



Inquiry

An Interdisciplinary Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/sinq20>

Individuation by agreement and disagreement

Torfinn Thomesen Huvenes

To cite this article: Torfinn Thomesen Huvenes (2022): Individuation by agreement and disagreement, *Inquiry*, DOI: [10.1080/0020174X.2022.2051203](https://doi.org/10.1080/0020174X.2022.2051203)

To link to this article: <https://doi.org/10.1080/0020174X.2022.2051203>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 14 Mar 2022.



Submit your article to this journal [↗](#)



Article views: 741



View related articles [↗](#)



View Crossmark data [↗](#)

Individuation by agreement and disagreement

Torfinn Thomesen Huvenes 

Department of Philosophy, University of Bergen, Bergen, Norway

ABSTRACT


It is common to explain agreement and disagreement in terms of relations among mental states. The main purpose of the present discussion is to present an alternative way of thinking about the relationship between mental states and agreement and disagreement. The idea is to connect agreement and disagreement with the individuation of mental states. More specifically, for at least some mental states, standing in the same relations of agreement and disagreement is both necessary and sufficient for identity. This provides us with a framework for discussing issues related to agreement and disagreement. It is for instance well-known that *de se* beliefs raise questions about agreement and disagreement. The present framework allows us to discuss such beliefs while relying on assumptions that are interestingly different from the standard assumptions in the literature. While most of the discussion focuses on beliefs, there are other mental states, such as desires, that can seemingly also stand in relations of agreement and disagreement. This raises questions about how far the framework can be extended.

ARTICLE HISTORY Received 17 September 2021; Accepted 4 February 2022

KEYWORDS Agreement; disagreement; mental states

1. Introduction

In the recent philosophical literature on agreement and disagreement, the standard approach has been to explain agreement and disagreement in terms of relations among mental states. The literature has seen the development of several accounts along these lines. For instance, using MacFarlane's (2014, ch. 6) terminology, a mental state may preclude the accuracy or satisfaction of another mental state.¹ That is one way to explain disagreement. Another option is to explain disagreement in terms of noncotenability. That is to say that it is impossible to coherently

CONTACT Torfinn Thomesen Huvenes  torfinn.huvenes@uib.no

¹See e.g. Stevenson (1963), Jackson (2008), MacFarlane (2014, ch. 6), Marques and García-Carpintero (2014), and Marques (2015) for further discussion.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

have both mental states at the same time.² As MacFarlane points out, these options do not have to be treated as competitors. It is possible that there are just different varieties of disagreement, perhaps with some features that they have in common.³

The purpose of the present discussion is to explore what happens if we approach the relationship between mental states and agreement and disagreement from a different direction. The plan is to start from a thesis that connects agreement and disagreement with the individuation of mental states. The idea is that, for at least some mental states, standing in the same relations of agreement and disagreement is both necessary and sufficient for identity.

Unlike much of the recent literature on agreement and disagreement, this is not an attempt to say something about what it is to agree or disagree. Instead, the idea is to use agreement and disagreement to give identity-conditions for mental states. That said, we can still use the present framework to discuss the kind of issues that have been prominent in the literature. It is for instance well-known that *de se* beliefs raise questions about agreement and disagreement. I am going to argue that the framework allows us to discuss such beliefs while relying on assumptions that are interestingly different from the standard assumptions in the literature.

Ultimately, the framework will have to be assessed on the basis of its treatment of a wider range of cases. For now, it is a matter of demonstrating the potential implications of the framework. That also takes precedence over discussing how the framework relates to other ways of individuating mental states. That said, we are going to see that the framework can be combined with different assumptions about the role of content in individuating mental states. These assumptions are particularly important insofar as we do not start out by assuming anything about how agreement and disagreement relate to content.

In section 2 I introduce the main thesis that provides the core of the framework that I will be employing throughout the subsequent discussion, as well as some additional assumptions. In section 3 I discuss the formal properties of the agreement and disagreement relations. A central question is whether agreement is an equivalence relation. In light of the discussion, I propose to distinguish between two relations: full agreement and partial agreement. In section 4 I use the framework to discuss *de se* beliefs. The conclusion is that the framework places

²See e.g. Dreier (2009, 105–106) and MacFarlane (2014, ch. 6) for further discussion.

³See e.g. Zeman (2020) for discussion of the idea that there are some features that the different notions of disagreement have in common.

constraints on an account of de se beliefs. However, these constraints are in the first instance constraints on the individuation of the relevant beliefs, and not their contents. In section 5 I discuss whether the framework can be extended to other mental states than beliefs. I argue that there are good prospects of extending the framework to desires. However, other mental states, such as imaginings and perceptual experiences, are more recalcitrant.

