; in other words, the agent displays attitude-behavior inconsistencies, which compromise the assumption of agential rationality or of the appropriate causal relations between belief and action.

2.3 | What then? The positive aspect of anti-doxasticism

Anti-doxastic approaches are thus unified by a negative thesis: delusions are not beliefs. However, they differ on the typological alternative they propose. Considering the bizarre character of some delusions, some authors have proposed *non-assertoric approaches*. For example, Berrios (1991) claims that delusional statements are *contentless* (they are “empty speech acts,” in his terms). Less radical, Sass (1994, 2004) proposes to understand delusions in metaphorical terms; delusional statements are contentful, but their content cannot be understood literally. Similarly, others have criticized standard doxasticism on the grounds that delusions, more than doxastic deviances, are best characterized in terms of their specific experiential properties (e.g., Hohwy & Rajan, 2012; Radden, 2011). In a stronger version of this claim, delusions would involve a whole experiential reality shift: they would constitute “alternative realities” (Gallagher, 2009). Although Sass’s and Gallagher’s perspectives introduce interesting elements, we will not discuss them further. The main reason is that these approaches might not be properly characterized as “anti-doxastic”; after all, claiming that some delusions involve certain alterations of the structure of experience is not incompatible with claiming that they are beliefs, taken as endorsements of such experiences (Sass, 2004, p. 77; see also Bayne & Pacherie, 2004b; see also López-Silva, 2020). Thus, we will focus here on anti-doxastic views that attempt to understand delusions in terms of other kinds of mental states different from belief. Here, two strands can be distinguished: (a) *commonsensical anti-doxasticism*, which rejects doxasticism about delusions, but not their interpretation in folk-psychological terms; and (b) *non-commonsensical anti-doxasticism*, which rejects both doxasticism and the folk-psychological conceptual framework altogether.

2.3.1 | Commonsensical and non-commonsensical anti-doxasticism

Some anti-doxasticists claim that delusions can be explained in terms of regular mental states other than belief. This *reclassification* strategy rejects doxasticism, but still assumes that delusions can be conceptualized in folk-psychological terms (Bayne & Hattiangadi, 2013). An example of this kind of *commonsensical anti-doxasticism* is the meta-cognitive view defended by Currie (Currie, 2000; Currie & Jureidini, 2001). This proposal states that delusions are not beliefs, but *imaginings* that the person mistakes for beliefs. This would explain why some delusions are not properly acted upon. When we imagine—instead of believing—that a certain state of affairs is the case, we might take the imagined content into action or not. If we just imagine that there is a divine force that might punish us if we don’t comply with a series of obligations, we might sometimes act, reason, or feel in accordance with the imagined content, but we are no longer expected to do so.

One major problem with this account is that it does not consider the fact that many people with delusions do display belief-like behaviors (see López-Silva, 2018). For example, many people with delusions try to back up their delusions with reasons, and they cannot but feel convinced and try to convince others about the truth of their thoughts. This feature is at odds with a purely imaginative account of delusions (Bayne & Pacherie, 2005; Bortolotti, 2010;

Radden, 2011). Furthermore, even if they did not, it would not be clear whether the meta-cognitive approach solves the problems it aims to solve. On this account, people with delusions believe that they believe the delusional content (hence the “meta-cognitive” character of delusions), but in fact they just imagine it. Yet one might then wonder: why do not they act in accordance with what they believe to believe? Why wouldn't they be rationally compelled to act in accordance with what they believe to believe?⁸

Alternatively, one might think that delusion attributions appear precisely “when we run out of the explanatory resources provided to us by our folk understandings of how the mind works” (Murphy, 2012, p. 22). *Non-commonsensical anti-doxasticists* thus assume that folk psychology offers poor explanatory resources for understanding delusions, since their features defy everyday assumptions about how we usually function. One possible way to implement this kind of non-commonsensical anti-doxasticism is to follow a *rebranding* strategy; if folk-psychological resources are not good enough for a proper scientific account for delusions, a new *sui generis* type could be crafted for them (see Bayne & Hattiangadi, 2013). Along these lines, some authors have proposed that delusions—together with other odd states—are *hybrid* states. For example, Egan (2008) has proposed to understand delusions as *bimagnations*, that is, mental states that share the functional profile of both beliefs and imaginations, but which fail to fit any of them fully. Egan's (2008) “bimagnations” are precisely supposed to account for the observed deviances from both belief-like and imagination-like causal stereotypes.

More radically, Schwitzgebel's (2012) dispositionalist anti-doxasticism claims that delusions are not “in-between” states in the sense that they are instances of some hybrid, middle-ground new mental type. Rather, delusions just don't amount to any folk-psychological state; they are just “fuzzy” or “half-way” states that fail to conform to our usual folk-psychological expectations. For Schwitzgebel, what is “in-betweenish” is our belief ascriptions: when the agent's dispositional profile does not fully meet the stereotypes that laypeople usually associate with a certain belief, it is just not fully correct nor incorrect to describe it in doxastic terms. But that should not worry us, nor scientists for that matter. Instead of crafting a new mental type for delusions, we might adopt an operationalist perspective: if the behavioral, cognitive, and phenomenological activity of people with delusions cannot be properly characterized in terms of usual mental states, then let's just exhaustively specify their dispositional layout. We might choose to coin a new name for it or not, but that's unessential; once this dispositional profile has been fully specified, scientists are left free for determining its natural causes.

2.4 | Believing my mind: doxasticism about delusions

Different pro-doxastic approaches have replied to the challenges posed by anti-doxasticists. We may here distinguish between revisionist and non-revisionist defenses of doxasticism. *Revisionist doxasticisms* assume that functionalism and interpretivism still grant a doxastic status to delusions when certain revisions are introduced in these frameworks. We will primarily discuss Bortolotti's (2010) and Bayne and Pacherie's (2005) approaches; these mostly focus on the folk-psychological notion of belief, taken to be at least partially individuated by its stereotypical or ideal rational/causal profile (see also Bayne & Hattiangadi, 2013). On the other hand, *non-revisionist doxasticism* assumes that functionalism and interpretivism are inadequate theories of belief and should thus be replaced by a different theoretical framework (Clutton, 2018; Núñez de Prado-Gordillo, 2022). Here we will focus on Clutton's (2018) approach, which focuses not so much on the folk-psychological notion of belief, but rather on cognitive-scientific understandings of it that emphasize its cognitive-phenomenological character (i.e., the feeling of “mental assent” that often characterizes believing) and scientific claims about their neural underpinnings.

2.4.1 | Revisionist doxasticism

Revisionist doxasticisms do not question the theoretical framework behind anti-doxasticism—namely, interpretivism or functionalism; instead, they examine to what extent these two theories of belief really motivate anti-doxasticism towards delusions and recommend local revisions of the theoretical background. Although the two kinds of revisionist doxasticism that we will examine here share many of the arguments in favor of doxasticism, we'll consider them separately depending on whether they put a greater emphasis on the articulation of a response to RC or to SCP.

