

Self-ascription and the *de se*

James Openshaw¹ 

Received: 24 August 2017 / Accepted: 3 April 2018 / Published online: 18 April 2018
© The Author(s) 2018

Abstract This paper defends Lewis’ (Philos Rev 88:513–543, 1979a) influential treatment of *de se* attitudes from recent criticism to the effect that a key explanatory notion—*self-ascription*—goes unexplained (Cappelen and Dever in The inessential indexical, Oxford University Press, Oxford, 2013; Holton, in: Loewer, Schaffer (eds) The Blackwell companion to David Lewis, Blackwell, Oxford, pp. 399–410, 2015). It is shown that Lewis’ treatment can be reconstructed in a way which provides clear responses. This sheds light on the explanatory ambitions of those engaged in Lewis’ project.

Keywords De se · Self-ascription · David Lewis · Centered worlds · Properties

1 Introduction

Rather than taking the contents of attitudes to be sets of possible worlds (Hintikka 1962)—ways a world might be—Lewis (1979a) proposed that we take them to be properties—ways an individual might be. A striking feature of Lewis’ discussion is its brevity. At face value, quick reflection on some basic thought experiments reveals that possible worlds contents cannot be the right objects of propositional attitudes after all. Recent writers have argued that the account Lewis (1979a) gives of *de se* attitudes suffers from a lacuna:

Special thanks to John Hawthorne, Tim Williamson, and anonymous referees for this journal for their comments. I am also grateful to the Arts and Humanities Research Council and the Royal Institute of Philosophy for their financial support during this period of research.

✉ James Openshaw
james.openshaw@philosophy.ox.ac.uk

¹ St. Edmund Hall, University of Oxford, Queen’s Lane, Oxford OX1 4AR, UK

...it is not the revisionary theory of content that does the solving, but the unexplained notion of *self-ascription of a property*... (Cappelen and Dever 2013: p. 103). [...] The role of self-ascription is little emphasized either by Lewis or in the subsequent literature. But the work of solving the informational puzzle[s] rests entirely on [...] self-ascription (2013: p. 108).

The second part [of Lewis' account], which is much less to the fore in Lewis's presentation and in the subsequent discussion, involves treating our attitude to these properties as that of *self-ascription* (Holton 2015: p. 400). Taking self-ascription as primitive is crucial to Lewis's account (Holton 2015: p. 403)

It is true that it is not just the revisionary theory of content that does the solving. But it is false that self-ascription is itself a primitive relation, incapable of being illuminated. In this paper, I provide the emphasis which Cappelen and Dever (2013) and Holton (2015) ask for. What those who endorse Lewis' influential move in characterizing *de se* attitudes need take as primitive is not some ad hoc attitude. In fact, whether or not Lewis' account of the *de se* is correct,¹ if it is problematic for the reasons to which recent critics have appealed, then theorists who attribute possible worlds (or centered worlds) contents to thinkers, and theorists who attribute contents which can only be evaluated for truth *at a context*, face the same concern. The complaints that “[t]aking self-ascription as primitive is crucial to Lewis's account” (Holton 2015: p. 403) and that this “unexplained notion” is what does the explanatory work (Cappelen and Dever 2013: p. 103) therefore miss their mark. Much more would need to be said in order to undermine Lewis' (1979a) characterization of *de se* attitudes.

I begin in Sect. 2 by outlining Lewis' (1979a) classic treatment of *de se* attitudes and clarifying the necessary details. Section 3 then returns to the matter of explaining the notion of self-ascription in a way which disarms the complaints of the above critics. Section 4 closes with a brief summary and conclusion.

2 Lewis, Lingens, and Frege

According to the framework in Lewis (1974), our folk-psychological talk and attitude-reporting practices are part of a ‘general theory of persons’, the underlying principles of which set out “how beliefs and desires and meanings are normally related to one another, to behavioral output, and to sensory input” (1974: p. 334). This tacit, general theory of persons may be systematized into a set of platitudes about belief and desire, and put this way it provides an implicit definition *of* the attitudes: they are whatever states play the right causal roles. Thus set out, Lewis thought, we are led to a program according to which contents are assigned “to characterize states of the head; to specify their causal roles with respect to behavior, stimuli, and one another” (1979a: p. 526).

