



Explicating the complexity of self-illness ambiguity

Anna Golova^{1,2,3} · Roy Dings^{4,5,6}

Received: 10 May 2025 / Accepted: 26 August 2025
© The Author(s) 2025

Abstract

Self-illness ambiguity (SIA) has been understood, roughly, as a difficulty in delineating one's self from one's mental illness. In this paper, we explicate some of the previously neglected complexity of SIA, by distinguishing two forms of the phenomenon: (a) identity-related SIA (*'How do I relate to my illness?'*) and (b) agential SIA (*'Is it me or my illness that makes me act/think/feel a certain way?'*). In addition, we differentiate general and particular varieties of these SIA-forms, as well as descriptive and analytic approaches to them. The resulting taxonomy allows (1) clarifying the growing SIA-literature, (2) supporting (self-)understanding in clinical contexts, and (3) drilling into the normatively significant features of SIA, e.g., enabling the better theorising of potential connections between agential SIA and questions of responsibility. Our taxonomy thus strengthens the conceptual foundation for future theories and applications of SIA.

Keywords Self · Mental illness · Ambiguity · Identity · Agency

✉ Anna Golova
anna.golova@philosophy.ox.ac.uk

¹ Faculty of Philosophy, University of Oxford, Oxford, UK

² Uehiro Oxford Institute, University of Oxford, Oxford, UK

³ Merton College, University of Oxford, Oxford, UK

⁴ Faculty of Philosophy, Theology and Religious Studies, Radboud University, Nijmegen, the Netherlands

⁵ Department of Psychiatry, University Medical Center Utrecht, Utrecht, the Netherlands

⁶ Lemon Tree Interdisciplinary Center for Psychiatry, Psychotherapy and Philosophy, Amsterdam University Medical Center, Amsterdam, the Netherlands

1 Introduction

[One] woman who was undecided about returning to work as a preschool teacher wondered out loud: ‘Do I not want to go back to it, or is it because of the illness that I don’t want to return?’ (Karp, 1992, 155)

This testimony illustrates one possible manifestation of a phenomenon that has been called ‘illness ambiguity’ (Karp, 1992), or more precisely, ‘*self*-illness ambiguity’ (Sadler, 2003; 2007). Self-illness ambiguity (*hereinafter*: SIA) has generally been understood in the literature to be a “difficulty in distinguishing one’s self, or ‘who one is,’ from a mental disorder or diagnosis” (Dings & Glas, 2020, 334). There has recently been a surge in philosophical attention paid to SIA (e.g., Dings, 2020; Dings & Glas, 2020; Glas & Dings, 2020; de Haan, 2020; Dings & de Bruin, 2022; Bluhm & Cabrera, 2022; Bortolan, 2022; Maiese, 2022; Tekin, 2022; Jeppsson, 2022b; Carls-Diamante, 2023; Drożdżowicz, 2023; McConnell & Golova, 2023; Snoek, 2023; Ervas, 2024; Russell, 2025; Crippen, 2025; Dings et al., 2025; Dings & Golova, 2025). Among this, have been some justified calls for further clarification of the term (de Haan, 2023; Schechtman, 2022; Dings & De Bruin, 2023).

Indeed, a closer look at the fast-growing SIA-literature reveals several sources for tensions, equivocations, and confusions. For instance, as we shall discuss, previous work has been insufficiently precise regarding (1) the heterogeneity of the phenomena discussed under the heading of SIA, and (2) the link between descriptive work on SIA and theoretical commitments concerning the conceptual relation of selfhood and psychopathology.

Addressing the literature’s imprecisions and tensions, the present paper seeks to clarify this conceptual landscape by explicating some previously neglected nuance and complexity in SIA. Regarding the heterogeneity of SIA, we argue for a distinction between (a) ‘identity-related’ SIA (*hereinafter*: ‘identity-SIA’, illustrated by the slogan ‘*How do I relate to my illness?*’), and (b) ‘agential’ SIA (*hereinafter*: ‘agency-SIA’, illustrated by ‘*Is it me or my illness that makes me act/think/feel a certain way?*’). Regarding the theoretical commitments involved in SIA, we argue it is necessary to distinguish levels of abstraction (general vs particular) and approaches (description vs analysis) that can be applied to both identity-SIA and agency-SIA.¹

In doing so, the goal of this paper is to provide more precise language and conceptual tools for capturing SIA as a (heterogenous) phenomenon. We do not mean to develop a full-fledged theory of SIA, but rather to offer a taxonomy that will allow asking more nuanced questions about SIA as a concept, and thus serve as a steppingstone for theorising its underlying philosophical commitments in the future (see Sect. 5).

Providing this taxonomy is important for at least three reasons. (1) We aim to help foster constructive scholarly debate about SIA. Here, taxonomic clarity is impor-

¹ We limit the scope of this paper to self-*illness* ambiguity, although some of our distinctions may also apply more widely to ‘self-ambiguity’ (cf. Dings & de Bruin, 2022). We briefly touch on a more deflationary interpretation of SIA in Sect. 3.

tant in its own right, to avoid authors talking past each other and to identify where real disagreements lie. (2) Our taxonomy can have important implications for mental health service users and providers. For service users, the distinctions we introduce have potential hermeneutical benefits (cf. Kidd et al., 2022; Hassall, 2024), e.g., refining their resources for understanding themselves and their actions, and for being understood and appropriately reacted to by others. For clinicians, our distinctions may be helpful in assessing and supporting their clients' attitudes, diagnoses, and treatment decisions (cf. Dings & Golova, 2025; Dings & de Haan, 2022).² Finally, (3) our taxonomy shall help drill into the normatively and philosophically significant features of SIA, e.g., enabling the better theorising of the implications of agency-SIA for questions regarding service users' agency and responsibility.

With these theoretical and practical goals in mind, in the following, we turn to developing our refined taxonomy of SIA-forms and of different ways of engaging with them. We will first explain and motivate our primary distinction between identity-SIA and agency-SIA in Sect. 2, and situate it among existing distinctions (including 'perspectives' and 'modes' of SIA) in Sect. 3. In Sect. 4, we will consider SIA at different degrees of abstraction (general vs particular), which will lead us to distinguishing approaches to SIA (descriptive vs analytic) in Sect. 5. This will reveal the kinds of philosophical (epistemological, metaphysical, and ethical) questions about SIA that our distinctions can help uncover, especially regarding the potential connections between agency-SIA and responsibility. We will close by considering further directions for research in Sect. 6.

2 Agential and identity-related self-illness ambiguity

The example from the start of this paper has been taken from a longer passage in sociologist David Karp's (1992) report on self-help groups for affective disorders such as depression and bipolar disorder:

...One woman whose husband was a manic depressive said that sometimes 'it is unclear whether the illness is talking or whether [he] is talking.' In a different meeting, a woman, who described a history of beginning to take college courses only to drop out, offered the following analysis of her behavior: 'I would always start courses and then drop out. I used to think I was easily bored. Now I don't blame myself or the class. It's not me or the class, it's the illness.' Yet another woman who was undecided about returning to work as a preschool teacher wondered out loud: 'Do I not want to go back to it, or is it because of the illness that I don't want to return?' [...] [T]he continuing ambiguity about which self was responsible for particular behaviours was evidenced by the comment of a regular who said, 'I'm trying to discover my basic personality aside from the

² In Dings and Golova (2025), we discuss the usefulness of SIA, and of some of our conceptual distinctions within it, for psychiatry. For clinical SIA-aspects, see also Dings and Glas (2020), as well as Glas (2023) who develops a primarily clinical typology of SIA.

disease. That's why I'm at these meetings, to discover which is which' (Karp, 1992, 155).

This entire excerpt is frequently used to illustrate SIA in conference presentations and published articles (e.g., Dings & Glas, 2020). While doing so, however, presenters and authors do not usually acknowledge the rich differences among Karp's examples. This inadvertently construes the examples as representing one homogenous SIA-phenomenon. To demonstrate the heterogeneity of SIA, we propose thinking about this central excerpt in a new way. To this end, we shall analyse it, isolating the following vignettes:

The puzzled wife: 'Sometimes it is unclear whether the illness is talking or whether [my husband] is talking.'

The college dropout: 'I would always start [college] courses and then drop out. I used to think I was easily bored. Now I don't blame myself or the class. It's not me or the class, it's the illness.'

The doubtful teacher: 'Do I not want to go back to [work as a preschool teacher], or is it because of the illness that I don't want to return?'

The personality seeker: 'I'm trying to discover my basic personality aside from the disease. That's why I'm at these meetings, to discover which is which.'

Using these examples, we develop a taxonomy of SIA-forms (summarized in Table 1 at the end of Sect. 4). The distinctions we draw here within expressions of SIA by service users are prima facie neutral with respect to possible underlying philosophical commitments (e.g., regarding the metaphysical reality of a separation between self and mental disorder).