2.4.2 | Modest doxasticism

Bortolotti's (2010, 2011, 2012), Miyazono & Bortolotti, 2014; see also Bortolotti, 2018) modest doxasticism is one of the most widely discussed defenses of doxasticism. Bortolotti focuses on RC, rejecting both of its premises. Her strategy is thus twofold: (i) she asks if delusions really fail to meet the standards of epistemic, procedural, and agential irrationality (empirical question, concerning premise 2)⁹; and (ii) she asks if interpretivism is able to accommodate our straightforward understanding of many non-clinical phenomena in doxastic terms (conceptual question, concerning premise 1).

Regarding the empirical question, Bortolotti points out that many delusional cases do fit well with the criteria imposed by interpretivism for a mental state to qualify as a belief, *contra* the second premise of the RC argument (see also Bayne & Pacherie, 2005; Reimer, 2010). First, many delusions might be grounded on an insufficient evidential basis, but this does not seem to preclude our interpretation of them in doxastic terms; after all, many ordinary beliefs are formed on the basis of insufficient evidence, and do not change easily when confronted with counterevidence. Second, regarding the bad integration and double-bookkeeping objections, it seems that many people with delusions do in fact reason and act upon their delusions (see Young, 1999). Bortolotti (2010, pp. 69–70, 164–165) gathers evidence of this from the clinical literature. For example, people with Cotard delusion (i.e., the delusion that one is dead) sometimes stop eating and bathing; in addition, they sometimes justify why, despite being allegedly dead, they can move and talk (for example, because they already are in Heaven). In this sense, Young and Leafhead (1996) pointed out that all Cotard patients showed some form of delusion-related behavior. In fact, a great number of people experiencing delusions do act on their delusional beliefs (de Pauw & Szulecha, 1988), which, on certain occasions, might even lead to severe consequences. Blount (1986) reports a case of a patient suffering from Capgras delusion who decapitated his stepfather trying to find the batteries in his head. Similarly, a patient with perceptual delusional bicephaly tried to remove his second head through various means (first with an axe, later shooting at it), suffering a number of injuries (see Ames, 1984). Several patients with delusions of superhuman strength have been reported injured after acting on their delusions (Petersen & Stillman, 1978). Finally, cases of paranoia usually entail a wide pattern of avoidance and escape behaviors. These cases show that, at least sometimes, delusions do guide action.

Regarding the conceptual question, related to premise 1 of RC, Bortolotti claims that classical interpretivism yields a too stringent theory of belief; one that not only excludes delusions from the realm of belief, but also a vast amount of other non-clinical phenomena that we naturally interpret in terms of irrational beliefs (such as superstitious, contradictory, or poorly acted-upon beliefs, etc.). In this sense, Bortolotti argues that the interpretivist's rationality constraints should not be seen as *constitutive* of belief (i.e., as determining what can count as a belief), but as *guiding* criteria for belief ascription: procedural, epistemic, and agential rationality are just *regulative* ideals that laypeople usually (but not necessarily) follow when ascribing beliefs to one another. Relatedly, Bortolotti (2010, p. 262) emphasizes the context-relative nature of belief ascription, that is, the idea that belief ascription is influenced by considerations about how certain variables in the person's environment or psychological state might favor or difficult their ability to meet the relevant rationality constraints. Finally, Bortolotti (2010, p. 99) argues that classical interpretivism fails to distinguish “between two notions of rationality—rationality as conformity or subscription to epistemic norms, and intelligibility of observed behaviour.” In the former, stronger sense of the term, to be a rational agent involves an achievement, which entails the systematic ability to draw correct inferences from the available evidence, to act upon one's self-professed mental states, and to ground one's judgements in intersubjective good reasons. In the weaker sense of the term, none of this is necessary: a rational (i.e., intelligible) agent is just one whose behavior can be regarded as meaningful or purposeful; that is justified by some *reason*, regardless of whether such reason is intersubjectively good or whether the agent displays an overall rational (in the strong sense) pattern of activity.

Taking all this into account, Bortolotti proposes a relaxation of the interpretivist's requirements for a mental state to count as a belief. First, instead of being fully procedurally rational, Bortolotti proposes that beliefs must just have *some* inferential connections with other beliefs and mental states. Second, regarding the standard of epistemic rationality, Bortolotti holds that beliefs need not be responsive to evidence (i.e., they need not change in light of contradictory evidence); they just need to be *sensitive* to it; in other words: all it takes for an intentional state to count as a belief is that it can potentially change in light of contradictory evidence, even if it does not in many occasions. Finally, an intentional state doesn't have to be action-guiding, in a strong sense of the term, nor endorsed on the basis of intersubjectively good reasons; for Bortolotti, an intentional state might count as a belief if it is (a) *behaviorally manifestable*, that is, it must potentially lead to action in some of the relevant circumstances; and (b) endorsed on the grounds of *subjectively* good reasons (i.e., reasons the agent regards as good reasons).

2.4.3 | Dispositionalist doxasticism

Bayne and Pacherie's (2005); see also Bayne & Pacherie, 2004a, 2004b) approach aims at rejecting the SCP argument against doxasticism. They also attack both the empirical and conceptual premises of SCP, as Bortolotti (2010) does with the RC argument. First, Bayne & Pacherie gather evidence showing that many people with delusions do in fact exhibit the behavioral, cognitive, and phenomenological patterns that one would expect of a standard believer. Consequently, they claim that the second premise of SCP is not empirically warranted (see also Bayne & Pacherie, 2004a, p. 6).

Second, Bayne and Pacherie (2004a, 2005) also revise the first conceptual premise. Specifically, Bayne and Pacherie's defense builds up on Schwitzgebel's (2002) functionalist-dispositionalist account of belief¹⁰; however, they emphasize a *ceteris paribus* clause (i.e., an "all things equal" condition) that is present in this theoretical framework but is often overlooked in the debates about the typology problem. According to this clause, a mental state is a belief if it displays belief-like causal roles *in standard or normal circumstances*. Just like introverts may display extrovert behavior in non-standard situations (e.g., under the effect of substances that lower social inhibition), there are circumstances where a person's belief-dispositions cannot be manifested (e.g., not expressing one's political beliefs when under coercion). Such deviations from the dispositional profile can nonetheless be *excused* by these non-standard circumstances. In those cases, we might still claim that the agent has the relevant disposition (and therefore the relevant belief); they just fail to manifest it due to the excusing condition. Furthermore, Schwitzgebel also states that, even in cases where some of the relevant dispositions are actually missing, we might still ascribe the belief to the agent as an "ascriptive shorthand," in order to facilitate communication with a given audience (see also Tumulty, 2011, 2012). Here, the ascriber's interests and those of their audience play a decisive role in determining whether the belief ascription is *convenient* (rather than fully true) or not. Child rearing often involves ascribing beliefs to infants even when they have not yet learned how to fully act upon them; for instance ascribing a child the belief that sharing toys with a sibling is something good-even when they are not precisely inclined to share them-is often a way to teach them how to be good siblings.

According to Bayne and Pacherie (2004a, 2004b, 2005), the same applies to delusions; even when they systematically deviate from the relevant dispositional stereotype, these deviations can often be *excused* by appeal to some non-standard circumstance. Furthermore, when "a deviation from the stereotype cannot be excused or explained in this way, whether or not the attributor ascribes the belief will depend on the context of the belief ascription and *what her interests are*" (Bayne & Pacherie, 2005, p. 181). They discuss several possible factors that could explain the apparent inconsistencies that some people with delusions exhibit, mainly focusing on two: environmental pressures and non-standard or disrupted perceptual, motivational, or affective conditions. People with delusions might not systematically engage in delusion-consistent behaviors to avoid the risk of being hospitalized or detained, or to avoid being labeled "crazy" and consequently stigmatized, and so on. In addition, since people with delusions allegedly have anomalous perceptual and affective and motivational experiences,¹¹ Bayne & Pacherie hold that it would not be strange that similar anomalous processes could be responsible for the agent's deviations from the stereotypical profile.