The propositions Lewis (1974) had in mind were possible worlds contents. But that conception of the objects of propositional attitudes was revised in his (1979a) in which

¹ For a presentation of other considerable difficulties facing Lewis' (1979a) account of mental content, see Magidor (2015). I will not be defending the account against these objections. Neither will I be addressing scepticism as to whether there *is* any distinctive kind of *de se* representation (*inter alia*, Cappelen and Dever 2013).

he cites several arguments-by-thought-experiment for the claim that an adequate theory of intentionality must recognise an irreducible kind of indexicality: the phenomenon of *de se* attitudes. As Lewis later summarized:

The contentful unity is the entire system of beliefs and desires [...]. Its content is defined, insofar as it is defined at all, by constitutive rationality on the basis of its typical causal role. The content is in the first instance narrow and *de se* (1999: p. 324).²

Consider the plight of Rudolf Lingens, an amnesiac lost in Main library, Stanford. In an attempt to discover who and where he is, he reads many books, including one which is in fact a biography of his own life and another which is in fact a description of Main library. Unfortunately for Lingens, while he may learn that Main library is next to Hoover tower,

[h]e still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, 'This place is aisle five, floor six, of Main Library, Stanford. I am Rudolf Lingens.' (Perry 1977: p. 492).

The conclusion drawn is that not all knowledge is propositional knowledge, where propositions are stipulated to be sets of possible worlds. Lingens might have all of *that* sort of knowledge and yet still fail to know who and where he is. Lingens needs more information than is given by a partitioning of the set of worlds. Call the distinctive sort of knowledge lacked by Lingens *de se* knowledge. A more fine-grained sort of content is required to capture the *de se* state expressed by the sentence 'I am Rudolf Lingens'. The kind of content we must seek will partition not only possible worlds but the individuals at those worlds.³

Lewis (1979a) proposed that we conveniently take the objects of *de se* states to be *properties*. A possible worlds content determines a set of worlds: those at which the content is true. And for any set of worlds we may talk of the corresponding *world set property*: the property of being at a world in that set. This suffices to establish that properties are at least as fine-grained as possible worlds contents. However, properties also go further: if *x* instantiates property *F* while some worldmate *y* does not, *F* cannot correspond to a proposition. According to Lewis (1979a), we can then say that what Lingens needs to do, in order to find out who and where he is, is not to believe some possible worlds content but to *self-ascribe* some property: to 'locate' himself not only among worlds but within them. In particular, Lewis' (1979a) conclusion is that Lingens' revelation involves something more than just the self-ascription of a world set property. After all, learning that Main library is next to Hoover tower did not help Lingens find out where *he* is. Lingens' revelation must involve the self-

² See also Lewis (1986: p. 36).

³ Perry's (1977) own proposal was to instead suggest that the belief relation is tertiary: we have the agent, the content believed, and the agent's belief state.

ascription of a more fine-grained property, one which cuts across the individuals *at* worlds.⁴

The notion of self-ascription is introduced casually by Lewis (1979a: p. 518) as the relation a subject must bear to a property in order to count as entertaining a *de se* belief.⁵ We may also follow Lewis (1979a) in translating talk of property-contents into an analogue of what Quine (1968) called *centered worlds* contents, trading in a property for a set of ordered pairs of a world and an individual at that world. As Lewis suggests, “[a] class of centered worlds corresponds to a property” (1979a: p. 532).⁶ The idea is that both kinds of content are eligible to play the theoretical role of accommodating the *de se*: “centered worlds [also] amount to presentations of possible individuals” (1983a: p. 25, n. 18). That these are intended by Lewis as two ways of saying the same thing despite the latter involving no talk of self-ascription might raise *prima facie* suspicions about the alleged essentiality of that notion for Lewis’ (1979a) account.

Granting the assumption that properties can in general be matched up one–one with sets of centered worlds, the results, for Lewis, are the following two equivalent characterisations of *de se* belief.

BELIEF DE SE:

S believes *de se* (at *w*) that she is *F* iff

- (i) every centered world $\langle w', s' \rangle$ in the set of centered worlds compatible with *S*’s beliefs (at *w*) is such that *s'* is *F* at *w'*.