With that in mind, our chief distinction between identity-SIA and agency-SIA can, as a first approximation, be represented by the following slogans:

Identity-SIA: 'How do I relate to my illness? Is it part of who I am?'

Agency-SIA: 'Is it me or my illness that makes me behave/think/feel a certain way?'

Readers might immediately wonder how identity-SIA and agency-SIA may bear on each other. We will return to this question later, after first illustrating these SIA-forms with Karp's examples.

Karp's (1992) *personality seeker* expresses what we call *identity-SIA* in their attempt to distinguish their 'basic personality' from 'the disease'. They are trying to determine a line between who they are and their illness. The *doubtful teacher*, on the other hand, utters *agency-SIA* in wondering whether it is genuinely her or rather her illness that makes her not want to do something – in this case, return to work as a preschool teacher. She is attempting to make sense of a motive for action (or lack

thereof), struggling to determine how it relates to her, her real wishes, and/or her illness.

The *college dropout* has settled on one answer to her previous agency-SIA regarding the behaviour of dropping out of classes, now assuming that this behaviour is due to her illness. She thus no longer experiences SIA because she has settled on how she thinks her illness relates to her and to her actions. Still, imagining the *college dropout's* previous ambiguity will provide a useful example of agency-SIA in the following.

As illustrated by these examples, our distinction between identity-SIA and agency-SIA pertains to the relations at stake in different cases of SIA. Identity-SIA concerns a *dyadic* relation, that is, a relation between two relata: the person and their illness. The person wonders how their (whole) identity relates to their mental illness, including whether the illness is part of who they are. By contrast, agency-SIA concerns a *triadic* relation between three relata: the person, their illness, and an action. The person wonders how an action relates (or as in the case of the *college dropout*, how certain repeated actions relate) to them and to their mental illness. These relations are schematically represented in Fig. 1.

We understand ‘action’ in the broadest possible sense here. That is, in employing the term ‘action’, we stress that agency-SIA can apply to overt actions, but we think it can equally apply to thoughts, feelings, desires, inclinations, dispositions, attitudes, choices, judgements etcetera. Here, we roughly follow the ‘characteristics’ included in Schechtman’s (2018, 73) ‘characterization question’ about which “actions, experiences, beliefs, values, desires, character traits, and so on (hereafter abbreviated ‘characteristics’) are to be attributed to a given person”.

In conceptually distinguishing identity-SIA and agency-SIA, we do not mean to imply that identity and agency are divorced from each other. Undoubtedly, actions (and questions such as ‘is this action *mine*?’) can be connected to identity.³ Who a person is, and what it is that makes them act are distinct questions,⁴ but they can interact. Agency-SIA and identity-SIA may likely feed into each other to the extent that, for most people, actions flow from one’s identity, and may conversely, also help construe identity. Especially repeated actions may be taken as evidence for one’s identity. For instance, the more frequently the *college dropout* drops out of classes, the more she may start to think of this behaviour as defining what she is like as a person. Similarly, the *personality seeker* may well probe into their usual behaviours in their attempt to determine their ‘basic personality’.⁵

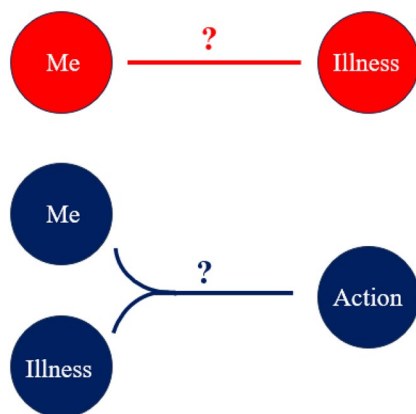
Nonetheless, it is important to conceptually draw apart identity-SIA and agency-SIA, firstly, to capture their potential interactions and to clarify the SIA-literature. Given the close connection between identity and agency, it is perhaps not surprising that authors in the SIA-literature have so far elided the distinction between identity-

³ See again Schechtman’s (2018, 81) characterization question, which concerns e.g., “the degree to which an action expresses someone’s identity”. Sadler’s (2007) heuristic ‘personal self’ also includes both identity and agency.

⁴ See also a Pugh’s (2020b) argument regarding the need to differentiate personality, identity, autonomy, authenticity, agency, and self, which are summarized under the ‘PIAAAS’-acronym in the literature on ‘personality changes’ after deep-brain-stimulation.

⁵ In this vein, Durant (2006 [1926]) took Aristotle to say that ‘we are what we repeatedly do’.

Fig. 1 The elements involved in identity-SIA (top) and agency-SIA (bottom)



SIA and agency-SIA. Notably, Karp (1992, 155) himself inadvertently conflates the two when he introduces the *personality seeker* – who we think expresses identity-SIA – as an example of ambiguity over one’s responsibility for certain behaviours, which seems more typical of agency-SIA as we conceive of it (more on this in Sects. 3 and 5).

Much of the SIA-literature treats examples of (what we would identify as) either identity-SIA or agency-SIA as exhausting the entire phenomenon of SIA, or conversely, inadvertently invokes examples of both forms but treats them as representing one homogenous phenomenon. Indeed, most existing characterizations of SIA, going back to Sadler (2003; 2007), move seamlessly from (1) defining it as a “difficulty patients with mental disorder may have to distinguish their selves (‘who they are’) from their illness” to (2) depicting the characteristic phenomenon of SIA as one of unclarity over “whether a given experience, inclination or behavior belongs to patients themselves or to the disorder” (Glas, 2023, 165; see also e.g., Bortolan, 2022).

Putting aside worries about the separability of self and mental illness (but see Sect. 5), there is an inconsistency within such existing descriptions of SIA. The second kind of depiction of SIA (as wondering whether to ascribe something to oneself or to the illness) may read as though the persons in question have already ‘distinguished themselves from their illness’, thus having resolved SIA in the sense of the first characterization above.⁶

With these observations in mind, we can elaborate on three reasons why it is important to distinguish identity-SIA and agency-SIA: the distinction (1) eases tensions in the SIA-literature, (2) serves as a prerequisite for further theorising SIA, and (3) promises to benefit service users.

⁶Sadler (2007, 116) variously describes SIA as a “confluence of the personal self with mental illness” or as questions like “Is this my illness talking?”. See also Bluhm and Cabrera (2022) for a refinement of Sadler; they, too, observe that “asking [‘is this me or my disorder?'] means that the patient does conceptualize their disorder as distinct from them” (2022, 372). Bluhm and Cabrera’s (2022) as well as de Haan’s (2023) papers constitute exceptions to our criticism of the SIA-literature; we return to these in Sections 3 and 4.

As for the first point, to explicate the outlined tension in the literature in more detail, and to show how our distinction can resolve it, it is helpful to look at how so-called ‘ambiguating factors’ have been described so far, including in some of our own work:

Description 1

Ambiguating factors are those factors with which one may struggle to determine whether [the factor] is internal (i.e. ‘part of the self’) or external (i.e. ‘not part of the self’), such as [...] mental illness (Dings & de Bruin 2022, 274).

Description 2

[T]he patient might wonder whether it is them or X [here: the mental illness] that makes them desire, feel or do Y. The X is the ambiguating factor; the thing that makes it unclear whether it is themselves or X that is making them desire, feel or do Y (Dings & de Haan 2022, 59).

These descriptions are again inconsistent with each other: the conceptualization of the ‘ambiguating factor’ X (here: mental illness) in description 2 appears to presuppose that the individual in question has already determined that X is ‘not part of the self’ in the very sense that is in question in description 1, where it is meant to be unclear whether the ambiguating factor X is part of oneself.

To ease this tension, we need to carefully distinguish between identity-SIA and agency-SIA conceptually. In particular, we suggest that the tension can be resolved if we suppose that description 1 above pertains to identity-SIA, whereas description 2 pertains to agency-SIA. In identity-SIA, a person may wonder whether X, their illness, is ‘part of them’ (as in description 1). However, for the purposes of agency-SIA, that is, when wondering whether it is ‘themselves or X [their illness] that is making them desire, feel or do Y’, a person might presuppose that X, ‘the illness’, is *not* part of themselves (as in description 2).

It may be true that one can only sufficiently answer agency-SIA questions once one has taken a view on whether one’s illness constitutes (part of) one’s identity. However, with our terminology, there is no longer a conceptual tension. It would be problematically question-begging to presuppose that the illness is not part of the self for identity-SIA; yet this assumption is not similarly question-begging for agency-SIA, since agency-SIA *can* occur in the absence of identity-SIA.

Consider, for example, Karp’s *doubtful teacher*, who, apparently assuming that her mental illness is not part of who she is, wonders whether it is her illness that makes her not want to return to work.⁷ The idea that agency-SIA can occur independently of identity-SIA is also supported by the observation that somatic conditions, which are often conceptualized as separate from who one is (cf. Sadler, 2007), can still give

⁷ Conversely, the *personality seeker* could, in principle, experience only identity-SIA without agency-SIA. However, as we will discuss in Sect. 3, in practice, it may often be unreflective agency-SIA that, along with a ‘not feeling like oneself’, can prompt reflective questions about one’s relation to one’s illness.

rise to agency-SIA, as for example, in the case of a diabetic wondering whether their angry outburst was caused by hypoglycaemia. As we will point out below, with mental illness, matters are often still more complicated.