2.5 | Non-revisionist doxasticism

Bortolotti's, Bayne and Pacherie's defenses of doxasticism draw from the very same theories of belief that first gave rise to RC and SCP, interpretivism and functionalism. On the contrary, several recent approaches dispute the adequacy of these theoretical frameworks for understanding delusions or providing a firm defense of doxasticism (Clutton, 2018; Núñez de Prado-Gordillo, 2022). *Expressivist* accounts of delusions (Núñez de Prado-Gordillo, 2022; Pérez-Navarro et al., 2019; Wilkinson, 2020), on the one hand, go deeper than interpretivism and certain varieties of functionalism in their rejection of any strong ontological commitments about the factual nature of delusions. Wilkinson (2020), for instance, argues that delusion attributions are non-descriptive, that is, they do not describe some fact about the person's psychology nor dispositional profile, but rather just express a (negative) evaluation of it. Diagnoses of delusions from this perspective would be more like moral condemnations (e.g., "killing is wrong") than factual statements about the person's behavioral profile or its neural underpinnings. Pérez-Navarro et al. (2019), apply a similar view to belief attributions in general, presenting the whole debate on the typology problem of delusions as an example of their evaluative, rather than descriptive nature. Finally, while neither of these two approaches explicitly accepts nor rejects doxasticism about delusions -for instance, Wilkinson (2020) just adopts it strategically-, Núñez de Prado-Gordillo (2022) draws from a similar view to defend it; specifically, Núñez de Prado-Gordillo (2022) emphasizes how expressivism and related

non-descriptivist frameworks are better positioned than functionalist and interpretivist theories to explain the usual conceptualization of delusions as beliefs.

However, in treating belief/delusion ascriptions more like moral condemnations than factual statements, these views are mostly agnostic or skeptical regarding the connection between the typology problem and its implications for the cognitive science of delusions. Here we are thus mostly interested in non-revisionist views that stress this connection. Clutton (2018) has recently advanced such account. Instead of introducing local adjustments to the interpretivist or functionalist frameworks, Clutton (2018) has proposed to reject both altogether, due to their alleged “anti-realist tendencies” toward belief (p. 11).¹² Specifically, Clutton problematizes their lack of a clear ontological individuation of beliefs as *real* entities, separated from the dispositional profiles that characterize them. According to Clutton (2018), this attitude is incompatible with what he calls *scientific doxasticism* about delusions, that is, the “robustly realist” (p. 11) doxastic approach that figures in prominent cognitive models of delusions (e.g., Alford & Beck, 1994; Coltheart, 2007; Coltheart et al., 2011; Ellis & Young, 1990; Freeman et al., 2002; Garety, 1991; Frith, 1992; Maher, 2005, 1974). On his view, the respectability of these scientific theories is in itself a very good reason to *assume* that delusions are beliefs, and that these are real entities with an independent ontological status. Clutton's view thus focuses on providing a philosophical account of belief that can fully accommodate the realist commitments at play in traditional cognitivist models of delusions.

Clutton's view appeals to Kriegel's (2015) cognitive phenomenological theory of beliefs. On this view, beliefs are cognitive-phenomenological dispositions, that is, dispositions to “mentally affirm” or “judge” that something holds true whenever one entertains that possibility (Clutton, 2018, p. 4). From this perspective, all it would take for someone to truly have anti-racist beliefs would be that they “mentally assent” to things like “all races deserve equal treatment” whenever they entertain that possibility in their minds. According to Clutton, when we judge that something is the case (i.e., when we consider some proposition *p* to be true), “we feel [...] a sense of mental affirming, as we entertain the proposition” (Clutton, 2018, p. 4). That feeling of assent, which might come in different degree of conviction, is the distinctive phenomenological mark of belief (one that is absent when the agent is just imagining something, for example). According to Clutton, these *epistemic feelings* are analogous—though irreducible—to their sensory counterparts (e.g., seeing red), but distinct in their epistemic nature.

Thus far, the cognitive phenomenological proposal would not be so different from Schwitzgebel's dispositionalist approach, except that the latter individuates beliefs in terms of complex sets of dispositions, where not only cognitive-phenomenological, but also other dispositions are considered. However, unlike Schwitzgebel, Clutton's stresses the link between cognitive-phenomenological dispositions and their neural bases. On Clutton's view, having the disposition to have occurrent episodes is for one's neural system to be set such as to trigger the relevant episode of “mental affirming” in response to the triggering conditions; such neural setting “is the ‘truth-maker’ of the disposition, the categorical grounds in virtue of which a belief ascription can be true.” (Clutton, 2018, p. 5). Thus, contrary to interpretivism and functionalism, Clutton's cognitive phenomenological approach is strongly committed to realism about beliefs: beliefs are real entities, neural states whose triggering produces in the subject a distinctive kind of cognitive-phenomenological experience, that is, that of judging that *p*, with its distinctive epistemic feeling of affirming the entertained proposition. Finally, in the cognitive phenomenological approach, the agent has a special epistemic access to their own mental states; while others can only access an agent's mental states through a mediate, inferential strategy (i.e., via the observation of the agent's behavior), the agent herself has direct or acquaintance knowledge of their own beliefs.

Importantly, these last two features of Clutton's proposal (i.e., the “privileged access” conception of self-knowledge and the strict identification of the phenomenological disposition to judge that *p* with a certain neural state) introduce a radical departure from traditional defenses of doxasticism. As aforementioned, interpretivist and functionalist views like those endorsed by revisionist doxasticists emphasize third-personal accounts of belief, which focus on the criteria that external observers apply in regular belief ascription. By contrast, the cognitive phenomenological approach emphasizes the first-person perspective. Specifically, on this account, there are only two *direct* sources of evidence for one's beliefs: (i) from a first-person perspective, one's privileged access to one's own “mental affirming” dispositions; and (ii) from a third-person perspective, the individual's neural states. Not only the person's behavior, but also any cognitive and phenomenological activity other than that of “privately judging that *p*” are regarded as mere indirect and inconclusive sources of evidence about the person's beliefs. From a third-person point of view, the strongest *indirect* source of evidence would be the person's sincere reports of their beliefs, since such reports are taken to communicate the person's subjective experience (Clutton, 2018, p. 6).

These features of the cognitive phenomenological view of belief straightforwardly deal with the SCP and RC arguments; since the most reliable indicators of an agent's beliefs are their sincere self-ascriptions and, more importantly,

their neural activity, whether they display attitude-attitude or attitude-behavior inconsistencies is to some extent irrelevant. If someone sincerely claims to be systematically disposed to mentally affirm something, or if a neurologist determines that such and such neural states correspond to such cognitive-phenomenological disposition, then we should take the person's self-report or their neurological analysis at face value. This way, the cognitive phenomenological theory of belief aims at accommodating the realist assumptions behind various cognitive models of delusions: beliefs are real entities, which cause belief-like patterns of behavior, cognition, and experience, and are supposed to be equivalent to certain states of the person's neural circuitry.