Or, equivalently,

- (ii) *S* self-ascribes (at *w*) $\lambda x \cdot x$ is *F* at *w*.⁷

Since properties self-ascribed are at least as fine-grained as possible worlds contents believed, it should be possible to reduce all belief to *de se* belief. Notoriously, this is what Lewis (1979a) proposes. To believe the proposition that cyanoacrylate glue dissolves in acetone is to self-ascribe the property of being located within the set of

⁴ As Holton (2015: p. 402) observes, what Lingens really needs to find out who and where he is is ‘break-through knowledge’: *de se* knowledge which connects up with the *de dicto* knowledge he gains in the library. What kinds of properties is it possible to self-ascribe? In short, any way an individual could be. If you were to believe you were a poached egg, you would thereby self-ascribe $\lambda x \cdot x$ is a poached egg. See Lewis (1979a: p. 530).

⁵ Note that self-ascribing a property-content is local to the characterisation of belief. Perhaps in the case of *desiring de se* that one eat pasta we could speak of self-prescribing the content $\lambda x \cdot x$ is eating pasta. On counterfactual attitudes such as imagining that, and the problems they raise for Lewis’ (1979a) account, see Ninan (2013).

⁶ On whether a one–one correspondence between properties and centered worlds holds, see Holton (2015: pp. 403–405). There is also a cluster of worries here familiar to unstructured views of mental content. If properties are only as fine-grained as centered worlds, will Sally’s ignorance that being a podiatrist is the same property as being a foot doctor result in us construing her as having irrational beliefs? And will she come out as trivially believing ‘logical truths’ like ‘I exist’ (Kaplan 1989)? I set these kinds of worries aside in this paper.

⁷ I use lambda calculus in the specifications of properties for lucidity. Holton (2015) observes that we will sometimes have to be resourceful with the properties chosen to be contents of attitudes. For example, the belief that “I chose myself to assess myself [...] seems to involve self-ascribing the property of self-assessing-choosing” (2015: p. 408). The following specification makes Lewis’ proposal much more tractable, however: $\lambda x \cdot x$ chose *x* to assess *x*.

worlds at which cyanoacrylate glue dissolves in acetone: that is, $\lambda x \cdot x \text{ inhabits some } w \in w'$ where $\{w': \text{cyanoacrylate glue dissolves in acetone at } w'\}$ (1979a: p. 518). In general, to have a *de dicto* belief is to self-ascribe what I earlier called a world set property.

While there are differences between the case of Lingens and others which Lewis (1979a) considers (the mad Heimson, the two gods, the insomniac...), the underlying argument is the same. In spite of the fact that the subject knows all of the relevant propositions, there is *some information* she lacks. It is tempting to draw an analogy between this sort of ignorance and the ignorance displayed in standard Frege cases. Indeed, one might worry that to the extent that there is a puzzle about propositional attitudes raised by cases like that of Lingens the amnesiac, it is reducible to familiar puzzles of cognitive significance.

The case [...] gives us no reason to accept a revisionary theory of content—it is, at best, just another instance of Frege’s puzzle, [and] will be solved by whatever our general theory of opacity is... (Cappelen and Dever 2013: p. 103).

According to the Fregean [...] Lingens has [not] learnt all the (relevant) true propositions: the proposition that is expressed (in the relevant context) using the sentence ‘You are in the Stanford library’ is not the same proposition as the one expressed by [‘Lingens is in the Stanford library’] [...]. The Lingens case thus poses no special puzzle (Magidor 2015: pp. 254–255).

In combination with the concerns as to whether the notion of self-ascription is in good standing, this line of objection forms a two-pronged attack. In response, those who endorse Lewis’ above characterization of *de se* attitudes must do two things: (a) explain why we should treat the two superficially similar informational puzzles differently; (b) explain the ‘unexplained’ notion of self-ascription. In what remains of the present section, I provide a brief explanation of the first kind, clarifying important details of Lewis’ (1979a) picture along the way. Section 3 addresses (b).