We here mean to merely clarify the SIA-literature's terminology of 'ambiguating factors', rather than to embrace its use going forward. The 'ambiguating factors' terminology harbours potential for further inaccuracies, if only because it captures just part of the factors that can give rise to SIA. For instance, a person's relation to their illness and actions can be ambiguous because of unclarity not just about mental illness but also about the self (cf. Dings & de Bruin, 2022).

What adds to this complexity of self-illness ambiguity in practice is that identity-SIA and agency-SIA can co-occur and overlap as parts of a non-sequential process. A person may reflect on identity-related and agential questions simultaneously or iteratively. That is, they may assume that their illness is separate from who they are when posing a given agential question, while also independently questioning at the identity-related level, whether the illness truly *is* separate from who they are.⁸

So, our *doubtful teacher* may rely on the assumption that her illness is separate from her, when wondering about her lack of motivation to return to work, while, in other instances, asking herself whether that assumption is really accurate. If she sometimes also experiences identity-SIA, or fluctuates over time in how she thinks of her relation to her illness, then this may exert influence at the agential level. Her agential ambiguity may be confounded if she wonders whether to ascribe her action to something that she waveringly conceptualizes as separate from herself.

Yet the very fact that identity-SIA and agency-SIA can affect and bear on each other, suggests that it can be helpful to draw the two apart. We are suggesting that agency-SIA and identity-SIA are distinct, yet interacting, phenomena that have so far been treated as one. We are undertaking a difficult and delicate endeavour here in attempting to neatly tease apart conceptual aspects of something that we suggest is very complex and indeed messy in practice. Yet it is precisely because the terrain is so 'messy' that we need clearer conceptual distinctions to capture important nuances. Recognizing when service users and authors in the SIA-literature invoke identity-SIA and when they rather speak of agency-SIA, allows identifying the different implicit assumptions that can underly these SIA-forms.

This connects to a second benefit of our distinction: as we will elaborate on in Sect. 5, distinguishing identity-SIA and agency-SIA is a steppingstone for further theorising SIA. For instance, given general connections between identity and agency, some might think that one of identity-SIA or agency-SIA will collapse into the other. However, that is a substantial claim that would need to be argued for rather than assumed, and it is only possible to do so once one has first distinguished between the relevant surface phenomena.

By dividing up the terminological terrain, our taxonomy allows asking – but leaves open – several theoretical questions, e.g., regarding the philosophical plausibility of separating self and mental illness and of ascribing actions to one or the other. Distinguishing identity-SIA and agency-SIA similarly leaves open what the exact relation

⁸See Dings and Glas (2020, 339): "The problem in self-illness ambiguity is that this ambiguating factor (i.e., the illness) itself is ambiguous."

between identity and agency will turn out to be in the psychiatric context. In this context, given the ways in which many diagnoses complicate selfhood, we might not be able to just presuppose that actions will flow from identity as straightforwardly as many philosophical theories would assume. Our terminological tools therefore allow scrutinizing this and other assumptions and developing nuanced theories of SIA.

Finally, the distinction between identity-SIA and agency-SIA can also have significant clinical-hermeneutical implications for service users. In the clinic, elucidating agency-SIA and its relation to identity-SIA could help clients achieve greater levels of self-understanding that might enable therapeutic progress. This process shall be aided by our refined terminology, including that of agency-SIA with its focus on actions and symptoms rather than just diagnostic labels. In talking to clients, clinicians can draw on agency-SIA and explicitly point out its interactions with identity-related issues (see our discussion in Dings & Golova, 2025).

For instance, the *college dropout* could be encouraged by a clinician to think about how the way in which she has resolved her agency-SIA may impact her relation to her illness as well as her sense of agency and responsibility for her actions. Or we can imagine the *doubtful teacher* fruitfully reflecting on her relation to her illness, prompted by a clinician's observation that her agency-SIA regarding her career decision may indicate that she has conceptualized her illness as separate from herself. Distinguishing identity-SIA and agency-SIA shows how the two can involve questions that are conceptually distinct from each other, but that can also interact in practice in ways that can contribute to a person's thoroughgoing ambiguity across several levels.

A further *theoretical* benefit of the distinction between identity-SIA and agency-SIA is its deliberate complementarity to and synergy with existing distinctions in the SIA-literature, as well as with another new distinction ('general' and 'particular' SIA) that we will add in Sect. 4. This allows forming a taxonomy of interacting distinctions. In the next section, we discuss interactions with existing distinctions, including 'perspectives' on and 'modes' of SIA.

3 Perspectives, modes, and other existing distinctions in self-illness ambiguity

Our distinction between identity-SIA and agency-SIA firstly sits alongside the existing distinction made in a previous paper between SIA from a first-person perspective (1PP) and SIA from a third-person perspective (3PP) (cf. Dings & de Bruin, 2022). In the literature, SIA is standardly described from the 1PP of service users: I wonder about *my* illness, *my* identity, and/or *my* actions. However, as pointed out in Dings and de Bruin (2022, 283), there can also be 3PP SIA, e.g., "when person A is not sure whether she should attribute a particular behavior or utterance to person B or to B's mental disorder".⁹

⁹SIA from this 3PP may be familiar to many. There could, in principle, also be 2PP SIA. We subsume it under 3PP SIA here, as just another instance of other- rather than self-directed SIA.

In our terminology in the present paper, this is a description of 3PP agency-SIA, as exemplified by the *puzzled wife* in the excerpt from Karp: at times, she wonders whether she is really listening to her husband talk, or whether it is rather ‘his bipolar disorder talking’. In addition to such 3PP agency-SIA, there can also be 3PP identity-SIA: when person A wonders how to distinguish person B from B’s illness. For instance, the *personality seeker’s* family or therapist could also be attempting to distinguish their relative’s or client’s illness from that individual’s personality.

Secondly, in addition to these perspectives, our distinction between identity-SIA and agency-SIA also supplements the existing distinction regarding what can be called ‘modes’ of SIA. In earlier work, Dings and Glas have distinguished a reflective and an unreflective mode in which SIA may be faced: reflective SIA involves conscious reflection about ambiguity, whereas unreflective SIA pertains to an immediately felt, embodied experience of ambiguity (Dings, 2020; Dings & Glas, 2020). For instance, in the unreflective mode, an action may, in an embodied sense, *feel* ambiguous, that is, not quite like one’s own and not quite alien. From this experience, one may shift into the reflective mode, where one consciously reflects on whether to attribute that action to oneself or to one’s illness (which may then again affect embodied experience and so on) (cf. *ibid.*).¹⁰

Dings and Glas (2020) argue that the modes-distinction is important for understanding SIA because SIA may arise when there is a mismatch between one’s reflective mode of self-understanding and one’s unreflective self-experience, that is, between how one understands oneself and how one experiences oneself. To illustrate, we could imagine such a mismatch in the case of the *college dropout* if she reflectively understood herself to be a studious person but unreflectively experienced herself feeling bored whenever she tried to do coursework.

Having established why the modes-distinction can be relevant for understanding SIA, it is important to explain further how these modes interact with our new distinction between identity-SIA and agency-SIA.¹¹ In some ways, the distinctions may appear to run closely together, insofar as identity-SIA seems to most commonly occur at a reflective level, potentially often prompted by unreflective agency-SIA. However, notwithstanding these empirical assumptions, we want to emphasize that there can clearly be reflective agency-SIA and perhaps also unreflective identity-SIA.

Agency-SIA can pertain to both how an action feels when embedded in the situation in which it occurs, as well as to the reflective question of how to attribute that action. The *doubtful teacher* can *feel* ambiguity in her motivations, say, when trying to approach the classroom; *and* she can (as she explicitly does) reflect on her ambiguous motives. From a 3PP, the *puzzled wife*, may also plausibly *feel* ambiguous about her husband’s observable actions or utterances, and she can explicitly reflect on the relation between her husband, his actions, and his bipolar disorder.

The existence of identity-SIA in the unreflective mode is less obvious, since as we mentioned, identity-SIA more commonly manifests as conscious self-reflection when one takes a step back to think about matters of identity. Nonetheless, we want to leave

¹⁰ Although this is an agency-based explanation of SIA-modes, we will explicate below how unreflective and reflective ambiguity may pertain to both agency and identity.

¹¹ We thank an anonymous reviewer for pressing us on this point.

open the possibility of a somewhat ‘deflated’ unreflective identity-SIA, which rather than being an explicit ambiguity about the relationship between oneself and one’s illness, pertains to ‘not feeling like oneself’ in the context of one’s illness (cf. Dings & de Bruin, 2022) (see Table 1 at the end of Sect. 4). The *personality seeker* might, for instance, have such an unreflective sense of not feeling quite ‘themselves’. From a 3PP, the *puzzled wife* could also have an unreflective sense of no longer knowing who her husband is. SIA may generally often start from an indistinct sense of something not feeling ‘quite right’ (cf. Dings & Glas, 2020).