3 | CONCLUDING REMARKS: THE FUTURE OF THE DOXASTIC APPROACHES TO DELUSIONS

This paper has critically examined the current available alternatives to the so-called typology problem of delusions, namely, the discussion about the type of mental state that underlies reports regarded as *delusional*. After clarifying the two main philosophical views in which the traditional formulation of the debate is supported (i.e., interpretivism and functionalism), we have followed the usual distinction between *doxastic* and *anti-doxastic* views. Furthermore, we have introduced two new sub-distinctions that offer guidance in the interpretation of the different proposals. On the doxastic camp, we have distinguished between *revisionist* and *non-revisionist* proposals. On the anti-doxastic camp, we have distinguished between *commonsensical* and *non-commonsensical* anti-doxasticisms. This mapping offers guidance for future in-depth assessments of the scientific and clinical implications of the different approaches considered. For instance, if delusions really are beliefs, then cognitive restructuring techniques that depend heavily on argumentation and on the analysis and treatment of information processing biases (e.g., Socratic dialog) could be a sensible therapeutic approach. If, by contrast, delusions are more like imaginings or illusions, not characterized for their reason-responsiveness, other therapeutic approaches focusing on imaginative or perceptual processes could be more effective (e.g., systematic desensitization techniques, perceptual learning-based techniques, etc.).¹³ Likewise, discussions about the relevance of folk psychological concepts for understanding delusions are crucial for determining what therapeutic approaches are needed. If delusions are hybrid states—for example, *bimagnations*—or in-between states that fall short of any given folk-psychological stereotype, more idiosyncratic assessment strategies might be needed. Specifically, this would invite a more careful consideration of the exact behavioral, cognitive, and phenomenological profile displayed by each person, tailoring interventions to each individual case.

Considering the current state of the discussion, doxastic approaches seem to remain the dominant positions for several reasons. Doxastic approaches fit better with much of contemporary empirical research on delusions (Coltheart et al., 2011; 2015; Corlett, 2018; Corlett & Fletcher, 2021; Miyazono & McKay, 2019). In addition, they accommodate well how patients tend to describe their own experiences, which may in turn help to prevent cases of epistemic injustice and other kinds of abuse (Kidd et al., 2022). Finally, they provide plausible responses to the various anti-doxastic arguments in the literature (RC and SCP), yielding more intuitive and conceptually sound frameworks. While the replies offered by revisionist doxasticisms provide better ways of understanding belief ascription practices, Clutton's (2018) non-revisionist doxasticism contributes by emphasizing the relevance of first-person experiences in the conceptualization of delusions. However, despite their strengths, doxastic approaches face several challenges. In the remainder of this section, we point out some of them.

First, it would be beneficial for doxasticists to explore the ways in which first- and third-person approaches to delusion can interact in order to inform their position. On the one hand, an excessive emphasis on third-person descriptions of delusions (as revisionist doxasticisms seem to do) might lead to an over-intellectualization of delusional phenomena (Berrios, 1991), which isolates them from the phenomenological contexts in which they emerge (Mishara & Zaytseva, 2019; López-Silva, 2022). Clutton's (2018) re-emphasizing of the first-person perspective might alleviate this worry; however, it needs to clarify the relationship between the content—or “*what*”—of delusions and the way in which delusions are given in consciousness—or the “*how*” of delusions. Traditionally, doxasticism has tended to focus on the justification and consistency in the ascription of delusional *contents*. However, delusional content might be accidentally true, or not even be the primary source of conflict for patients. In many cases, the content of the delusion might vary over time or could be unclear. As noted by phenomenological psychiatrists, one of the most important aspects of clinical delusions—specially in schizophrenia—is the experiential context in which they are adopted. As suggested by the Early Heidelberg School of authors such as Beringer, Grule, and Mayer-Gross (see Kendler & Mishara, 2019), clinically relevant delusions arise from multimodal perceptual alterations in the general structure of

consciousness where different sensory modalities tend to merge creating a phenomenologically rarified experience of the body, reality, and the self. As we have suggested earlier, claiming that some delusions involve alterations of the structure of phenomenal experience is not incompatible with claiming that delusions are beliefs, taken as endorsements (Sass, 2004, p. 77; see also Bayne & Pacherie, 2004b; see also López-Silva, 2020) or explanations (Sterzer et al., 2016, 2018; Corlett et al. 2020, 2022) of such experiences for example. In addition, clinical evidence shows that psychotic delusions are always accompanied by other symptoms such as auditory-verbal hallucinations and other delusions (Rosen et al., 2016; López-Silva, Harrow, et al., 2022). Arguably, it would be beneficial for doxastic approaches to explore the way in which such altered conditions might inform the phenomenality and ascription of delusions *qua* beliefs.¹⁴

This leads to a second challenge for doxasticism—specially revisionist doxasticism—, concerning the question of whether *folk psychology* really offers an adequate pre-theoretical framework to fully understand delusional experiences and behaviors (Porcher, 2016). The worry here is that, even if we accept that delusions are beliefs, their often-puzzling nature seems to pose a challenge for a full understanding of them using exclusively commonsensical or folk-psychological interpretative resources. This issue resonates with Jaspers' (1963) claim about the existence of a fundamental gap in understanding when it comes to psychotic vs. non-psychotic experiences. This is of utmost importance for ethical issues concerning mental health assessment practices, as this gap in understanding can be the source of important information losses, difficulties for empathy, and derived forms of unjust treatment (e.g., epistemic injustice; see Kidd et al., 2022). This calls for additional interpretative strategies to make sense of each person's lived experience. The use of metaphors, as emphasized by current contextual therapies (Törneke, 2017) or new conceptual resources afforded by phenomenological psychopathology (Kidd et al., 2022) could be a way to bridge this hermeneutical gap between patients, therapists, and other relevant social agents.

Finally, many of the discussed proposals—both pro- and anti-doxasticism—seem to leave the question of what makes delusions pathological somewhat aside (for an attempt to solve this question, see Petrolini, 2015, 2017; Miyazono, 2015). For instance, modest doxasticism explicitly assumes that there is no sharp divide between non-clinical irrational beliefs and delusional beliefs because the presumption of rationality is not constitutive of belief ascriptions (i.e., the *continuity thesis*). This relocates both types of beliefs in a continuum where epistemic features do not establish their main differences. If this is the case, modest doxasticism is required to clarify what is distinctive about delusions and the criteria through which different irrational beliefs can be located in different places of such a continuum.¹⁵ Not drawing clear distinctions between delusions from more common irrational beliefs—such as racist or misogynistic beliefs—might (i) pathologize certain beliefs, exempting their holders from personal responsibility; or (ii) trivialize clinical delusions, detaching them from some of the features that make them the target of clinical research and intervention, and potentially invisibilizing the psychological suffering that often accompanies them. A further exploration of the potential connections between doxasticism and other approaches might lead us too far from our target aim. However, it is important to note that there are plenty of paths for different disciplines and traditions to interact in order to understand the complexities of delusions. We hope to have motivated this task here.