The transition to the property-content framework above allows a unified solution to both Perry- and Frege-puzzles. Not only did Lewis reduce the *de dicto* to the *de se*, he reduced the (so-called) *de re* to the *de se*. On Lewis’ (1979a, 1983a) account subjects entertain *de re* thoughts only with respect to a *relation of acquaintance*, where *S* and *o* are acquainted just in case there exists an “extensive causal dependence of [*S*’s] states upon [*o*’s]”; and this causal dependence is of a sort apt for the reliable transmission of information” (1979a: p. 542). Consider the following example. Ba the Babylonian is familiar with the planet Venus under two different guises. He knows it as *Phosphorus*, the morning star, and also as *Hesperus*, the evening star. He is unaware that the two are identical and believes that Phosphorus is a young star while Hesperus is old. Ba bears both the *P*-relation and the *H*-relation to Venus, where *P* is the relation *x* bears to *y* iff *y* is the unique object which *x* sees shining brightly in the morning sky, and *H* is the relation *x* bears to *y* iff *y* is the unique object which *x* sees shining brightly in the evening sky. The contents of Ba’s beliefs are the properties $\lambda x \cdot x \text{ bears } P \text{ uniquely to a young star}$ and $\lambda x \cdot x \text{ bears } H \text{ uniquely to an old star}$. By characterising the contents of subjects’ belief states in this way, we capture the differences in cognitive significance and dispositions to behaviour between Ba’s belief that Hesperus is visible

($\lambda x \cdot x$ bears H uniquely to something which is visible) and his belief that Phosphorus is visible ($\lambda x \cdot x$ bears P uniquely to something which is visible).

In general, then, we have the following characterisation of *de re* belief (Lewis 1979a, 1983a).

BELIEF DE RE:

S believes *de re* (at w) that x is F iff:

- (a) there is some acquaintance relation R such that S bears R to x uniquely (at w)
- (b) every centered world $\langle w', s' \rangle$ in the set of centered worlds compatible with S 's beliefs (at w) is such that the thing to which s' bears R at w' is F at w' .

Or, equivalently (to (b)),

- (c) "...the subject self-ascribes [(at w)] the property of bearing relation [R] uniquely to something which has property [F (at w)]" (Lewis 1979a: p. 539).

So the way in which the Lewisian resolves to treat the Frege puzzles, in contrast with the *de se* puzzles, is justified by the fact that we ultimately achieve a unified account of both phenomena.⁸ The burden of explaining the elusive notion of self-ascription remains, however.

3 Self-ascription

...even if one accepts that there is a special problem to be solved [i.e. beyond the general problem of opacity], it is not the revisionary theory of content that does the solving, but the unexplained notion of *self-ascription of a property*... (Cappelen and Dever 2013: p. 103).

If the characterizations in (i) and (b) were available to Lewis, why did he introduce (ii) and (c) along with the elusive notion of self-ascription? Is self-ascription dispensable after all? This section clarifies the explanatory project in which Lewis was engaged and uses this clarification to resolve these and other critical questions raised at the outset of the paper. What those who endorse either of Lewis' characterizations of belief *de se* must take as primitive is not some ad hoc attitude.

Lewis (1983b: p. 230) observes that it is possible to represent the semantic notion of a *context of use* as an ordered pair of a world and a speaker at that world.⁹ Everything else, including 'standards of permissibility' (Lewis 1979b: pp. 340–341), 'rankings of comparative salience' (ibid. 348–50), 'standards of precision' (ibid. 351–4), *inter*

⁸ While Cappelen and Dever (2013: pp. 103–108) and Magidor (2015) mention this response, they take it to be unsuccessful because it fails to provide answers to further problems. Cappelen and Dever's complaints are the subject of Sect. 3. Magidor (2015: pp. 266–271) worries that the descriptivist elements of Lewis' treatment of *de re* attitudes are problematic (see also Holton (2015: pp. 405–406)). I focus on only the former complaint here, since my aim is to show that Lewis' (1979a) story can be reconstructed as a *bona fide* treatment of the *de se*, even if his broader picture of singular thought suffers defects.