Having indicated how identity-SIA and agency-SIA interact with perspectives and modes, we want to briefly digress here to acknowledge two further existing distinctions by other authors in the SIA-literature. Our distinction between identity-SIA and agency-SIA may remind some readers of distinctions suggested by de Haan (2023) as well as by Bluhm and Cabrera (2022). Here, and in the next section, we suggest that these authors’ distinctions are very valuable but have not yet been systematically developed in a manner that allows for a fine-grained taxonomy of interacting distinctions. As it stands, they risk being conflated with the described modes (as well as with our ‘levels of abstraction’ that we will discuss in the next section).

We briefly present de Haan’s as well as Bluhm and Cabrera’s distinctions here. First, de Haan, in a short commentary piece, indicates a distinction pertaining to “time and scope” (2023, 163, note 2), between (a) “long term” “existential SIA” of “mak[ing] sense of being someone with certain difficulties” (2023, 162), and (b) “acute” “experiential SIA of specific experiences” (2023, 162; abstract). Second, in a paper on ‘self-implant ambiguity’ in deep-brain-stimulation, Bluhm and Cabrera draw an analogy to the disease-illness distinction in the philosophy of medicine, in order to distinguish (a) SIA regarding one’s ‘disease-concept’ (SIA “focused on the disease or disorder itself, understood as an objective entity” or “the way in which an individual perceives the disease and its relationship to the self”), from (b) SIA regarding one’s ‘illness-experience’ (SIA focused on “the subjective experience of that disease” and of one’s mental states) (2022, 372).

Neither de Haan’s temporal distinction, nor Bluhm and Cabrera’s appeal to the disease-illness distinction immediately match our distinction between identity-SIA and agency-SIA. At the end of the next section, we shall be better placed to fully delineate their distinctions from our developed taxonomy. Before doing that, we here want to point out how de Haan’s as well as Bluhm and Cabrera’s distinctions may run the risk of being conflated with the different modes of SIA described above and especially may risk neglecting an important category of reflective agency-SIA.

Describing (what we call) agency-SIA as “experiential SIA of specific experiences” (de Haan, 2023, abstract) or as relating to the “experience of the illness”, “an individual’s experience of their own mental states” or a “phenomenal blurring among mental states” (Bluhm & Cabrera, 2022, 372; 377) may invite an unintended reading that obscures the fine-grained distinction between certain relata of SIA-forms (identity or agency), on the one hand, and the different modes of SIA (unreflective or reflective), on the other. In particular, such experience-related terminology may inadvertently lead to agency-SIA being conflated with or reduced to the unreflec-

tive – i.e., experiential – mode.¹² Yet as we discussed, agency-SIA can and does very much manifest in the reflective mode; and we will suggest further below that explicitly acknowledging (certain forms of) such reflective agency-SIA can help reveal a normative dimension of SIA.

In addition, we have also emphasized in Sect. 2 the importance of thinking about (reflective) agency-SIA in terms of a *triadic* relation between person, illness, and action (broadly conceived). Bluhm and Cabrera *sometimes* – but not consistently – adopt a triadic phrasing when they switch to describing SIA regarding illness-experience as including “the question of which mental states [emotions, beliefs, values, etc.] are caused or influenced by the disorder, and which arise from the individual themselves” (2022, 377). To develop a systematic SIA-taxonomy that distinguishes (reflective) agency-SIA from other SIA-forms, we think it is necessary to *consistently* characterize agency-SIA as triadic.

We also think it is important to point out that the triadic relation characteristically involved in agency-SIA *can* pertain to emotions, beliefs, values etcetera, as mentioned above, but we want to emphasize that agency-SIA can *also* pertain to (motives for) actions in the narrow sense, i.e., overt behaviours. One of the main reasons to emphasize a category of reflective agency-SIA that can pertain to actions in the narrow sense, is that this makes room for evaluating those actions, thus revealing a normative dimension of SIA.

It is often in reflective agency-SIA about actions in the narrow sense (especially from a 3PP, but also a 1PP) that folk-ascriptions of responsibility for actions and questions of appropriate reactions to them may come up (cf. Strawson, 1982) (see also Greenberg, 2024; Sadler, 2003). As Karp (1992, 153) puts it:

The ambiguity that people tried to resolve through discussion was whether they were responsible for their feelings, *behaviours*, and, ultimately, the direction of their lives or whether they were simply pushed around by forces over which they had little control (*emphasis added*).

For example, it is at the level of reflective agency-SIA that the *college dropout* used to wonder whether to blame herself, her classes, or her illness for the behaviour of dropping out of classes. From a 3PP, we can imagine the *college dropout's* professors wondering whether she dropped out of class because of her illness, and subsequently wondering how to appropriately react to this behaviour. Explicating a category of reflective agency-SIA that can pertain to overt actions renders such questions regarding agency less conspicuous and easier to ask clearly. In Sect. 5, we will indicate how these questions may connect to ongoing discussions in philosophy.

First though, we need to add another layer to our taxonomy. We want to point out that there can also be different levels of abstraction applied, e.g., to what Bluhm and Cabrera describe as SIA regarding “the disease or disorder itself” or “the way in which an individual conceptualizes their disorder and its relationship to their self”

¹² Again, we mean that the authors’ distinctions may be read in this way, not that they are intended as such. De Haan (2023, 163, note 2) points out that contrary to Dings and Glas’ modes-distinction, hers is intended as a “distinction of time and scope”, which we pick up on in Sect. 4.

(2022, 372; 377). We will show how reflective identity-SIA and agency-SIA questions can be posed at different levels of abstraction. Explaining what we mean by these levels in the next section will allow us to introduce relevant approaches to SIA and connections to philosophical questions in Sect. 5.

4 Levels of abstraction: general and particular self-illness ambiguity

The further distinction within (especially 3PP) reflective SIA that we wish to introduce here is a distinction between (more) *general* and (more) *particular* reflective questions. The kinds of SIA we have discussed thus far can be described as particular forms, that is, mostly cases of one specific individual wondering about a particular action, identity/self, and/or mental illness (symptom), in themselves or another person. As illustrated by Karp's participants, such particular questions are commonly asked from the 1PP by service users (identity-SIA: *'Where does my personality end and illness start?'* (cf. de Haan, 2020); agency-SIA: *'Is it me or my illness that makes me (not want to) do this?'*), or from the 3PP, e.g., by loved ones (*'Is this my partner's illness talking?'*).

Importantly, particular 3PP-questions can also be asked by clinicians with respect to their clients (identity-SIA: *'How does my client relate to their illness?'*; agency-SIA: *'Is what I am observing a symptom?'*, *'Is this their illness speaking?'* (Greenberg, 2024)). Clinicians may attempt to evaluate which 'aspects' of their clients to count towards their diagnosis (in ways that can diverge from the client's own assessment (cf. Sadler, 2007; see also Dings & Golova, 2025)). Prior to diagnosis, queries by both clinicians and lay people may also include the difficulty of whether to label someone as having a (including a particular) mental illness, say, as experiencing clinical depression rather than low mood, in the first place (cf. Bröer & Besseling, 2017).

It is now useful to see that there can also be very general forms of reflective identity-SIA and agency-SIA, that is, questions not about individual cases, but e.g., at the most abstract level, about the concepts 'agency', 'identity', 'selfhood' and 'mental disorder' or 'psychopathology', and the theoretical relations between them. Rather general questions (also hinted at by Sadler, 2007) can to some extent be asked from the 1PP by service users themselves. For instance, the *personality seeker's* attempt to determine a line between their personality and their illness may include reflection on the more abstract question of how to plausibly think about the complex conceptual relation between selfhood and mental disorder more generally. The unavailability of an uncontested theoretical concept of selfhood or of mental disorder may contribute to a person's ambiguity about *their* relation to *their* mental illness.

General queries play an even more characteristic role in SIA from the 3PP, since it is from this perspective that abstract philosophical or scientific questions are usually formulated. The aforementioned question of judging whether a specific individual has a (particular) mental illness will, for instance, be informed by a more general question, known as the 'demarcation question' in (the philosophy of) psychiatry. This is the "question of what distinguishes [mental] disorder from health" (Kingma, 2013, 364), e.g., where and how to generally draw the definitional line between depression and intense sadness (Horwitz & Wakefield, 2007).

SIA in this general sense can concern more abstract questions about a definitional rather than a merely empirical “elusiveness of the self-illness relation” (Dings & Glas, 2020, 337), and about the plausibility of a conceptual self-illness distinction (cf. Sass, 2007).