AUTHOR CONTRIBUTIONS

Pablo López-Silva: Conceptualization (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal). **Miguel Núñez de Prado-Gordillo:** Conceptualization (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal). **Victor Fernández-Castro:** Conceptualization (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal).

FUNDING INFORMATION

PLS was supported by the project FONDECYT regular n° 1221058 'The architecture of delusions' granted by the Chilean National Agency for Research and Development (ANID) of the Government of Chile, the Project FACSO 2/2021 granted by the Universidad de Valparaíso, Chile, and the Discretionary International Exchange Award (223452/Z/21/Z) 'Renewing Phenomenological Psychopathology' granted to Matthew Broome and Giovanni Stanghellini. MNPG was supported by the Dutch Research Council (NWO) through a postdoctoral fellowship associated to the VIDI project "Shaping Our Action Space: A Situated Perspective on Self-Control" (VI.Vidi.195.116). He also received partial support from the Spanish Ministry of Science and Innovation through the projects "The Social Roots of Mental Health" (PID2021-126826NA-I00), and "Functional Study of the Clinical Interaction in Patients Diagnosed with Mental Illness" (PSI2016-76551-R), as well as by an FPI-UAM Predoctoral Fellowship (2017), granted by the Autonomous University of Madrid. VFC received support from the BBVA Foundation through the Leonardo 2021 Grant for Researchers and Cultural Creators and by the Spanish Ministry of Science and Innovation through the Juan de la Cierva Research Grant

IJC2019-040199-I and the Project “Las Raíces Sociales de la Salud Mental: agencia y normatividad” PID2021-126826NA-I00.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Pablo López-Silva  <https://orcid.org/0000-0001-7457-7724>

Miguel Núñez de Prado-Gordillo  <https://orcid.org/0000-0001-5680-935X>

Victor Fernández-Castro  <https://orcid.org/0000-0001-7627-5738>

RELATED WIREs ARTICLE

[Are clinical delusions adaptive?](#)

ENDNOTES

- ¹ Some are skeptical or agnostic about the extent to which typological specification is informative for modeling the underlying causal (e.g., neurocognitive) architecture of delusions (e.g., Porcher, 2016); in fact, opinions are divided on this matter. However, major research programs and therapeutic traditions are based on this assumption (e.g., cognitive neuropsychiatry, cognitive behavioral therapy, etc.). The paper mostly focuses on approaches to the typology problem that share this view.
- ² The fundamental rationale behind the doxastic approach has its roots in Locke's notion of madness. In his *Essay on Human Understanding*, Locke (1961; 1689) suggests that madness was due to faulty associations in the process whereby sense data (experiential inputs) were transformed into “ideas” (beliefs) (for a contemporary version of this idea, see Maher, 1974). Porter (2003) claims that Locke's notion of madness as the result of different impairments in the process of formation of ideas became central to the new thinking about mental conditions in Britain and France around 1700 (p. 127). In fact, the term “delusion” was first used as referring to mental problems around the same date.
- ³ Here it is important to note that the subjective certainty of delusional reports varies considerably from subject to subject (Parnas, 2003).
- ⁴ It is worth noting that some dispositionalist positions reject functionalism and its ontological assumptions about the factual and causal nature of beliefs and similar mental states. Instead, they advocate a non-factualist view, according to which beliefs ascriptions are merely interpretative tools that we use in everyday interactions, but they do not capture any particular fact about the person's psychology nor its causal underpinnings (see Heras-Escribano 2017 for a contemporary view of non-factualist dispositionalism and Tanney (2009) for a non-factualist reading of Ryle's dispositionalism).
- ⁵ The ontological commitments of functionalism vary depending on the formulation. For example, Fodor's classical functionalism is committed to psychological realism, that is, the idea that beliefs and other mental states are discrete psychological entities in our cognitive architecture. By contrast, Schwitzgebel' dispositionalism explicitly rejects psychological realism about beliefs, equating these with “dispositional stereotypes” that a certain community associates with believing a certain content; in this sense, his approach is more akin to interpretivist proposals (see Clutton, 2018).
- ⁶ In this sense, it is a common assumption that a defense of doxasticism towards monothematic delusions will provide enough grounds for a defense of a similar account regarding polythematic ones (see Coltheart et al., 2011; López-Silva, 2020).
- ⁷ For some, the content of some delusional statements is itself bizarre enough to preclude an interpretation in literal terms (e.g., “I am dead”). This led Jaspers (1963); Jaspers, 1963 to declare them as empathically “ununderstandable,” in the sense that while they might be explainable in causal terms, they are nonetheless unintelligible from a rational point of view. Although the issue of content is indeed an important one, we'll leave it aside here to focus on the problems of epistemic, procedural, and agential irrationality, in the interpretivist terms, or the problem of the deviation from the causal stereotypical profile, in functionalist terms.
- ⁸ It is true that, for Currie and collaborators, the relevant patterns of action and reaction would be those associated to the second-order belief (e.g., the belief that I believe that my partner has been replaced by an impostor), not the

first-order one (e.g., the belief that my partner has been replaced). However, our point is that the appropriate patterns of actions and reactions for both beliefs partially overlap; for instance, if I believe that I believe that people from all races deserve equal treatment, I am at least expected to present myself as an anti-racist, to try to correct myself if I see signs that I have failed to behave or cognize appropriately, and so forth; all of which are also part of the patterns associated with the first-order anti-racist belief. We would like to thank an anonymous reviewer for helping us clarify this point.

- ⁹ Regarding the content rationality constraint, Bortolotti (2010) takes it to overlap with the epistemic and procedural rationality constraints, hence she doesn't discuss it separately (see pp. 57–58).
- ¹⁰ Note that, although Schwitzgebel (2012) himself is an antidoxasticist, Bayne and Pacherie (2005) account was based on his earlier dispositionalist account (Schwitzgebel, 2002).
- ¹¹ Importantly, Tumulty (2011) takes this argument by Bayne and Pacherie (2005) to motivate anti-doxasticism about delusions; according to her, the motivational and affective deficits that sometimes accompany delusions would prove that certain belief-like dispositions are in fact missing. The belief ascription in this case would not thus be true proper, but just pragmatically convenient at best.
- ¹² Clutton rejects interpretivism and *dispositionalist* functionalism. However, given that Clutton's theory of belief entails the view that beliefs are dispositions to entertain occurrent phenomenal states “before the mind's eye,” and that these have causal roles, Clutton's proposal can be construed as a particular kind of functionalist approach; specifically, as a hybrid kind of functionalism, which exhibits features of both occurrentism and dispositionalism (see Nottelmann, 2013).
- ¹³ We would like to thank an anonymous reviewer for pointing this out.
- ¹⁴ Here we see an especially fruitful path for the mutual enlightenment between the analytic traditional in philosophy of psychiatry and the phenomenological approach to psychopathology.
- ¹⁵ For a potential answer to this issue, see Bortolotti (2020).