⁹ Purely for convenience of presentation here, I assume that agents are (necessarily) time-slices of individuals. There are alternatives, of course, depending on one's metaphysical views. If agents are (necessarily) world-bound then the world element here is redundant: a context or centered world is simply an individual's singleton set. On the other hand, if we deny that agents are (necessarily) time-slices of individuals, we will need a time element. The important point is that contexts and centered worlds have the same structure, whatever structure that is.

alia, may be recovered from these minimal coordinates. According to Lewis (1980), the input to compositional semantics for languages eliciting context-sensitivity will be a pair of a sentence-type S and a context of use c . The output will be an assignment of a truth-value to the semantic value of S at every index (every ordered pair of a world and speaker). At this point we face a challenge. If what gets evaluated for truth is a sentential semantic value with respect to each index, what does such a notion—truth *at an index*—have to do with any sort of truth which might plausibly norm linguistic practice? What we want is for the semantic values of natural language sentences to receive truth-values which, by default, tell us something about our de facto utterance situations. On the orthodox way of resolving this challenge, we choose to set or *initialize* the index parameter values, by default, to those corresponding to the given utterance situation.¹⁰ As Lewis puts it,

...we must distinguish two sorts of indices: *Original indices*, in which the shiftable features are as determined by the context [...]; and *shifted indices*, in which that is not so. Truth in a context is truth at an original index, and this is the semantic notion that is directly relevant to truthful speech (Lewis 1983b: p. 231).

By accessing the context of use, we can extract a notion of truth with communicative import—truth at a context of use—from the notion of truth at an index. Specifically: for all sentences S , S is true at c iff the semantic value of S is true at $\langle w_c, s_c \rangle$.

Now what does all this have to do with self-ascription? In a footnote to the (1983b) postscript to ‘General Semantics’, Lewis remarks:

It will not escape the reader of ‘Attitudes *De Dicto* and *De Se*’ [...] that these ‘contexts’ are the same as the ‘subjects’ that self-ascribe properties, and that sets of them are the self-ascribed properties (1983b: p. 230, n. 2).¹¹

This telling remark suggests the following picture. Contexts are centered worlds. And sets of centered worlds, or indeed of contexts, are equivalent to property-contents. A subject’s total belief state, the content of which is narrow and *de se* according to Lewis, determines a set of centered worlds (equivalently, a collection of property-contents). For example, the content expressed by the belief that cyanoacrylate glue dissolves in acetone is a set of centered worlds $\langle W', S' \rangle$ (equivalently, the property $\lambda x \cdot x$ inhabits some $w \in W'$ where $\{W'\}$: cyanoacrylate glue dissolves in acetone at W').

Notice that *none* of these contents, centered worlds or properties, can be evaluated for truth simpliciter. If property-contents, for instance, are to be truth-evaluable at all, we need a specification of an object which serves to determine the truth-value of the property-content. By analogy, what would it be for us to evaluate the sentence-type ‘It is raining’, independent of any specification of time or place? This is where the semantic resources above enter in. It is the process of *initialization*—of defining truth at a context of use via truth at an index—which forces our hand to evaluate belief

¹⁰ See also Heim and Kratzer (1998: p. 243), Kaplan (1989: p. 547) and MacFarlane (2014: p. 53).

¹¹ This remark again indicates Lewis’ intended equivalence between properties and sets of centered worlds.

contents in a way which privileges the world and agent of the context of the belief over the coordinates of some arbitrary index.

Recall that Lewis introduced the notion of self-ascription casually just to denote the relation a subject must bear to a property in order to count as having a *de se* belief (1979a: p. 518). In light of the proposal suggested by this (1983b) remark, it is this process of initialization which secures the right relation between subject and set of centered worlds (or property-content) believed.¹² Talk of self-ascription is then just an intuitive gloss on work performed by the stipulative process of initialization. This is our first insight into the character of Lewis' (1979a) proposal. Self-ascription is not some ad hoc attitude. It is a phrase used to express the relation that a believer, *S*, bears to a property-content which is true at *c* where *S* is the subject of *c*.¹³

In the remainder of this section I provide responses to three worries. Each of these responses will serve to clarify the reconstruction of Lewis' (1979a) account of *de se* belief I have begun to suggest. The third worry, in particular, will clarify what notions Lewis' characterisation does take as primitive.