2. The framework

The following thesis provides the core of the framework that I will be employing throughout the discussion:

Thesis 1: Necessarily, for all mental states σ_1 and σ_2 , $\sigma_1 = \sigma_2$ if and only if σ_1 and σ_2 agree and disagree with all the same mental states.

[Thesis 1] gives us necessary and sufficient identity-conditions for mental states. In what follows, I am not going to attempt to motivate [Thesis 1] by arguing for it directly. Instead, its motivation ultimately depends on whether it can provide a useful framework for theorising about the individuation of mental states. In that sense, the methodology is abductive. The main purpose of the present discussion is to present the framework and discuss some of its implications. That is the first step towards assessing [Thesis 1].

However, as it stands, [Thesis 1] raises several questions. Some of these will have to be addressed later. For instance, there is the question of what we can and cannot assume about the agreement and disagreement relations. That question will be addressed in section 3. [Thesis 1] also seems more plausible for some kinds of mental states than others. It might for instance be that it holds for beliefs, but not for, say, perceptual experiences. That is something that we will return to in section 5.

That said, there are four points of clarification that are worth addressing before we proceed. The first point of clarification concerns the worry that [Thesis 1] is circular due the expression ‘same mental states’ which appears on the right-hand side of the biconditional. This might seem like just another way of talking about the identity of mental states. However, there is no circularity. The right-hand side of the biconditional can be stated without using identity. To say that σ_1 and σ_2 agree and disagree with all the same mental states is just to say that for every mental state σ_0 , σ_1 agrees with σ_0 if and only if σ_2 agrees with σ_0 and σ_1 disagrees with σ_0 if and only if σ_2 disagrees with σ_0 . That is how [Thesis 1] is to be understood.

The second point of clarification concerns the type-token distinction. [Thesis 1] only holds for types of mental states. For instance, let us suppose that *A* and *B* believe that it is snowing.⁴ While *A* and *B* have distinct token mental states, these can still be tokens of the same type. In that sense, *A* and *B* may still have the same mental state, the belief that it is snowing. Furthermore, that belief may be distinct from another belief, such as the belief that it is windy. It is in this sense that we will be discussing the individuation of mental. For instance, if we ask whether *A*'s belief is identical to *B*'s belief, the question is whether they are beliefs of the same type. In other words, it is the individuation of types of mental states that matter for the present discussion, not token mental states.

The third point of clarification concerns the relationship between [Thesis 1] and other ways of individuating mental states. It is important to note that, as it is stated, [Thesis 1] does not rule out other constraints on the individuation of mental states. That is particularly relevant with regards to the role of content in individuating mental states like beliefs. For instance, one might hold that having the same the content is necessary for identity in the case of beliefs. Unless there are beliefs with distinct contents that agree and disagree with all the same mental states, that is compatible with [Thesis 1]. In fact, in section 4 we are going to discuss how we can derive conclusions about the content of *de se* beliefs by combining [Thesis 1] with additional assumptions about content.

The fourth and final point of clarification concerns the fact that [Thesis 1] is formulated in terms of agreement and disagreement between mental states.⁵ That is arguably not the way we ordinarily talk about agreement and disagreement.⁶ It is more common to talk about agreement between an individual and another individual or an action or attitude. For instance, there is nothing problematic about saying that *A* agrees with *B* or that *A* agrees with what *B* says or believes. It is perhaps less natural to say that *A*'s belief agrees with *B*'s belief. This raises the legitimate worry that we could end up pushing the relevant notions too far by talking about agreement and disagreement between mental states.

⁴For the sake of simplicity, we are ignoring the time and location when describing the relevant beliefs. These should be held fixed throughout the discussion.

⁵It also worth noting that we are interested in agreement or disagreement as states, as opposed to the activity of agreeing or disagreeing. See e.g. Cappelen and Hawthorne (2009, 60–61) and MacFarlane (2014, 119–120) for relevant discussion.

⁶See e.g. Kölbel (2014, 102) for a similar observation.

One way to address this worry is to connect agreement and disagreement between mental states with agreement and disagreement between individuals. [Thesis 2] and [Thesis 3] allow us to do just that.

Thesis 2: Necessarily, for all individuals x and y , there is something that x agrees with y about if and only if there is a mental state σ_1 and a mental state σ_2 such that σ_1 agrees with σ_2 and x has σ_1 and y has σ_2 .