FURTHER READING

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5[®])*. American Psychiatric Pub.
- Bortolotti, L., & Miyazono, K. (2015). Recent work on the nature and development of delusions. *Philosophy Compass*, 10(9), 636–645.
- Connors, M. H., & Halligan, P. W. (2020). Delusions and theories of belief. *Consciousness and Cognition*, 81, 1–14.
- Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., & Powers, A. R., III. (2019). Hallucinations and strong priors. *Trends in Cognitive Sciences*, 23(2), 114–127.
- Davidson, D. (2001). *Subjective, intersubjective, objective*. Oxford University Press.
- Gerrans, P. (2014). *The measure of madness*. MIT Press.
- Goldman, A. I. (2012). Theory of mind. In E. Margolis, R. Samuels, & S. P. Stich (Eds.), *Oxford handbook of philosophy of cognitive science*. Oxford University Press.
- Hamilton, A. (2006). Against the belief model of delusion. In M. C. Chung, K. W. M. Fulford, & G. Graham (Eds.), *Reconceiving schizophrenia* (pp. 217–234). Oxford University Press.
- Kusters, W. (2020). *A philosophy of madness: The experience of psychotic thinking*. MIT Press.
- López-Silva, P., & Cavieres, A. (2021). Salto a conclusiones y formación de delirios en psicosis: Un análisis crítico. *Psiquiatría Biológica*, 28(1), 9–13. <https://doi.org/10.1016/j.psiq.2020.12.002>
- López-Silva, P., de Núñez Prado-Gordillo, M., & Fernández Castro, V. (2022). On doxasticisms and anti-doxasticisms: Mapping the typology problem of delusions. *Resistances. Journal of the Philosophy of History*, 3(6), e21095. doi:10.46652/resistances.v3i6.95
- Mölder, B. (2010). *Mind ascribed: An elaboration and defence of interpretivism*. John Benjamins Publishing.
- Morton, A. (2009). Folk psychology. In B. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The oxford handbook of philosophy of mind*. Oxford University Press.
- Sanchez-Curry, D. S. (2020). Interpretivism and norms. *Philosophical studies*, 177(4), 905–930. <https://doi.org/10.1007/s11098-018-1212-6>
- Schneider, K. (1959). *Clinical psychopathology*. Grune & Stratton.
- Schwitzgebel, E. (2021). The pragmatic metaphysics of belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.
- Slors, M. (2015). Interpretivism and the meaning of mental state ascriptions. *Studia Philosophica Estonica*, 10(2), 18–27.
- Stephens, G. L., & Graham, G. (2006). The delusional stance. In M. C. Chung, K. W. M. Fulford, & G. Graham (Eds.), *Reconceiving schizophrenia* (pp. 193–216). Oxford University Press.
- von Eckardt, B. (1994). Folk psychology. In S. Guttenplan (Ed.), *A companion to the philosophy of mind* (pp. 300–307). Blackwell.

REFERENCES

- Alford, B., & Beck, A. (1994). Cognitive therapy of delusional beliefs. *Behaviour Research Therapy*, 32(3), 369–380.
- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders (DSM-5-TR®)*. American Psychiatric Pub.
- Bayne, T. (2010). Delusions as doxastic states: Contexts, compartments, and commitments. *Philosophy, Psychiatry, & Psychology*, 17(4), 329–336.
- Bayne, T., & Pacherie, E. (2004a). Bottom-up or top-down? Campbell's rationalist account of monothematic delusions. *Philosophy, Psychiatry, & Psychology*, 11(1), 1–11. <https://doi.org/10.1353/ppp.2004.0033>
- Bayne, T., & Pacherie, E. (2004b). Experience, belief, and the interpretive fold. *Philosophy, Psychiatry, & Psychology*, 11(1), 81–86.
- Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language*, 20(2), 163–188. <https://doi.org/10.1111/j.0268-1064.2005.00281.x>
- Bayne, T., & Hattiangadi, A. (2013). Belief and its bedfellows. In N. Nottelman (Ed.), *New essays on belief* (pp. 124–144).
- Berrios, G. E. (1991). Delusions as “wrong beliefs”: A conceptual history. *The British Journal of Psychiatry*, 159(14), 6–13.
- Berrios, G. E., & Luque, R. (1995). Cotard's syndrome: Analysis of 100 cases. *Acta Psychiatrica Scandinavica*, 91(3), 185–188.
- Billon, A. (2023). What is it like to lack mineness? In M. García-Carpintero & M. Guillot (Eds.), *Self-experience: Essays on inner awareness* (pp. 314–340). Oxford University Press.
- Bisiach, E., & Geminiani, G. (1991). Anosognosia related to hemiplegia and hemianopia. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury: Clinical and theoretical issues* (pp. 17–39). Oxford University Press.
- Block, N. J., & Fodor, J. A. (1972). What psychological states are not. *Philosophical Review*, 81(2), 159–181. <https://doi.org/10.2307/2183991>
- Blount, G. (1986). *Dangerousness of patients with Capgras syndrome*. Nebraska Medical.
- Bongiorno, F. (2022). Spinozan doxasticism about delusions. *Pacific Philosophical Quarterly*, 103, 720–752. <https://doi.org/10.1111/papq.12373>
- Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford University Press.
- Bortolotti, L. (2011). Double bookkeeping in delusions: Explaining the gap between saying and doing. In J. H. Aguilar, A. A. Buckareff, & K. Frankish (Eds.), *New Waves in Philosophy of Action*. Palgrave Macmillan. https://doi.org/10.1057/9780230304253_12
- Bortolotti, L. (2012). In defence of modest doxasticism about delusions. *Neuroethics*, 5(1), 39–53. <https://doi.org/10.1007/s12152-011-9122-8>
- Bortolotti, L. (2018). *Delusions in context*. Springer Nature.
- Bortolotti, L., & Broome, M. (2009). Delusional beliefs and reason giving. *Philosophical Psychology*, 21(3), 1–21.
- Bortolotti, L. (2020). *The epistemic innocence of irrational beliefs*. Oxford University Press.
- Byrne, A. (1998). Interpretivism. In *European Review of Philosophy* (Vol. 3, pp. 199–223).
- Campbell, J. (1999). Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist*, 82(4), 609–625.
- Campbell, J. (2001). Rationality, meaning, and the analysis of delusion. *Philosophy, Psychiatry, & Psychology*, 8(2), 89–100.
- Carruthers, P. (2013). On knowing your own beliefs: A representationalist account. In N. Nottelman (Ed.), *New essays on belief* (pp. 145–165). Palgrave Macmillan.
- Clutton, P. (2018). A new defence of doxasticism about delusions: The cognitive phenomenological defence. *Mind & Language*, 33(2), 198–217. <https://doi.org/10.1111/mila.12164>
- Coltheart, M. (2002). Cognitive neuropsychology. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology (Vol. 4): Methodology in experimental psychology* (3rd ed., pp. 139–174). Wiley.
- Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology*, 60(8), 1041–1062. <https://doi.org/10.1080/17470210701338071>
- Coltheart, M. (2015). Delusions. In R. Scott & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences* (pp. 1–12). John Wiley and Sons.
- Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, 62, 271–298. <https://doi.org/10.1146/annurev.psych.121208.131622>
- Corlett, P. (2018). Delusions and prediction error. In L. Bortolotti (Ed.), *Delusions in Context* (pp. 35–66). Springer International Publishing.
- Corlett, P. R., & Fletcher, P. C. (2021). Modelling delusions as temporally-evolving beliefs. *Cognitive Neuropsychiatry*, 26(4), 231–241. <https://doi.org/10.1080/13546805.2021.1938984>
- Corlett, P. R., Mohanty, A., & MacDonald, A. W. I. I. (2020). What we think about when we think about predictive processing. *Journal of Abnormal Psychology*, 129(6), 529–533. <https://doi.org/10.1037/abn0000632>
- Corlett, P. R., Mollick, J. A., & Kober, H. (2022). Meta-analysis of human prediction error for incentives, perception, cognition, and action. *Neuropsychopharmacol.*, 47, 1339–1349. <https://doi.org/10.1038/s41386-021-01264-3>
- Currie, G. (2000). Imagination, delusion and hallucinations. *Mind & Language*, 15(1), 168–183. <https://doi.org/10.1111/1468-0017.00128>
- Currie, G., & Jureidini, J. (2001). Delusion, rationality, empathy: Commentary on Martin Davies et al.: *Philosophy, Psychiatry & Psychology*, 8(2), 159–162. <https://doi.org/10.1353/ppp.2001.0006>
- Davidson, D. (1986). A coherence theory of truth and knowledge. In E. Lepore (Ed.), *Truth and interpretation: perspectives on the philosophy of Donald Davidson* (pp. 307–319). Basil Blackwell.
- Davies, M. (2000). *Pathologies of belief*. Blackwell.
- Dennett, D. C. (1979). True believers: The intentional strategy and why it works. In D. C. Dennett (Ed.), *The intentional stance* (pp. 13–42). MIT Press.
- Dennett, D. (1987). *The intentional stance*. MIT Press.