First, the reconstruction above draws on remarks Lewis made about language. In spite of the encouragement in the quote above (1983b: p. 230, n. 2), conclusions about the nature of thought drawn from conclusions about the nature of language require serious scrutiny. Specifically, indices are devices used by linguistic semanticists to capture the displacement of sentences across parameters such as time and location. It is arguably a necessary condition on positing an index-parameter in a semantic theory for a language that there be a sentential operator which shifts that parameter. Does that mean that we have to not only make sense of but *posit* 'Mentalese' index-shifting operators to support Lewis' claims about the index-relativity of belief?

No. To see why, let us consider the distinction Lewis (1980) emphasises between sentential semantic value and content.¹⁴ First, suppose we chose not to recognise any such distinction, affirming that the semantic values of sentences are contents (à la Kaplan (1989)). In that case, it would not be a plausible necessary condition on positing an index-parameter that there be a sentential operator which shifts that parameter. For not all constraints on the nature of content come from semantic theorising. Lewis' (1979a) point was precisely that thought experiments like Perry's (1977) give us good, and good enough, reason to think of mental content as an agent-relative entity.

On the other hand, suppose we were to recognise a distinction between the semantic values of sentences and contents. In that case, despite Lewis' (1983b) use of the term 'indices' in the remark quoted above, it is clear that Lewis' intent would be to talk about points of evaluation: sequences of parameters against which contents are to be

¹² Specifically, when $\langle W', S' \rangle$ is the content of a belief of *s* at world *w*—which is to say at context *c*, where *c* is $\langle w, s \rangle$ —this centered worlds content is evaluated for *truth at c*. It is *true at c* just in case $\langle w, s \rangle$ is a member of $\langle W', S' \rangle$.

¹³ A referee for this journal has noted that similar suggestions about the role of self-ascription in Lewis' (1979a) account appear in Recanati (2007) and Pagin (2016, especially p. 277): "[w]henver a representation is tokened in the subject's mind, the content of the representation is evaluated with respect to a situation involving the subject and the time and place of the tokening. In other words, the content is construed as a property which is ascribed to the context in which the representation is tokened" (Recanati 2007: p. 269).

¹⁴ Lewis appreciates the possibility of theorizing about the nature of mental content independent of 'psychological' theorizing and its would-be semantic values at (1980: p. 83).

evaluated.¹⁵ The notions of index and point of evaluation were not yet distinguished at the time of the (1970) paper which his (1983b) postscript concerns. But Lewis did later emphasise a distinction between these two notions.¹⁶ Once that distinction is in place, we open the possibility to evaluating contents with respect to a parameter—for example, an agent or possible world—for reasons somewhat independent of compositional semantics. And once that possibility opens up, it becomes possible to theorise about the nature of mental content somewhat independent of ‘psycholinguistic’ semantic theorising.¹⁷ Here, again, Lewis’ (1979a) point was that thought experiments like Perry’s (1977) give us good, and good enough, reason to think of mental content as an agent-relative entity.

A second worry concerns the relationship between believer and content believed. If property-contents are to be truth-evaluable, we need a specification of an object. Since Lewis (1979a) claims that all thought is *de se* thought, we need the property to be evaluated for truth with respect to the believer at the context of the belief in question. What determines that it is the believer? Likewise, on the centered worlds characterization of mental content, in virtue of what does the centre of a centered world represent the believer?

I said above that initialization—the process of defining truth at a context of use via truth at an index (or point of evaluation)—is what enables us to evaluate contents in a way which privileges, by default, the world and subject of the context of use. As Ninan (2008) has pointed out, there is something merely stipulative about the answer to this second worry: “[there is] an implicit stipulation that, when we consider a centered world as a candidate doxastic alternative for some agent *x*, the centre of that centered world represents *x*” (59). Truth at an index, or truth at a point of evaluation, bears little familiarity with the notion of truth which norms linguistic practice and inquiry. Opting to define truth *at a context* is what allows us to recapture the familiar notion of truth. In Lewis’ words: “Truth at a context is truth at an original index [or context-initialized point of evaluation], and this is the semantic notion that is directly relevant to truthful speech” (1983b: p. 231). So the fact that the centre of a centered world content represents the believer is no coincidence. This is explicitly stipulated in the intentional theory. Moreover, because Lewis (1979a: p. 518) introduces self-ascription just to denote the relation a subject must bear to a property in order to count as having a *de se* belief, and because all belief is *de se* belief, what it is for *S* to self-ascribe property-content *P*, it turns out, is for the process of initialization for *P* to determine that *S* is the believer at the context at which *P* is to be evaluated for truth.