Here, it is important to clarify that particularity and generality sit on a spectrum, ranging from the ‘most particular’ to the ‘most general’ forms of SIA. For agency-SIA, at the most particular end, we could imagine the *college dropout* wondering whether one particular instance of her dropping out of a class is due to her affective disorder (for another example see Participant 23 in Hope et al., 2011, 25). From here, one could generalize along two axes: the number of agents and/or the number of actions. Agency-SIA can, for instance, pertain to one individual’s many actions, such as when the *college dropout* wonders whether her *repeated* behaviour of dropping out of classes is due to her illness.¹³ Generalizing over both axes, one could ask questions, say, about all actions by all individuals with such affective disorders, through to asking – at the highest level of abstraction – how agency and psychopathology conceptually relate to each other in general.

Identity-SIA is in a sense already more general than agency-SIA: we described in Sect. 2 that it pertains to a relation between someone’s illness and their identity as a whole, for which especially their repeated actions can in turn provide evidence.¹⁴ Still, identity-SIA admits of more or less general permutations, ranging from questions about one specific individual’s identity in relation to their particular mental illness, to how a number of people with a similar diagnosis relate to their illness, all the way to the quite different abstract question of how psychopathology and selfhood conceptually relate.

At this point, we are in a position to clarify further how our now developed taxonomy differs from the existing distinctions by de Haan as well as Bluhm and Cabrera that we mentioned in Sect. 3. We can now see how their distinctions pertaining to (1) “time and scope” ((a) “long term”/“overarching” vs (b) “acute”/“specific, singular” SIA) (de Haan, 2023, 163, note 2; 162), as well as to (2) ‘disease-concept’ ((a) how “an individual conceptualizes their disorder” vs (b) how they experience their illness/particular mental states) (Bluhm & Cabrera, 2022, 372; 377) may also blend identity-SIA and agency-SIA with SIA at different levels of generality or abstraction.

Contrary to the authors’ distinctions regarding time and scope and disease-concept, the scope of identity-SIA and agency-SIA questions can cover both specific situations and more general considerations, including abstract conceptual ones. Unlike de Haan’s distinction, ours is also not temporal: although identity-SIA may more often be diachronic, we want to allow for ‘long-term’ agency-SIA, e.g., in temporally

¹³ ‘Generality’ here pertains to the content of the question, that is, wondering about repeated actions, not to repeated instances of asking the question itself.

¹⁴ Identity-SIA usually pertains to one identity per person. One could ask about different identity-*aspects*, e.g., personality traits, but triadic questions about the attributability of such traits to oneself or one’s illness may rather fall under agency-SIA. Depending on how one thinks about the relation between identity, personality dispositions, and patterns of action, some might take identity-SIA to just be a more general form of agency-SIA. This is a possible theoretical position that we, as per our discussion in Sect. 2, do not mean to settle with our taxonomy.

extended ‘actions’ like pursuing a career. Similarly, identity-SIA *can* pertain to one’s disease-concept (in relation to one’s self-concept), as Bluhm and Cabrera rightly suggest, but not exclusively so: it can pertain to (i) how a person experiences themselves in the context of their illness; (ii) how they think their illness relates to themselves (or how another’s illness relates to that person); (iii) how they conceptualize what mental disorder and selfhood are, *and*, most abstractly, (iv) how these concepts are construed from the 3PP, e.g., by philosophers. We mean to provide a fuller taxonomy that emphasizes how different distinctions in SIA interact to form and allow asking these more nuanced questions.

Now, it is at the most general, abstract level, pertaining to how the various concepts relevant to SIA fit together, that one can wonder whether a separation between self and mental illness is conceptually coherent on a plausible theoretical understanding of selfhood and of psychopathology. Indeed, de Haan (2023, 160) has raised the worry that the very term ‘self-illness ambiguity’ may invite a particular view of “both illness and self as entities that [...] can be separated”.

This leads us to the next section. In Sect. 2, we drew distinctions within lay expressions of SIA. The nuance we have been adding allows us to now probe into the meaning and coherence of these expressions and to ask what theoretical presuppositions may lie behind these distinctions. For instance, despite our claims to the contrary, some may suggest that our description of reflective agency-SIA as a question of ascribing actions either to oneself or to one’s illness (or to person B or their illness), at least seems to have overtones of the simplified theoretical view of mental disorder and of the self as separable things that de Haan (2023) appears to criticize.

Indeed, although there is no theoretical agreement on how to construe selfhood, mental disorder/psychopathology, or their relation, many nuanced philosophical views of selfhood and mental disorder seem to emphasize the intertwinement of the two (e.g., Tekin, 2021; Gallagher, 2024). Such abstract analyses of the relation between selfhood and mental disorder are not always straightforwardly corresponding to or compatible with individuals’ expressions of SIA. In the next section, we address this apparent discrepancy by distinguishing different approaches to SIA.

Before doing so, Table 1 summarizes the taxonomy of SIA that we have been developing over the course of Sects. 2 to 4. Identity-SIA and agency-SIA can be integrated into a systematic SIA-taxonomy of interacting distinctions, including modes, perspectives, and levels of abstraction. (The ‘general’ and ‘particular’ forms featured in Table 1 – as well as in Table 2 in the next section – denote the maximally general/particular ends of the spectra described above, where questions located at opposite ends need not directly correspond to each other and different kinds of questions are possible between the extremes).

Table 1 Relata, modes, and perspectives of SIA (cf. Dings & Golova, 2025)

Relatum	Mode	Persp.	Illustration
<i>Identity</i>	Reflective	1PP	'How do I relate to my illness?'
		3PP	Particular: 'What is the relation between B and their illness?' General: 'How does selfhood relate to psychopathology?'
	Unreflective	1PP	<i>Feeling</i> ambiguous about oneself in the context of one's illness
		3PP	<i>Feeling</i> ambiguous about who B is in the context of their illness
<i>Agency</i>	Reflective	1PP	'Is it me or my illness that makes me do/think/feel this?'
		3PP	Particular: 'Is it B or their illness that makes them do/think/feel this?' General: 'How does psychopathology relate to agency (and to selfhood)?'
	Unreflective	1PP	<i>Feeling</i> ambiguous about one's action/thought/feeling in the context of one's illness
		3PP	<i>Feeling</i> ambiguous about B's action etc. in the context of their illness

Note: A version of Table 1 also appears in Dings and Golova (2025), in which we focus on the relevance of these distinctions for psychiatry. There, we refer to the distinction between 'general' and 'particular' SIA as a distinction between 'type' and 'token' forms of SIA

5 Descriptive and analytic approaches to self-illness ambiguity

We want to propose that the potential discrepancy between expressions of more particular SIA and abstract conceptualizations of the relation between selfhood and mental disorder can be addressed by distinguishing between descriptive and analytic (and partially prescriptive) projects that one may pursue regarding SIA. We will also suggest that distinguishing agency-SIA from identity-SIA as well as more general from more particular questions about them can aid both these projects. In particular, as we have begun discussing earlier, these distinctions can support the analytic project by enabling one to ask refined conceptual questions about SIA and connecting SIA to important philosophical debates, e.g., regarding responsibility.

One way to think of discussions about SIA is as a purely descriptive project. On this understanding, the aim is to try to capture the phenomenon of SIA as accurately as possible by faithfully describing people's reports of it. Such reports are the only authoritative benchmark for describing how SIA is experienced and reflected upon by service users (and third parties).

Our distinction between identity-SIA and agency-SIA can aid this descriptive project by helping to capture the nuances of people's expressions of SIA. Especially, the agency-SIA slogan '*Is it me or my illness that makes me do/think/feel this?*' is primarily descriptive in this way, intended to reflect utterances that lay people appear to make, as evidenced e.g., by Karp's (1992) study.

Here, if agency-SIA appears to rely on construing self and illness as separable at the identity-level, then this is appropriate to the extent that such assumptions can be part of people's folk-concepts, especially given popular biomedical, essentialist ideas of mental disorder (cf. Karp, 1992; Haslam & Ernst, 2002; Tse & Haslam, 2024; see also Lakoff, 1992). When such conceptions are seemingly perpetuated in the SIA-literature, then this may need to be interpreted as part of a descriptive approach to SIA.

Table 2 Approaches to SIA

	Description	Analysis
Particular	Describing instances of how specific service users (and third parties) voice SIA.	Suggesting the most helpful way for specific service users to think about the relations between their illness, themselves, and their actions.
General	Describing common patterns in service users' (and third parties') expressions of SIA.	Analysing the most coherent conceptualization of the relations between psychopathology, selfhood, and agency. Societal/professional linguistic norms regarding the self-illness relation

It is, however, important to distinguish the description of people's utterances of SIA from a theory thereof; describing how lay people divide up the world is distinct from theorising whether that way of dividing up the world is conceptually coherent. Hence, it is a separate question whether reports of people's SIA (be it in individuals' descriptions or in depictions of how SIA commonly manifests) match, say, what philosophers would consider a coherent conceptualization of the relation between psychopathology, selfhood, and/or agency.