- De Pauw, K. W., & Szulecka, T. K. (1988). Dangerous delusions: Violence and the misidentification syndromes. *The British Journal of Psychiatry*, *152*(1), 91–96.
- Egan, A. (2008). Imagination, delusion, and self-deception. In T. Bayne & J. Fernandez (Eds.), *Delusion and self-deception: Affective and motivational influences on belief formation (macquarie monographs in cognitive science)* (pp. 263–280). Psychology Press.
- Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *The British journal of psychiatry: the journal of mental science*, *157*, 239–248. <https://doi.org/10.1192/bjp.157.2.239>
- Frankish, K. (2009). Delusion: A two-level framework. In M. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience: Philosophical perspectives* (pp. 269–280). Oxford University Press.
- Frankish, K. (2012). Delusions, levels of belief, and non-doxastic acceptances. *Neuroethics*, *5*(1), 23–27.
- Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, *41*, 331–347.
- Frith, C. (1992). *The cognitive neuropsychology of schizophrenia*. Erlbaum.
- Gallagher, S. (2009). Delusional realities. In M. R. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience: Philosophical perspectives*. Oxford University Press.
- Gallagher, S. (2015). Relations between agency and ownership in the case of schizophrenic thought insertion and delusions of control. *Review of Philosophy and Psychology*, *6*, 865–879. <https://doi.org/10.1007/s13164-014-0222-3>
- Garety, P. (1991). Reasoning and delusions. *The British Journal of Psychiatry*, *159*(Suppl 14), 14–18. <https://doi.org/10.1192/S0007125000296426>
- Gibbs, P. (2000). Thought insertion and the inseparability thesis. *Philosophy, Psychiatry, & Psychology*, *7*(3), 195–202.
- Graham, G. (2010). Are the deluded believers? Are philosophers among the deluded? *Philosophy, Psychiatry, & Psychology*, *17*(4), 337–339. <https://doi.org/10.1353/ppp.2010.0033>
- Green, H., Hauser, L., & Troyakov, V. (2018). Are delusions beliefs? A qualitative examination of the doxastic features of delusions. *Psychosis*, *10*(4), 319–328. <https://doi.org/10.1080/17522439.2018.1528298>
- Guillot, M. (2017). I me mine: On a confusion concerning the subjective character of experience. *Review of Philosophy and Psychology*, *8*, 23–53.
- Hamilton, A. (2007). Against the belief model of delusion. In M. C. Chung, K. W. M. Fulford, & G. Graham (Eds.), *Reconceiving schizophrenia* (pp. 217–234). OUP.
- Heras-Escribano, M. (2017). Non-factualist dispositionalism. *Philosophia*, *45*(2), 607–629.
- Hohwy, J., & Rajan, V. (2012). Delusions as forensically disturbing perceptual inferences. *Neuroethics*, *5*(1), 5–11. <https://doi.org/10.1007/s12152-011-9124-6>
- Hohwy, J. (2013). Delusions, illusions and inference under uncertainty. *Mind and Language*, *28*(1), 57–71.
- Humpston, C. S. (2022). Isolated by oneself: Ontologically impossible experiences in Schizophrenia. *Philosophy, Psychiatry, & Psychology*, *29*(1), 5–15. <https://doi.org/10.1353/ppp.2022.0001>
- Jaspers, K. (1963). *General psychopathology*. Manchester University Press.
- Kendler, K. S., & Mishara, A. (2019). e prehistory of Schneider's rst-rank symptoms: Texts from 1810 to 1932. *Schizophrenia Bulletin*, *45*(5), 971–990.
- Kidd, I. J., Spencer, L., & Carel, H. (2022). Epistemic injustice in psychiatric research and practice. *Philosophical Psychology*, 1–29. <https://doi.org/10.1080/09515089.2022.2156333>
- Kriegel, U. (2015). *The varieties of consciousness*. Oxford University Press.
- Locke, J. (1961). *An essay concerning human understanding*. J.M. Dent & Sons Ltd. Original work from 1689.
- López-Silva, P. (2016). The typology problem and the doxastic approach to delusions. *Unisinos Journal of Philosophy*, *17*(2), 202–211. <https://doi.org/10.4013/fsu.2016.172.15>
- López-Silva, P. (2018). Mapping the psychotic mind: A review on thought insertion. *Psychiatric Quarterly*, *89*(1), 957–968. <https://doi.org/10.1007/s11126-018-9593-4>
- López-Silva, P. (2020). Atribuciones de agencia mental y el desafío desde la psicopatología. *Kriterion*, *61*(147), 1–19. <https://doi.org/10.1590/0100-512X2020n14713pls>
- López-Silva, P., Harrow, M., Jobe, T. H., Tufano, M., Harrow, H., & Rosen, C. (2022). 'Are these my thoughts?': A 20-year prospective study of thought insertion, thought withdrawal and thought broadcasting and their relationship to auditory verbal hallucinations. *Schizophrenia Research*.
- López-Silva, P., & Cavieres, A. (2022). El delirio psicótico como objeto de estudio multidisciplinar para la filosofía de la mente. *Sophia*, *33*, 71–90.
- López-Silva, P. (2022). La marca de la Psicosis: Hacia una Breve Síntesis del Problema Tipológico de los Delirios. *Revista Colombiana de Psiquiatría*. <https://doi.org/10.1016/j.rcp.2021.11.002>
- Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual psychology*, *30*(1), 98.
- Maher, B. (2005). Delusional thinking and cognitive disorder. *Integrative Physiological & Behavioral Science*, *40*(3), 136–146. (Original work published in 1974). <https://doi.org/10.1007/BF03159710>
- McKay, R., & Dennett, D. (2009). The evolution of misbelief. *Behavioural and Brain Sciences*, *32*(6), 493–561.
- Mishara, A. L., & Zaytseva, Y. (2019). Hallucinations and phenomenal consciousness. In G. Stanghellini, A. Raballo, M. R. Broome, A. V. Fernandez, P. Fusar-Poli, & R. Rosfort (Eds.), *Oxford handbook of phenomenological psychopathology*. Oxford University Press.
- Miyazono, K. (2019). *Delusions and beliefs: A philosophical inquiry*. Routledge.
- Miyazono, K., & Bortolotti, L. (2014). The causal role argument against doxasticism about delusions. *Avant: Trends in interdisciplinary studies*, *3*, 30–50.