A third and final worry. Even if this story elucidates the mechanics of self-ascription, how does it get us beyond a treatment of mere ‘*x*-ascription by *x*’, a set of attitudes of which the *de se* is a proper subset?¹⁸ For it is possible to ascribe a property to oneself

¹⁵ To use Kaplan’s (1989) distinction, Lewis (1983b) is not talking about content-generating parameters but content-evaluating parameters.

¹⁶ On Lewis’ distinction between sentential semantic value and content, see his (1980) and Rabern (2017).

¹⁷ As Rabern (2012) puts it: “Theorists working on the nature of assertoric content (or mental content and information) should welcome this distinction, as it allows them to theorize about the nature of content, somewhat liberated from the confines of the strict compositionality principle” (94–5).

¹⁸ This objection, and its phrasing, is due to Cappelen and Dever (2013).

without *self-ascribing* that property. Consider Kaplan's (1989: p. 533) example of ascribing the property of being on fire to *that man*, where the man in question is, unbeknownst to Kaplan, himself.

It will serve us well to remember how things work on the classical possible worlds semantics for belief (Hintikka 1962). A subject's belief state is taken to determine a set of possible worlds: those worlds which are *compatible with* what the subject believes. This notion of compatibility with what a subject believes is taken as primitive. The goal, in short, is to characterize the subject's belief state in a way which makes various doxastic and logical relations explicit. For Lewis (1979a), of course, belief simpliciter will not do. This is precisely the point of the informational puzzles about Lings and other characters, recalled in Section 2. What Lewis' (1979a) decision amounts to is to take as primitive the notion of a set of centered worlds (or, equivalently, a property-content) being compatible with what a subject believes *de se*. In other words, Lewis' (1979a) primitive *just is de se belief*.¹⁹

The biconditionals in BELIEF DE SE, above, do not have the status of reductive analyses of *de se* belief. Lewis (1979a) simply does not pursue any such analysis. The target notion is in fact implicitly appealed to on the right side of the biconditionals. As we might say more explicitly:

BELIEF DE SE (EXPLICIT):

S believes *de se* (at *w*) that she is *F* iff:

- (i) every centered world $\langle w', s' \rangle$ in the set of centered worlds compatible with *S*'s *de se* beliefs (at *w*) is such that *s'* is *F* at *w'*

Or, equivalently,

- (ii*) *S* believes *de se* (at *w*) $\lambda x \cdot x$ is *F* at *w*

For Lewis, all belief is *de se* belief. And all belief expresses a content which is to be evaluated for truth at a context. This is the import of the irreducible sort of indexicality needed to treat the puzzle about Lings. The *de se* is not analysed but taken as a fundamental component of Lewis' intentional theory. The aim is to provide a *non-reductive* characterization of belief *de se*, analogous to Hintikka's (1962) non-reductive characterization of belief '*per se*'. This, according to Lewis (1979a) is what we need if we are 'to characterize the states of Lings' head; to specify their causal roles with respect to behavior, stimuli, and one another'.²⁰

4 Conclusion

To summarise, whichever way you slice it—(i*) or (ii*)—what is sought is a non-reductive characterisation of belief *de se*. Lewis achieves this characterisation by stipulating (in the way of a possible worlds semanticist with belief simpliciter) that the phenomenon of belief *de se* be tracked by the specification of a certain sort of content. Such content can only be evaluated for truth *at a context*; a context providing a believer. Self-ascribing is not something the subject does. It is a phrase used to

¹⁹ Weber (2016: p. 257) makes a similar observation about the status of Lewis' proposal as against Hintikka's (1962).