In addition, it is yet another question whether people's reports, or philosophers' analyses, match clinicians' assessments in specific cases, that is, what clinicians would suggest to a given client as a helpful way to think about their relation to their illness and actions. The general-particular distinction then cuts across the distinction between descriptive and analytic approaches to SIA. The potential combinations of approaches are summarized in Table 2.

We shall now further consider the different potential approaches to SIA from Table 2, with an emphasis on the analytic (and partially prescriptive) ones.¹⁵ To begin with descriptive approaches, some researchers (e.g., empirical psychologists) may be interested in capturing cases of SIA in specific individuals, while others (e.g., sociologists) may describe patterns in SIA occurring across society. The expressions of SIA to be described may, in turn, be influenced by (changing) societal norms, e.g., regarding how to talk about the relation between persons and their diagnoses (cf. American Psychiatric Association, 2000).¹⁶

¹⁵ Some of the approaches in the 'analytic' category, may be rather prescriptive. There are norms involved in what philosophers think is the right way (for philosophers) to construe the self-illness relation, in what clinicians recommend to their clients as valuable ways of construing this relation, and of course, in what, say, the media communicate to society as appropriate ways of talking about this relation. We prefer the more neutral umbrella term 'analytic' here to avoid confusion between the term 'prescriptive' and the subcategory of 'normative implications' of (agency-)SIA, in the sense of implications to do with the moral evaluation of actions. Similarly, it is beyond the scope of this paper to discuss the question of who should get to prescribe what in the mental health context.

¹⁶ General norms can also conflict with individual assessments: e.g., despite official recommendations to say 'persons with autism' (American Psychiatric Association, 2000), some in the referred to category state that being called 'autistic persons' better captures their relation to their diagnosis (cf. Kenny et al., 2016; Jorba et al., 2024).

Clinicians (e.g., psychiatrists or clinical psychologists) may engage in description as well as analysis and advice in SIA-cases. As a starting point for treatment, clinicians are likely interested in capturing descriptively how their client thinks about the relation between themselves, their illness, and their actions. Yet clinicians may also assess what ‘aspects’ of their client they, as clinicians, should consider illness symptoms, and when a given utterance or even treatment decision may be ‘the illness talking’ (cf. Tan, 2003). In addition, clinicians (perhaps not unlike the client’s friends and family) may also want to take a view on how it would be most therapeutically valuable or helpful for a given client to think about their relation to their illness and actions (cf. Dings & Golova, 2025).

Such advice will in any case need to be individually tailored: for instance, separating oneself from one’s illness or disowning certain (patterns of) actions, may be helpful for some clients and harmful for others, or have both positive and negative effects on the sense of agency of yet others (cf. Karp, 1992; Tekin, 2014).¹⁷ Whatever the concrete individual recommendation, the terminology of identity-SIA and agency-SIA (and their interactions) is meant to provide a language for clinicians to help their clients in the process of making sense of their identity and agency in the context of their illness (cf. Dings & Golova, 2025).

Finally, zooming out from particular cases, the distinction between identity-SIA and agency-SIA is also meant to support conceptual analysis regarding SIA, of the kind that especially philosophers (but also some psychiatrists or psychologists) may engage in. For conceptual work, abstract questions will normally be more central than specific cases. Philosophers, for one, are usually interested in truths of a more general nature (although they may investigate these by invoking real or hypothetical cases).

This also means that the goal of conceptual analysis regarding SIA will be quite different and separate from that of therapeutic advice in SIA-cases. Practising clinicians aim to support their clients in thinking about the relation to their illness and actions in whichever way is individually most therapeutically helpful. By contrast, for those conducting conceptual analysis, the question of helpfulness may likely be secondary to that of conceptual coherence or accuracy. That is, the most coherent or accurate conceptualization of psychopathology, selfhood, agency, and their theoretical relation may be independent of and differ from both lay concepts and from therapeutically helpful ideas regarding the elements involved in expressions of SIA.

The right way to think about these elements conceptually is not settled by distinguishing identity-SIA from agency-SIA or analytic from descriptive general or particular approaches to them. However, these distinctions allow theorists to formulate and investigate more nuanced questions in order to develop theories about potential conceptual presuppositions behind these SIA-forms.

The distinctions also make room for evaluating the normative significance of asking these refined questions and for identifying potential routes to scepticism about

¹⁷For some clients, externalizing the illness and certain behaviours may support viewing their difficulties not as identity-constituting (Hope et al., 2011; Pedley et al., 2017) but as manageable (Rego, 2004; Tomm, 2019). For others, externalization may harm their belief in being able to control and change their behaviours (Pickard, 2015; Karp, 1992; Sass, 2007). For yet others, it may be helpful to embrace ambiguity during recovery (McConnell & Golova, 2023; Drożdżowicz, 2023).

SIA. Some might, for instance, hold the sceptical view that the analytic level will ultimately collapse into the descriptive (or as we discussed in Sect. 2, that agency-SIA will collapse into identity-SIA), but again, our terminological distinctions are needed in order to be able to argue about such theoretical claims in the first place.

We have already pointed out some further sceptical abstract questions about SIA in Sect. 4. With regard to identity-SIA, for instance, one source of scepticism at the level of conceptual analysis concerns the separability of selfhood and psychopathology. Independently of whether assuming a self-illness distinction is sometimes therapeutically valuable, philosophers may examine whether such a separation is coherent on plausible views of selfhood and psychopathology. On multidimensional views of both concepts (Gallagher, 2024), for instance, self-aspects will conceivably overlap with illness-aspects (cf. Dings & de Bruin, 2022; see also Jeppsson 2022b; Tekin, 2022; Schechtman, 2022).

For agency-SIA, one may similarly examine whether it must presuppose any implausible conceptualization of mental disorder as separable, objectified, anthropomorphized, or agency-undermining. At the abstract, analytic level, agency-SIA can indeed be a concern for several branches of philosophy, including epistemology and metaphysics, and as we will discuss further below, ethics.

For instance, the *doubtful teacher's* agential question '*Is it because of the illness that I don't want to return?*' could be interpreted as an epistemic question aimed at determining (perhaps with the help of a therapist or philosophers) an extant truth about her real motive for (in)action. This appears to be a widespread interpretation in the SIA-literature we critiqued in Sect. 2, but it is not the only possible view.¹⁸ Theorists with (metaphysical) worries regarding the idea that there is such a truth to be discovered (cf. Jeppsson 2022b), might develop alternative interpretations of what it is that is being communicated in agency-SIA questions.¹⁹

Perhaps most significantly, as we began pointing out in Sect. 3, abstract concerns about agency-SIA, including the question whether a mental disorder could really 'make' someone act a certain way, can relate to issues in ethics. Demarcating agency-SIA from identity-SIA helps isolate the agential aspects of SIA that are especially relevant for ethical assessments. This shall make it easier in the future to explore important connections between SIA and bodies of literature in ethics, particularly on autonomy and responsibility.

With regard to autonomy, for instance, clinicians' 3PP worries that a client's treatment decision may be 'the illness talking',²⁰ can be usefully connected to the literature on decision-making capacity (cf. Appelbaum & Grisso, 1988; Pugh 2020a) and to theoretical questions regarding how mental disorders (as well as ambiguity) can

¹⁸ See also McConnell and Golova (2023, 67f) on the issue of conflating the idea that factors can be genuinely ambiguous and the epistemic issue of knowing how ambiguous they are.

¹⁹ Epistemic and metaphysical questions about agency-SIA also include issues of causation, e.g., the extent to which behaviours can be causally explained by reference to descriptive diagnostic criteria (that may include the very behaviours in question (cf. Giordano, 2005)).

²⁰ For instance, Tan (2003, 1246) points out how some treatment refusals by clients with anorexia nervosa can prompt psychiatrists to ask: "Are such refusals the patient's real wishes, or are they 'the anorexia talking'?"

plausibly be said to affect the ability to make autonomous (treatment) decisions (cf. Dings & de Haan, 2022).

Agency-SIA, especially regarding overt actions, also offers potential connection points to central discussions on responsibility. We mentioned earlier that e.g., the *college dropout's* agency-SIA case can include the question of what kinds of reactions her behaviour warrants, both from a 1PP (should she feel guilty/blame herself for the behaviour of dropping out of classes?), and from a 3PP (should her professors feel angry about this behaviour and/or blame her for it?). One way of capturing this philosophically could be in terms of P. F. Strawson's 'reactive attitudes', from a 1PP ("self-reactive attitudes", including guilt, remorse, and shame) and from a 3PP (e.g., "gratitude, resentment, forgiveness, love, and hurt feelings") (Strawson, 1982, 75).²¹

Beyond such reactive attitudes, the implications of agency-SIA for responsibility could also be usefully analysed further in terms of theories on which responsibility requires control over one's behaviour (cf. Nagel, 1979; Fischer & Ravizza, 1998; see again Karp, 1992), or expression of one's 'real self' (cf. Shoemaker, 2015; Frankfurt, 1988). As Schechtman (2018, 78) points out, whatever one's particular view of responsibility, a "person can be held responsible only for *his own* actions". Notably, it is an open question what exactly the connection between agency-SIA, responsibility, and blame-related reactive attitudes will turn out to be, since some accounts allow divorcing the therapeutically helpful presumption of responsibility from illness-ascriptions and from others' affective blame-reactions (see Pickard, 2017).