Thesis 3: Necessarily, for all individuals x and y , there is something that x disagrees with y about if and only if there is a mental state σ_1 and a mental state σ_2 such that σ_1 disagrees with σ_2 and x has σ_1 and y has σ_2 .

In other words, having mental states that agree is both necessary and sufficient for individuals to agree.⁷ The same holds for disagreement.⁸

This is not intended as a definition of agreement and disagreement between mental states in terms of agreement and disagreement between individuals.⁹ Instead, it is meant to provide a useful constraint on agreement and disagreement between mental states. For instance, let us suppose that it is possible that A has a mental state and B has a mental state, but A does not agree with B . In that case, A 's mental state does not agree with B 's mental state.

This constraint will be important throughout the following discussion. Even if it turns out that we do not have a firm grasp of agreement and disagreement between mental states, we can fall back on our grasp of agreement and disagreement between individuals. There are of course still going to be difficult cases.¹⁰ However, that does not prevent us from making claims about more straightforward cases.

⁷For the following discussion, 'A and B agree' should be taken to mean that there is something that A and B agree about. The same holds for disagreement.

⁸It is sometimes suggested that people can agree or disagree in virtue of the speech acts they perform, and not just in virtue of their mental states. See e.g. MacFarlane (2007, 2014, 22). However, I am going to follow Jackson and Pettit (1998, 251) in taking agreement and disagreement between individuals to be psychological phenomena. For instance, let us suppose that two individuals assert that it is raining, but only one of them is sincere. In that case, the person who is sincere can agree with what the insincere person asserted even if she does not agree with the insincere person. See e.g. Huvenes (2017, 273–274) for relevant discussion.

⁹[Thesis 2] and [Thesis 3] do not say anything about how judgments about agreement and disagreement are formed. While that is an empirical question, it would for instance be unsurprising if judgments about agreement and disagreement between individuals were to some extent based on judgments about the relationship between the mental states of the relevant individuals. However, these questions are orthogonal to the present discussion, which concerns the project of giving identity-conditions for mental states in terms of agreement and disagreement. [Thesis 1] is not a thesis about our judgments about agreement and disagreement. If the relevant judgments are reliable, that is enough.

¹⁰Let us suppose that while A believes it is raining, B believes that it is raining and $2+2=4$. Is that sufficient for A and B to agree? It is not immediately clear how we should answer that question. For instance, the relevant beliefs are true at the same possible worlds, but is that enough? That is not to say that there is no way of answering questions like this within the present framework. However, it means that some questions will require further discussion. Thanks to an anonymous referee for raising this issue.

3. Full and partial agreement

In the discussion that follows, it will be useful to know what we can and cannot assume about the agreement and disagreement relations. The goal is not to provide necessary and sufficient conditions for agreement and disagreement, but we can still say something about the formal properties of these relations. In section 4, we will see that even minimal assumptions about the formal properties of the agreement relation are sufficient to derive interesting conclusions. However, we will also see that some conclusions depend on stronger assumptions about the agreement relation.

While most of the discussion is going to focus on agreement, some brief remarks about disagreement are also in order. As far as disagreement is concerned, it seems clear that it is not reflexive.¹¹ It is not the case that every mental state disagrees with itself. It also seems clear that disagreement is not transitive. In other words, it can be that σ_1 disagrees with σ_2 and σ_2 disagrees with σ_3 , but σ_1 does not disagree with σ_3 . Let us suppose that two beliefs disagree with each other. In other words, σ_1 disagrees with σ_2 and σ_2 disagrees with σ_1 . That could easily be true. But if disagreement were transitive, then each belief would also disagree with itself.¹² That seems wrong.

It is more plausible that disagreement is symmetric. If there are cases in which symmetry fails, then such cases are not as obvious as for reflexivity and transitivity. For instance, if the belief that it is sunny disagrees with the belief that it is not sunny, then the belief that it is not sunny disagrees with the belief that it is sunny. That said, symmetry does not give us much by itself. As a result, the following discussion will focus on agreement.

Unlike disagreement, it seems that plausible that agreement is reflexive. If a mental state agrees with anything, it should be itself. The question it is whether agreement is also symmetric and transitive. In other words, is it an equivalence relation?