²⁰ Lewis (1979a: p. 526).

express the relation borne by S to a property-content which is true at c , where S is the believer at c ; an arrangement stipulated to track S 's *de se* beliefs. Cappelen and Dever (2013) are right that a mere shift in the so-called objects of belief is not what solves the thought experiments to which Lewis (1979a) appealed as motivation, and Holton (2015) is right that a new primitive is introduced in order to support the solution he provided. But that primitive is *compatibility with a subject's beliefs de se*, replacing the primitive of the classical possible worlds account, *compatibility with a subject's beliefs per se*. This is made explicit in (i*) and (ii*), above.

Whether or not Lewis' characterization of *de se* belief is correct, then, if it is problematic for the reasons to which recent critics have appealed, then theorists who attribute even possible worlds contents to thinkers, and theorists who in general attribute contents which can only be evaluated for truth at a context, face the same problems. This is not, of course, to say that such critics would be wrong to have gripes with such primitives, or with the explanatory interest of the resulting projects. Insofar as these constitute their gripes, however, they apply far wider than to Lewis' (1979a) characterization of *de se* belief. Correctly understood, then, Lewis' (1979a) characterization of *de se* belief is resistant to much of the criticism it has faced in the recent literature.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Cappelen, H., & Dever, J. (2013). *The inessential indexical*. Oxford: Oxford University Press.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Hintikka, J. (1962). *Knowledge and belief: An introduction to the logic of the two notions*. Ithaca, New York: Cornell University Press.
- Holton, R. (2015). Primitive self-ascription: Lewis on the *de se*. In B. Loewer & J. Schaffer (Eds.), *The Blackwell companion to David Lewis* (pp. 399–410). Oxford: Blackwell.
- Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.
- Lewis, D. (1970). General semantics. *Synthese*, 22(1), 18–67.
- Lewis, D. (1974). Radical interpretation. *Synthese*, 27, 331–344.
- Lewis, D. (1979a). Attitudes *de dicto* and *de se*. *Philosophical Review*, 88, 513–543.
- Lewis, D. (1979b). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339–359.
- Lewis, D. (1980). Index, context, and content. In S. Kanger & S. Öhman (Eds.), *Philosophy and grammar* (pp. 79–100). Dordrecht: Reidel.
- Lewis, D. (1983a). Individuation by acquaintance and by stipulation. *Philosophical Review*, 92, 3–32.
- Lewis, D. (1983b). *Postscripts to 'General semantics'*. In *his philosophical papers* (Vol. 1, pp. 230–232). Oxford: Oxford University Press.
- Lewis, D. (1986). *On the plurality of worlds*. Oxford: Blackwell.
- Lewis, D. (1999). *Reduction of mind*. In *his papers in metaphysics and epistemology* (pp. 291–324). Cambridge: Cambridge University Press.
- MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.
- Magidor, O. (2015). The myth of the *de se*. *Philosophical Perspectives*, 29, 249–283.
- Ninan, D. (2008). *Imagination, content, and the self*. Ph.D. thesis submitted to the Massachusetts Institute of Technology.

- Ninan, D. (2013). Self-location and other-location. *Philosophy and Phenomenological Research*, 87(2), 301–331.
- Pagin, P. (2016). *De se* communication. In M. García-Carpintero & S. Torre (Eds.), *About oneself: De se thought and communication* (pp. 272–306). Oxford: Oxford University Press.
- Perry, J. (1977). Frege on demonstratives. *Philosophical Review*, 86, 474–497.
- Quine, W. V. O. (1968). Propositional objects. *Crítica: Revista Hispanoamericana de Filosofía*, 2, 3–29.
- Rabern, B. (2012). Against the identification of assertoric content with compositional value. *Synthese*, 189(1), 75–96.
- Rabern, B. (2017). A bridge from semantic value to content. *Philosophical Topics*, 45(2), 181–207.
- Recanati, F. (2007). *Perspectival thought*. Oxford: Oxford University Press.
- Weber, C. (2016). Being at the centre: Self-location in thought and language. In M. García-Carpintero & S. Torre (Eds.), *About oneself: De se thought and communication* (pp. 246–271). Oxford: Oxford University Press.