More generally, there is, of course, a vast body of literature with differing views on agency, responsibility, and blame in the context of mental disorder (cf. King & May, 2022; Kennett, 2009; Shoemaker, 2015; Wilkinson, 2022; Pickard, 2017 to name but a few); we only mean to indicate here that agency-SIA offers a specific new angle and terminology for connecting ambiguities in the relation between self, actions, and disorder to important normative questions. Queries about the extent of control one has over one's actions, the degree to which these express who one really is, and whether one can or should be held responsible and/or be blamed for them, appear to be at the heart of many expressions of SIA. Zooming in on agency-SIA promises helping to theorise these dimensions of SIA in the future.

6 Further directions and conclusion

This paper has differentiated identity-related self-illness ambiguity (identity-SIA) from agential self-illness ambiguity (agency-SIA), in their more general and more particular forms, allowing for a plurality of descriptive and analytic approaches that can to some extent be considered independently of each other. Our taxonomy sug-

²¹The distinctions between agency-SIA and identity-SIA, as well as particular and general SIA, here can also help separate evaluation of actions from evaluation of the agent. Service users express self-reactive attitudes in agency-SIA and should be considered subjects of others' reactive attitudes, despite P. F. Strawson's (1982) indication that the 'objective attitude' should be adopted towards them. Jeppsson (2022a) reads Strawson in this way, criticizing him for depicting service users as generally exempt from responsible agency rather than excused for particular actions. Though not invoking SIA in *that* paper, Jeppsson (2022a) discusses 'unstable' responsibility (self-)exemptions, using Karp's *college dropout* case.

gests several routes for future empirical and philosophical research, especially into agency-SIA:

(1) There may be a need for systematic empirical research that focuses explicitly on agency-SIA, rather than including it incidentally in broader research on illness-ambiguity (Karp, 1992) or illness-identity (Estroff, 1989; Yanos et al., 2010). (2) Theoretical presuppositions behind agency-SIA in relation to identity-SIA may usefully be interrogated and examined conceptually and as possible discussion points in psychoeducation and clinical encounters. (3) Further investigating agency-SIA can be especially important in normative contexts, including assessments of service users' (sense of) agency and responsibility (and its relation to blame), as well as the capacity to make autonomous (treatment) decisions.²² (4) Philosophers may further pursue sceptical questions regarding the distinction between selfhood and mental disorder. One question here is also what, if any, role philosophical analysis should play in influencing or revising lay discourse. (5) Theorists sceptical of the self-illness distinction may investigate alternative (including less literal) interpretations, e.g., of agency-SIA as attempts at negotiating one's agency and responsibility for certain actions (cf. Ervas, 2024; cf. Jeppsson 2022b). (6) The extent to which agency-SIA (and identity-SIA) may look different for different diagnoses could be investigated (e.g., comparing episodic to continuous and short-term to chronic conditions).²³ (7) Finally, one may investigate how agency-SIA might manifest differently for different kinds of actions and mental states.

By providing a taxonomy of SIA, we aim to have laid the foundation for exploring these and other conceptual, clinical, and empirical questions in the future. The distinctions we have introduced shall support careful discussion and further theorising about the various forms of SIA. We also hope to prompt reflection on the role of philosophers in adding theoretical nuance, while remaining attentive to the complexity of the heterogenous phenomenon of SIA itself.

Acknowledgements We wish to thank two anonymous reviewers of this journal for their careful and constructive feedback. We also thank Prof Edward Harcourt as well as Dr Evert Merckx and Helena Ward for their very valuable comments on the manuscript; we owe special thanks to Dr Jonathan Pugh and Paul Heller for particularly extensive comments that helped shape this paper. Finally, we thank Prof Marya Schechtman and the participants of the workshop 'Narrative and selves: From memory to AI – Themes from the work of Marya Schechtman' for helpful discussion of ideas pertaining to this paper.

Author contributions AG and RD both conceived of the central idea for this paper. AG wrote the manuscript and revised it, in close consultation with RD.

Funding This work was supported by the Arts and Humanities Research Council through the Open-Oxford-Cambridge DTP (AG, ref. AH/R012709/1, studentship no. 2744275) and by the Dutch Research Council (NWO) (RD, project number 406.23.FHR.006).

Data availability Not applicable.

²² Agency-SIA may also contribute to adjacent research on 'responsibility gaps', where it may be ambiguous whether to ascribe responsibility to a person or e.g., their brain implant (Matthias, 2004; Bluhm & Cabrera, 2022).

²³ See Dings et al. (2025) for an application of some of our distinctions to personality disorders.

Declarations

Competing interests The authors have no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, Fourth edition, text revision (DSM-IV-TR)* (4th ed., Vol. 1). American Psychiatric Association.
- Appelbaum, P. S., & Grisso, T. (1988). Assessing patients' capacities to consent to treatment. *New England Journal of Medicine*, 319(25), 1635–1638. <https://doi.org/10.1056/NEJM198812223192504>
- Bluhm, R., & Cabrera, L. Y. (2022). Self-implant ambiguity? Understanding self-related changes in deep brain stimulation. *Philosophical Explorations*, 25(3), 367–385. <https://doi.org/10.1080/13869795.2022.2065342>
- Bortolan, A. (2022). Selves Hijacked: Affects and personhood in 'self-illness ambiguity'. *Philosophical Explorations*, 25(3), 343–362. <https://doi.org/10.1080/13869795.2022.2093393>
- Bröer, C., & Besseling, B. (2017). Sadness or depression: Making sense of low mood and the medicalization of everyday life. *Social Science and Medicine*, 183(June), 28–36. <https://doi.org/10.1016/j.socscimed.2017.04.025>
- Carls-Diamante, S. (2023). Know thyself: Bipolar disorder and self-concept. *Philosophical Explorations*, 26(1), 110–126. <https://doi.org/10.1080/13869795.2022.2147982>
- Crippen, M. (2025). The emergence, loss, and reemergence of individuated self: Aesthetic flow and narrative in self-illness ambiguity. *Philosophical Explorations*, 1–20. <https://doi.org/10.1080/13869795.2025.2515053>
- de Haan, S. (2020). The need for relational authenticity strategies in psychiatry. *Philosophy, Psychiatry, & Psychology*, 27, 349–351. <https://doi.org/10.1353/ppp.2020.0044>
- de Haan, S. (2023). What do my problems say about me? *Philosophical Explorations*, 26(2), 159–164. <https://doi.org/10.1080/13869795.2023.2205416>
- Dings, R. (2020). *Not being oneself? Self-ambiguity in the context of mental disorder* (Doctoral dissertation, Radboud University Nijmegen).
- Dings, R., & de Bruin, L. C. (2022). What's special about 'Not Feeling like Oneself'? A deflationary account of self(-illness) ambiguity. *Philosophical Explorations*, 25(3), 269–289. <https://doi.org/10.1080/13869795.2022.2051592>
- Dings, R., & De Bruin, L. C. (2023). Editorial: Self-illness ambiguity and narrative identity. *Philosophical Explorations*, 26(2), 147–154. <https://doi.org/10.1080/13869795.2023.2203709>
- Dings, R., De Bruin, L. C., & De Boer, N. (2025). Self-illness ambiguity in personality disorder: Is it me or my disorder (that makes me do x)? In K. Banicki & P. Zachar (Eds.), *Conceptualizing personality disorder: Perspectives from philosophy and psychological science*. Cambridge University Press.
- Dings, R., & de Haan, S. (2022). The role of self-illness ambiguity and self-medication ambiguity in clinical decision-making. *American Journal of Bioethics*, 22(6), 58–60. <https://doi.org/10.1080/15265161.2022.2063436>
- Dings, R., & Glas, G. (2020). Self-management in psychiatry as reducing self-illness ambiguity. *Philosophy, Psychiatry, & Psychology*, 27(4), 333–347. <https://doi.org/10.1353/ppp.2020.0043>
- Dings, R., & Golova, A. (2025). 'Is it me or my illness?': Self-illness ambiguity as a useful conceptual lens for psychiatry. *The British Journal of Psychiatry*, published online. <https://doi.org/10.1192/bjp.2025.10367>