This has some important consequences. For instance, if agreement is an equivalence relation, then for all mental states σ_1 and σ_2 , σ_1 and σ_2 agree with all the same mental states if and only if σ_1 agrees with σ_2 . From that, it follows that if we can establish that some mental states

¹¹It is a further question whether disagreement is irreflexive. For instance, could one say that the belief that the liar sentence is true disagrees with itself? That would entail that there is disagreement between two individuals who believe that the liar sentence is true. That seems questionable, but a thorough discussion is beyond the scope of the present discussion.

¹²Transitivity does not require the values of the variables to be distinct. In effect, we are considering the special case where $\sigma_1 = \sigma_3$.

agree with each other, then that suffices to establish that they agree with all the same mental states. That is something that becomes relevant in section 4.2.

That said, it is by no means obvious that agreement is an equivalence relation. To see why, let us suppose that while *A* believes that it is cold and raining, *B* only believes that it is cold. Is that sufficient for *A* to agree with *B*? If it is, then [Thesis 2] says that they must have mental states that agree. But does the belief that it is cold and raining agree with the belief that it is cold? If agreement is an equivalence relation, that would mean that their beliefs agree with all the same mental states. However, as we will soon see, that cannot be right. If the belief that it is cold and raining agrees with the belief that it is cold, it presumably also agrees with the belief that it is raining. There is no obvious reason for saying that it agrees with one, but not the other. If agreement is an equivalence relation, it then follows that the belief that it is cold also agrees with the belief that it is raining. But that cannot be right if [Thesis 2] is correct. In that case, if *A* believes that it is raining and *B* believes that it is cold, that is sufficient for *A* to agree with *B*. But that is surely wrong. It may very well be that *A* believes that it is raining, and *B* believes that it is cold, but they do not agree.

There are two ways of responding to this line of reasoning. The first response, which is perhaps the most obvious, is to deny that agreement is an equivalence relation. In that case, transitivity appears to be the obvious culprit.¹³ If one accepts that the belief that it is cold and raining agrees with the belief that it is cold, then it is unclear what reason one would have to deny that the belief that it is cold agrees with the belief that it is cold and raining or that the belief that it is cold and raining agrees with the belief that it is raining. But then transitivity leads to the conclusion that the belief that it is cold agrees with the belief that it is raining. If the belief that it is cold agrees with the belief that it is cold and raining, and the belief that it is cold and raining agrees with the belief that it is raining, then the belief that it is cold agrees with the belief that it is raining. But that conclusion is clearly wrong.

¹³One could also deny that agreement is symmetric. That gives us two ways of resisting the problematic reasoning. The first is to say that the belief that it is cold and raining agrees with the belief that it is cold, but not vice versa. The second is to say that the belief that it is cold agrees with the belief that it is cold and raining, but not vice versa. However, it is unclear what would motivate one of these options over the other. Insofar as the first belief agrees with the second, it seems just as plausible that the second belief agrees with the first. Without such a motivation, I am going to set this alternative aside.

However, there is a second response. That is to say that the fault lies with the assumption that the belief that it is cold and raining agrees with the belief that it is cold or vice versa. It may seem plausible that *A* and *B* agree if *A* believes that it is cold and raining and *B* believes that it is cold. But that arguably depends on what further assumptions we make about their beliefs. The key question is whether *A* also believes that it is cold.

It is natural to assume that *A* also believes that it is cold. After all, someone who believes a simple conjunction typically believes each of the conjuncts. If that is the case, then it seems clear that *A* and *B* agree. But in this case, there is no need to say that *A*'s belief that it is cold and raining agrees with *B*'s belief that it is cold. From [Thesis 2], it follows that if *A* and *B* agree, they have mental states that agree. However, it does not have to be the belief that it is cold and raining that agrees with the belief that it is cold. In this case, *A* and *B* both believe that it is cold and that is presumably sufficient for them to have mental states that agree. Since we do not have to say that *A*'s belief that it is cold and raining agrees with *B*'s belief that it is cold, we can take agreement to be transitive without having to say that the belief that it is cold agrees with the belief that it is raining.

But what if *A* does not believe that it is cold?¹⁴ In that case, it is no longer as clear that *A* and *B* agree. For instance, it is unclear how we should describe their agreement. It seems plausible that *A* and *B* agree about a proposition only if they both believe it be true or both believe it to be false. But then what is it that they agree about? It is not that it is cold. *B* believes that, but *A* does not. It is not that it is cold and raining. *A* believes that, but *B* does not. Perhaps they can agree even if there is no proposition that they agree about. That would be relatively unsurprising if we are dealing with other mental states, but it would be interesting if it were true for agreement in belief as well. While these remarks are far from conclusive, we can see how someone could deny that *A* and *B* agree in such a case.