- Miyazono, K., & McKay, R. (2019). Explaining delusional beliefs: A hybrid model. *Cognitive Neuropsychiatry*, 24, 335–346. <https://doi.org/10.1080/13546805.2019.1664443>
- Miyazono, K. (2015). Delusions as harmful malfunctioning beliefs. *Consciousness and Cognition*, 33, 561–573.
- Mullins, S., & Spence, S. (2003). Re-examining thought insertion. *British Journal of Psychiatry*, 182, 293–298.
- Murphy, D. (2012). The folk epistemology of delusions. *Neuroethics*, 5(1), 19–22. <https://doi.org/10.1007/s12152-011-9125-5>
- Nottelmann, N. (Ed.). (2013). *New essays on belief: Constitution, content and structure*. Springer.
- Núñez de Prado-Gordillo, M. (2022). *Mental health without mirrors: A non-descriptivist approach to mental health and the intervention with people with delusions* (Doctoral dissertation. Autonomous University of Madrid). Autonomous University of Madrid.
- O'Brien, L., & Soteriou, M. (2009). *Mental actions*. Oxford University Press.
- Parnas, J. (2003). Self and schizophrenia: A phenomenological perspective. In T. Kircher & A. David (Eds.), *The self in neuroscience and psychiatry* (pp. 217–241). Cambridge University Press.
- Pérez-Navarro, E., Fernández-Castro, V., González de Prado-Salas, J., & Heras-Escribano, M. (2019). Not expressivist enough: Normative disagreement about belief attribution. *Res Philosophica*, 96(4), 409–430. <https://doi.org/10.11612/resphil.1794>
- Petersen, R., & Stillman, R. (1978). *Phencyclidine abuse: An appraisal*. National Institute of Drug Abuse.
- Petrolini, V. (2015). When emotion and cognition do (not) work together: Delusions as emotional and executive dysfunctions. *Behavioral and Brain Sciences*, 38, E84. <https://doi.org/10.1017/S0140525X14000995>
- Petrolini, V. (2017). What makes delusions pathological? *Philosophical Psychology*, 30(4), 498–519. <https://doi.org/10.1080/09515089.2017.1288899>
- Porcher, J. E. (2016). Delusion as a folk psychological kind. *Filosofia Unisinos*, 17(2), 212–226.
- Porter, R. (2003). *Madness: A brief history*. OUP.
- Putnam, H. (1975). The nature of mental states. In H. Putnam (Ed.), *Mind, language and reality: Philosophical papers*. Volume 2 (pp. 429–440). Cambridge University Press (Original work published 1967).
- Radden, J. (2011). *On delusions*. Routledge.
- Reimer, M. (2010). Only a philosopher or a madman: Impractical delusions in philosophy and psychiatry. *Philosophy, Psychiatry, & Psychology*, 17(4), 315–328.
- Rosen, C., Harrow, M., Humpston, C. S., Tong, L., Jobe, T. H., & Harrow, H. (2022). 'An experience of meanings': A 20-year prospective analysis of delusional realities in schizophrenia and affective psychoses. *Frontiers in Psychology*, 1538, 1–14. <https://doi.org/10.3389/fpsyg.2022.940124>
- Rosen, C., Jones, N., Chase, K. A., Gin, H., Grossman, L. S., & Sharma, R. P. (2016). e intra- subjectivity of self, voices and delusions: A phenomenological analysis. *Psychosis*, 8(4), 357–368.
- Rothenfluch, R. (2020). A modified self-knowledge model of thought insertion. *Review of Philosophy and Psychology*, 11, 157–181. <https://doi.org/10.1007/s13164-019-00449-6>
- Sass, L. A. (1994). *The paradoxes of delusion: Wittgenstein, Schreber and the schizophrenic mind*. Cornell University Press.
- Sass, L. A. (2004). Some reflections on the (Analytic) philosophical approach to delusion. *Philosophy, Psychiatry, & Psychology*, 11(1), 71–80. <https://doi.org/10.1353/ppp.2004.0047>
- Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Noûs*, 36(2), 249–275. <https://doi.org/10.1111/1468-0068.00370>
- Schwitzgebel, E. (2012). Mad belief? *Neuroethics*, 5(1), 13–17. <https://doi.org/10.1007/s12152-011-9127-3>
- Schwitzgebel, E. (2013). A dispositional approach to attitudes: Thinking outside of the belief box. In N. Nottelmann (Ed.), *New essays on belief* (pp. 75–99). Palgrave Macmillan. https://doi.org/10.1057/9781137026521_5
- Stephens, G. L., & Graham, G. (2000). *When self-consciousness breaks: Alien voices and inserted thoughts*. MIT Press.
- Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological psychiatry*, 84(9), 634–643.
- Sterzer, P., Mishara, A. L., Voss, M., & Heinz, A. (2016). Thought insertion as a self-disturbance: an integration of predictive coding and phenomenological approaches. *Frontiers in Human Neuroscience*, 10, 502.
- Strawson, G. (2003). Mental ballistics or the involuntariness of spontaneity. *Proceedings of the Aristotelian Society*, 103, 227–256.
- Tanney, J. (2009). Rethinking Ryle: a critical discussion of the concept of mind. In G. Ryle, J. Tanney, Ed., *The concept of mind* (pp. ix – lvii). Routledge.
- Törneke, N. (2017). *Metaphor in practice: A professional's guide to using the science of language in psychotherapy* (Illustrated edition). ed.). Context Press.
- Tumulty, M. (2011). Delusions and dispositionalism about belief. *Mind & language*, 26(5), 596–628. <https://doi.org/10.1111/j.1468-0017.2011.01432.x>
- Tumulty, M. (2012). Delusions and not-quite-beliefs. *Neuroethics*, 5(1), 29–37. <https://doi.org/10.1007/s12152-011-9126-4>
- Wilkinson, S. (2020). Expressivism about delusion attribution. *European Journal of Analytic Philosophy*, 16(2), 59–77. [10.31820/ejap.16.2.3](https://doi.org/10.31820/ejap.16.2.3)
- Young, A. W. (1999). Delusions. *The Monist*, 82(4), 571–589. <http://www.jstor.org/stable/27903656>
- Young, A. W., & Leafhead, K. M. (1996). Betwixt life and death: Case studies of the Cotard delusion. In P. W. Halligan & J. C. Marshall (Eds.), *Method in madness: Case studies in cognitive neuropsychiatry* (pp. 147–171). Lawrence Erlbaum.

How to cite this article: López-Silva, P., de Prado-Gordillo, M. N., & Fernández-Castro, V. (2024). What are delusions? Examining the typology problem. *WIREs Cognitive Science*, e1674. <https://doi.org/10.1002/wcs.1674>