- Drozdźowicz, A. (2023). Self-illness ambiguity and Anorexia Nervosa. *Philosophical Explorations*, 26(1), 127–145. <https://doi.org/10.1080/13869795.2023.2168033>
- Durant, W. (2006). *The story of philosophy: The lives and opinions of the greater philosophers*. Pocket Books.
- Ervas, F. (2024). ‘It was the illness talking’: Self-illness ambiguity and metaphors’ functions in mental health narrative. *Philosophical Explorations*, 1–19. <https://doi.org/10.1080/13869795.2024.2437348>
- Estroff, S. E. (1989). Self, identity, and subjective experiences of schizophrenia: In search of the subject. *Schizophrenia Bulletin*, 15(2), 189–196. <https://doi.org/10.1093/schbul/15.2.189>
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511814594>
- Frankfurt, H. (1988). *The importance of what we care about: Philosophical essays*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511818172>
- Gallagher, S. (2024). *The self and its disorders*. Oxford University press.
- Giordano, S. (2005). *Understanding eating disorders: Conceptual and ethical issues in the treatment of Anorexia and Bulimia Nervosa*. Oxford University Press.
- Glas, G. (2023). Dimensions of self-illness ambiguity – a clinical and conceptual approach. *Philosophical Explorations*, 26(2), 165–178. <https://doi.org/10.1080/13869795.2023.2199013>
- Glas, G., & Dings, R. (2020). Dealing with self-illness ambiguity: A rebuttal. *Philosophy, Psychiatry, & Psychology*, 27(4), 353–354. <https://doi.org/10.1353/ppp.2020.0045>
- Greenberg, N. R. (2024). When the illness speaks. *BJPsych Advances*, 30(3), 164–165. <https://doi.org/10.1192/bja.2023.3>
- Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social & Clinical Psychology*, 21(6), 628–644. <https://doi.org/10.1521/jscp.21.6.628.22793>
- Hassall, R. (2024). Sense-making and hermeneutical injustice following a psychiatric diagnosis. *Journal of Evaluation in Clinical Practice*, 30(5), 848–854. <https://doi.org/10.1111/jep.13971>
- Hope, T., Tan, J., Stewart, A., & Fitzpatrick, R. (2011). Anorexia nervosa and the language of authenticity. *The Hastings Center Report*, 41(6): 19–29. <https://doi.org/10.1002/j.1552-146x.2011.tb00153.x>
- Horwitz, A. V., & Wakefield, J. C. (2007). *The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder*. Oxford University Press.
- Jeppsson, S. M. I. (2022a). Exemption, self-exemption, and compassionate self-excuse. *International mad Studies Journal*. <https://api.semanticscholar.org/CorpusID:256144108>
- Jeppsson, S. M. I. (2022b). Solving the self-illness ambiguity: The case for construction over discovery. *Philosophical Explorations*, 25(3), 294–313. <https://doi.org/10.1080/13869795.2022.2051589>
- Jorba, M., Petrolini, V., & Cepollaro, B. (2024). Person-first and identity-first approaches to autism: Meta-physical and linguistic implications. *Synthese*, 204(3), 111. <https://doi.org/10.1007/s11229-024-04749-y>
- Karp, D. A. (1992). Illness ambiguity and the search for meaning: A case study of a self-help group for affective disorders. *Journal of Contemporary Ethnography*, 21(2), 139–170. <https://doi.org/10.1177/089124192021002001>
- Kennett, J. (2009). Mental disorder, moral agency, and the self. In B. Steinbock (Ed.), *The oxford handbook of bioethics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199562411.003.0005>
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK Autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>
- 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>
- King, M., & May, J. (eds.). (2022). *Agency in mental disorder: Philosophical dimensions* (1st ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780198868811.001.0001>
- Kingma, E. (2013). Naturalist accounts of mental disorder. In K. W. M. Fulford (Ed.), *The oxford handbook of philosophy and psychiatry* (p. 363). Oxford University Press.
- Lakoff, G. (1992). Multiple selves: The metaphorical models of the self inherent in our conceptual system. <https://api.semanticscholar.org/CorpusID:148674829>
- Maiese, M. (2022). Self-illness ambiguity, affectivity, and affordances. *Philosophical Explorations*, 25(3), 363–366. <https://doi.org/10.1080/13869795.2022.2094999>
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175–183. <https://doi.org/10.1007/s10676-004-3422-1>

- McConnell, D., & Golova, A. (2023). Narrative, addiction, and three aspects of self-ambiguity. *Philosophical Explorations*, 26(1), 66–85. <https://doi.org/10.1080/13869795.2022.2115532>
- Nagel, T. (1979). *Mortal questions* (Vol. 89). Cambridge University Press.
- Pedley, R., Bee, P., Berry, K., & Wearden, A. (2017). Separating obsessive-compulsive disorder from the self: A qualitative study of family member perceptions. *BMC Psychiatry*, 17(1), 326. <https://doi.org/10.1186/s12888-017-1470-4>
- Pickard, H. (2015). Stories of recovery: The role of narrative and hope in overcoming PTSD and PD. In J. Z. Sadler, K. W. M. Fulford, & W. (C.W.) van Staden (Eds.), *The Oxford handbook of psychiatric ethics* (Vol. 2). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198732372.013.45>
- Pickard, H. (2017). Responsibility without blame for addiction. *Neuroethics*, 10(1), 169–180. <https://doi.org/10.1007/s12152-016-9295-2>
- Pugh, J. (2020a). *Autonomy, rationality, and contemporary bioethics* (1st ed.). Oxford University Press.
- Pugh, J. (2020b). Clarifying the normative significance of ‘personality changes’ following deep brain stimulation. *Science and Engineering Ethics*, 26(3), 1655–1680. <https://doi.org/10.1007/s11948-020-00207-3>
- Rego, M. D. (2004). Externality in psychiatry and the paradox of agency. *Philosophy, Psychiatry, & Psychology*, 11(4), 313–321. <https://doi.org/10.1353/ppp.2005.0013>
- Russell, J. L. (2025). Prescriptive ‘selves’ and self-illness ambiguity. *Synthese*, 206(2), 64. <https://doi.org/10.1007/s11229-025-05147-8>
- Sadler, J. Z. (2003). Mental health III: Issues in diagnosis. In S. F. Post (Ed.), *The encyclopedia of bioethics*. MacMillan Reference USA.
- Sadler, J. Z. (2007). The psychiatric significance of the personal self. *Psychiatry: Interpersonal & Biological Processes*, 70(2), 113–129. <https://doi.org/10.1521/psyc.2007.70.2.113>
- Sass, L. A. (2007). ‘Schizophrenic Person’ or ‘Person with Schizophrenia’?: An essay on illness and the self. *Theory & Psychology*, 17(3), 395–420. <https://doi.org/10.1177/0959354307073152>
- Schechtman, M. (2018). 4. The characterization question. In M. Schechtman (Ed.), *The constitution of selves* (pp. 73–92). Cornell University Press.
- Schechtman, M. (2022). Comment on ‘What’s special about “not feeling like oneself”?’: *Philosophical Explorations*, 25(3), 290–293. <https://doi.org/10.1080/13869795.2022.2077978>
- Shoemaker, D. (2015). *Responsibility from the margins*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198715672.001.0001>
- Snoek, A. (2023). How to overcome self-illness ambiguity in addiction: Making sense of one’s addiction rather than just rejecting it. A reply to McConnell and Golova. *Philosophical Explorations*, 26(1), 86–90. <https://doi.org/10.1080/13869795.2022.2140186>
- Strawson, P. (1982). Freedom and resentment. In G. Watson (Ed.), *Free will*. Oxford University Press.
- Tan, J. (2003). The anorexia talking? *Lancet*, 362(9391), 1246. [https://doi.org/10.1016/S0140-6736\(03\)14534-5](https://doi.org/10.1016/S0140-6736(03)14534-5)
- Tekin, Ş. (2014). Self-insight in the time of mood disorders: After the diagnosis, beyond the treatment. *Philosophy, Psychiatry, & Psychology*, 21(2), 139–155. <https://doi.org/10.1353/ppp.2014.0019>
- Tekin, Ş. (2021). Self and mental disorder: Lessons for psychiatry from naturalistic philosophy. *Philosophy Compass*, 16(1), e12715. <https://doi.org/10.1111/phc3.12715>
- Tekin, Ş. (2022). My illness, my self, and I: When self-narratives and illness-narratives clash. *Philosophical Explorations*, 25(3), 314–318. <https://doi.org/10.1080/13869795.2022.2097300>
- Tomm, K. (2019). Externalizing the problem and internalizing personal agency. *Journal of Systemic Therapies*. <https://api.semanticscholar.org/CorpusID:33441293>
- Tse, J. S. Y., & Haslam, N. (2024). Broad concepts of mental disorder predict self-diagnosis. *SSM - Mental Health*, 6(December), 100326. <https://doi.org/10.1016/j.ssmmh.2024.100326>
- Wilkinson, S. (2022). Psychiatric fictionalism and narratives of responsibility. *Philosophical Explorations*, 1–19. <https://doi.org/10.1080/13869795.2022.2116473>
- Yanos, P. T., Roe, D., & Lysaker, P. H. (2010). The impact of illness identity on recovery from severe mental illness. *American Journal of Psychiatric Rehabilitation*, 13(2), 73–93. <https://doi.org/10.1080/15487761003756860>