I am not going to attempt to determine which response is correct. Instead, I propose to distinguish between two relations: full agreement and partial agreement. While full agreement is an equivalence relation,

¹⁴It might be argued that it is impossible to believe a conjunction without believing the conjuncts. For instance, one might hold that belief is closed under entailment. See e.g. Stalnaker (1984, ch. 4) for relevant discussion. However, if that is true, that just means that we are left with the scenario in which *A* and *B* both believe that it is cold. As we have seen, in that scenario, there is no need to say that the belief that it is cold and raining agrees with the belief that it is cold to account for the agreement between *A* and *B*.

partial agreement is reflexive and symmetric, but not transitive. Moreover, partial agreement is necessary, but not sufficient, for full agreement. We can then say that there is partial agreement between the belief that it is cold and raining and the belief that it is cold, but not full agreement.

It is natural to suppose that there is a similar distinction to be drawn between full and partial disagreement. For instance, we might want to say that there is partial disagreement between the belief that it is not cold and the belief that it is cold and raining, but not full disagreement. However, since even full disagreement is neither reflexive nor transitive, full and partial disagreement cannot be distinguished based on their formal properties. That is another reason why it is easier to focus on agreement.

For the following discussion, I am going to leave it open whether we are talking about full or partial agreement when we talk about agreement. In section 4, we are going to see that it is possible to derive interesting conclusion from [Thesis 1] without taking a stand on that. But we will also see that some conclusions depend, at least in part, on whether we treat agreement as full or partial agreement.

4. De se beliefs

In this section, I will use the case of de se beliefs to illustrate some features of the present framework. Roughly speaking, de se beliefs are beliefs that one has about oneself and that one would typically express using a first-person pronoun. Such beliefs have received considerable attention in the philosophical literature. However, what matters here are the questions that de se beliefs raise regarding agreement and disagreement.

The focus will be on Lewis' (1979) influential account of de se beliefs.¹⁵ According to the account, the content of a de se belief is a property. Following Lewis, we can treat properties as sets of centred worlds, that is, world-individual pairs.¹⁶ For instance, if I believe that I am being attacked by a bear, the object of my belief can be represented as the content p such that $p = \{ \langle w, x \rangle : x \text{ is being attacked by a bear in } w \}$. Let us call such contents 'centred contents'.

It is well-known that Lewis' account of de se beliefs raises questions regarding agreement and disagreement. For instance, according to Ninan (2016), de se beliefs show that there cannot be a single notion of

¹⁵See e.g. Perry (1979) for another influential discussion of de se beliefs.

¹⁶Centred worlds are more commonly identified with world-time-individual triples. For the present discussion, the temporal dimension is ignored for the sake of simplicity.

content that can play a role in the explanation of action and the characterisation of agreement. De se beliefs are also used to argue against specific accounts of agreement and disagreement. For instance, MacFarlane (2014, 126), uses beliefs with centred contents to argue that noncoherence 'is not going to give us everything we might have wanted in a notion of disagreement'.¹⁷

In what follows, we are going to see that the present framework imposes certain constraints on an account of de se beliefs. The goal is neither to argue against Lewis' account nor is it to vindicate it. The discussion is mainly intended to show that it is possible to derive interesting conclusions from [Thesis 1] while relying on fairly minimal assumptions about agreement. I will do that by presenting two arguments. The first argument focuses on beliefs that have the same content according to Lewis' account. The second argument focuses on beliefs that have distinct contents according to Lewis' account.

4.1. The first argument

It will be useful to have an example to work with. Let us therefore suppose that *A* believes of herself that she is being attacked by a bear and that *B* also believes of herself that she is being attacked by a bear. In that case, their beliefs will have the same centred content according to Lewis' account. But it should be clear that this is not sufficient for *A* and *B* to agree. In fact, it is just as likely that *A* and *B* disagree. For instance, it could be that *A* believes that *B* is not being attacked by a bear or vice versa. In that case, it is plausible that *A* and *B* disagree.

The question is what the present framework says about *A*'s and *B*'s beliefs. It is useful to separate two questions. The first question is whether *A*'s belief that she is being attacked by a bear is the same belief as *B*'s belief that she is being attacked by a bear. That is a question about the individuation of the relevant mental states. The second question is whether *A*'s belief and *B*'s belief have the same content. That is a question about the content of the relevant mental states. While it might seem that it is the second question that is most critical for the evaluation of Lewis' account, it will be convenient to start with the first question.

¹⁷See e.g. Egan (2014), Kölbel (2014), Marques (2014), and Kindermann (2019) for further discussion connecting de se beliefs with agreement and disagreement.

Let us use ' β_A ' for A 's belief that she is being attacked by a bear and ' β_B ' for B 's belief that she is being attacked by a bear. In that case, the following argument can be used to argue that β_A and β_B are distinct mental states.

Premise 1: If $\beta_A = \beta_B$, then β_A and β_B agree with all the same mental states.

Premise 2: If β_A and β_B agree with all the same mental states, then β_A agrees with β_B .

Premise 3: If β_A agrees with β_B , then A having β_A and B having β_B is sufficient for A to agree with B .

Premise 4: A having β_A and B having β_B is not sufficient for A to agree with B .

Conclusion: $\beta_A \neq \beta_B$.

The first premise follows from [Thesis 1]. More specifically, it follows from the left-to-right direction. That is of course something that can be rejected. But it would be a departure from the present framework. I will therefore focus on the other premises.

The second premise only relies on the assumption that agreement is reflexive. If agreement is reflexive, then β_B agrees with itself. If β_A and β_B agree with all the same states, it then follows that β_A also agrees with β_B . That means that one cannot resist the argument just by denying that agreement is an equivalence relation. In other words, the premise holds for both full agreement and partial agreement. Given the plausibility of taking agreement to be reflexive, rejecting this premise seems unappealing.

The third premise follows from [Thesis 2]. One could deny that by saying that β_A agrees with β_B , but that it is not a kind of agreement that makes it the case that A agrees with B . However, in the absence of [Thesis 2], we again face the worry that we are extending the notion of agreement too far by talking about agreement between mental states. Furthermore, even if we set aside the question of whether A and B agree, it is unclear whether there is any independent motivation for thinking that β_A and β_B agree apart from the fact that they have the same content according to Lewis' account.

The fourth premise is simply based on the initial observations about the example. As already noted, the fact that A believes of herself that she is being attacked by a bear and B believes of herself that she is being attacked by a bear, does not make it the case that A agrees with B . That is of course not to deny that A could agree with B about something else. However, that is beside the point. The point is that if β_A agrees with β_B , then that should be sufficient for A to agree with B . But it is not.

The upshot is that there is limited scope for resisting the argument within the present framework. However, even if the argument is sound,

it only establishes that the beliefs are distinct. As a thesis about the individuation of mental states, [Thesis 1] does not by itself imply anything about the contents of the relevant beliefs. To make a claim about the content, we must assume that beliefs are individuated in terms of their contents. Let us therefore assume, at least for the time being, that sameness of content is sufficient for identity in the case of beliefs. In other words, for all beliefs β_1 and β_2 , if β_1 and β_2 have the same content, then $\beta_1 = \beta_2$. In that case, it follows from the conclusion that β_A and β_B are distinct, that they also have distinct contents.

This conclusion is incompatible with an account, like that of Lewis, that assigns the same content to *A*'s and *B*'s beliefs. However, it is possible to question the assumption that sameness of content is sufficient for identity. Perhaps the most obvious way to do that, is to say that beliefs have multiple contents. That is for instance similar to a line that Ninan (2016) takes in response to a related line of argument.¹⁸ Simplifying somewhat, in addition to their centred contents, *A*'s belief also has the non-centred content q such that $q = \{w: A$ is being attacked by a bear in $w\}$ and *B*'s belief also has the non-centred content r such that $r = \{w : B$ is being attacked by a bear in $w\}$.¹⁹ Beliefs can then be individuated in terms of their non-centred contents. The assumption that sameness of content is sufficient for identity holds for non-centred contents, but not for centred contents.

There are also other ways of questioning the assumption. For instance, one could say that beliefs are partly individuated by what they are about while denying that aboutness is determined by content.²⁰ In that case, beliefs with the same content may be distinct in virtue of being about distinct individuals. One could then say that *A*'s belief that she is being attacked by a bear is about *A* while *B*'s belief that she is being attacked by a bear is about *B* even if the beliefs have same content.

For the present discussion, it does not matter how, or even whether, one rejects the assumption that sameness of content is sufficient for identity. That assumption goes beyond [Thesis 1] and the present framework. This illustrates an important point. Within the present framework, the issue is not primarily the content of the relevant beliefs. The primary

¹⁸See e.g. Recanati (2007) and Moss (2012) for discussion of related views.

¹⁹Ninan (2016, 115–116) makes a similar point by drawing on Egan's (2006, 107) distinction between boring and interesting centred contents. An interesting centred content is a centred content that varies in truth-value across individuals in the same world. A boring centred content is a centred content that is not interesting.

²⁰See e.g. Sandgren (forthcoming) for further discussion of the view that aboutness is not determined by content.

issue is the individuation of the relevant beliefs. This makes the present discussion different from discussions that start out by making assumptions about content. For instance, we have not assumed that agreement consists in having beliefs with the same content.²¹ The preceding discussion shows that such assumptions are not necessary to derive interesting conclusions.

The preceding discussion also serves as an illustration of the differences between the present framework and other ways of individuating beliefs. One way to motivate Lewis' account of *de se* attitudes, is based on the observation that *A*'s and *B*'s beliefs appear to have the same causal role.

Suppose, for example, that you and I are both in a belief state that we could each express by saying, "I am being attacked by a bear" and that both of us believe that the best way to avert a bear attack is to curl up into a ball and remain still. Then we will, other things being equal, both be disposed to curl up and remain still. [...] Since your belief that you will be attacked by a bear and my belief that I will be so attacked have the same causal role in our respective cognitive systems, they ought to be assigned the same content. (Ninan 2013, 302)

Ninan uses this to motivate the claim that *A*'s and *B*'s beliefs have the same content, but one could also use it to motivate the claim that the beliefs are identical. Let us suppose that for all beliefs β_1 and β_2 , if β_1 and β_2 have the same causal role, then $\beta_1 = \beta_2$. In other words, having the same causal role is sufficient for identity.²² If we assume that *A*'s and *B*'s beliefs have the same causal role, as Ninan argues, then the beliefs are identical. However, this contradicts the conclusion of the argument above. Unless there is something wrong with the other premises, that means that beliefs cannot both be individuated in terms of causal role and agreement.²³

A thorough discussion of other ways of individuating beliefs, is beyond the scope of the present discussion. That said, even if there is a conflict, it is not obvious that the fault lies with [Thesis 1]. We should consider the possibility that inter-personal relations like agreement and disagreement play a more significant role in the individuation of beliefs than has been commonly recognised. It is also worth noting that even if that is true, beliefs can still be partly individuated in terms of their causal roles. For all that has been said so far, the present framework is compatible with the idea

²¹See e.g. Cappelen and Hawthorne (2009, 63) and Ninan (2016, 100) for the assumption that agreement consists in having beliefs with the same content.

²²See e.g. Armstrong (1968) and Lewis (1972) for influential discussions of the view that mental states are individuated in terms of causal roles.

²³See e.g. Ninan (2016) for a related argument to the effect that a single notion cannot play a role in both explaining action and characterizing agreement.

that having the same causal role is necessary for identity. In other words, for all beliefs β_1 and β_2 , if $\beta_1 = \beta_2$, then β_1 and β_2 have the same causal role. It is just that distinct beliefs may also have the same causal role.

4.2 The second argument

So far, everything that has been said is also compatible with sameness of content being necessary for identity. In other words, for all beliefs β_1 and β_2 , if β_1 and β_2 have the same content, then $\beta_1 = \beta_2$. Is it possible to construct a similar argument that also calls that into question? That is not obvious. It turns out that such an argument requires somewhat stronger assumptions than the previous argument, but it is nevertheless instructive to see what those assumptions are.

Let us modify the previous example. As before, B believes of herself that she is being attacked by a bear. But unlike the previous case, A does not have a belief about herself. Instead, she believes that B is being attacked by a bear. In that case, it seems that A and B agree. However, according to Lewis' account, their beliefs do not have the same content. While the content of B 's belief is still a centred content, the content of A 's belief can be represented as the non-centred content r such that $r = \{w : B \text{ is being attacked by a bear in } w\}$. These contents are clearly distinct. The former can vary in truth-value across individuals in the same world, but that is not true for the latter.

Let us use ' β_A ' for A 's belief that B is being attacked by a bear and ' β_B ' for B 's belief that she is being attacked by a bear. In that case, the following argument purports to demonstrate that β_A and β_B are identical.

Premise 1: If β_A and β_B agree with all the same mental states, then $\beta_A = \beta_B$.

Premise 2: If β_A agrees with β_B , then β_A and β_B agree with all the same mental states.

Premise 3: If A agrees with B , then β_A agrees with β_B .

Premise 4: A agrees with B .

Conclusion: $\beta_A = \beta_B$.

If the argument succeeds in establishing that the beliefs are identical, then someone who wants to maintain that the beliefs have the same content, must also deny that sameness of content is necessary for identity. The question is whether the argument is successful.

As before, there is not much to say about the fourth premise, which is supported by the observations about the example. The other premises require a bit more discussion.

The first premise follows from [Thesis 1] along with an additional assumption. The right-to-left direction of [Thesis 1] entails that if β_A and β_B agree and disagree with all the same states, then $\beta_A = \beta_B$. However, as stated, the argument only deals with agreement. To derive the first premise, we need to assume that if β_A and β_B agree with all the same states, then β_A and β_B also disagree with all the same states. That seems plausible. It would be surprising if two mental states agreed with all the same mental states, but one of them disagreed with a mental that the other did not disagree with. But even if it is plausible, it introduces another assumption that was not needed for the previous argument.

The third premise can be motivated by [Thesis 2]. According to [Thesis 2], *A* and *B* agree only if they have mental states that agree. That means that if β_A does not agree with β_A , then *A* and *B* must have some other mental states that agree. However, apart from β_A and β_B , we have not made any assumptions about the mental states of *A* and *B*. In the absence of other candidates, it seems reasonable to assume that β_A agrees with β_B .

The most serious question concerns the second premise. While the second premise of the previous argument only relied on the assumption that agreement is reflexive, this premise relies on the stronger assumption that agreement is an equivalence relation. In other words, it requires us to treat agreement as full agreement, as opposed to partial agreement. As we have seen, that is a substantial assumption. It opens another way of resisting the argument by arguing that we should treat agreement as partial agreement.

This shows that this argument relies on significantly stronger assumptions than the previous argument. It also shows that the choice between full and partial agreement has consequences for the implications of the present framework. It is possible to argue for a distinctness claim without taking a stand on that issue, but it becomes relevant if we want to argue for an identity claim. Since it is unclear how the choice between full and partial agreement should be settled, it remains an open question whether the second argument is successful.

5. Other mental states

In this section, the goal is to say something about how far [Thesis 1] can be extended. While some mental states, like beliefs, stand in relations of agreement and disagreement, that does not seem to be true of all mental states. That raises the question of which mental states [Thesis 1]

holds for. It is beyond the scope of the present discussion to come up with a comprehensive list or necessary and sufficient conditions. That would require a discussion of a wide range of mental states. Instead, the goal is to focus on some illustrative examples.

So far, the focus has been on beliefs. However, [Thesis 1] does not have to be restricted to beliefs. Following Stevenson (1944, 1963), it is common to distinguish between disagreement in belief and disagreement in attitude. While disagreement in belief involves a conflict of beliefs, disagreement in attitude involves a conflict of other mental states, such as desires or preferences.²⁴ A similar distinction can be drawn in the case of agreement.

For instance, following Stevenson (1963, 2), let us suppose that *A* and *B* have decided to have dinner together. While *A* wants them to go to a certain restaurant, *B* does not want them to go to that restaurant. As Stevenson observes, this appears to be a case of disagreement. However, insofar as *A* and *B* do not appear to have any beliefs that disagree, it seems plausible to conclude that it is their desires that disagree. To obtain a case of agreement, we only need to modify the example such that *A* and *B* both want to go the same restaurant. In that case, it appears that *A* agrees with *B*. In the absence of any beliefs that agree, it again seems plausible that their desires agree.

In these cases, the agreement and the disagreement have a practical dimension insofar as *A* and *B* are trying to decide on a joint course of action. If that were necessary for desires to agree or disagree, that could prevent us from extending [Thesis 1] to desires across the board. However, this is arguably an accidental feature of these examples. Following Ninan (2016, 100) if *A* and *B* want Elizabeth Warren to run for President, it seems plausible that they agree. In this case, the agreement does not have to have any practical dimension. There is no reason to suppose that *A* and *B* are trying to coordinate their actions in any way. In fact, they could be unrelated individuals who have no influence on whether Elizabeth Warren runs for President. Furthermore, we have still not assumed anything about their beliefs. It may be that they also have beliefs that agree, but that does not seem to be necessary for them to

²⁴The idea that there can be disagreement involving other mental states than beliefs, has become widely accepted. See e.g. Blackburn (1998), Gibbard (2003), Dreier (2009), Björnsson and Finlay (2010), Sundell (2011), Huvenes (2012), MacFarlane (2014), Marques and García-Carpintero (2014), McKenna (2014), Marques (2015), and Richard (2015) for relevant discussion.

agree.²⁵ It seems sufficient that they both want Elizabeth Warren to run for President.²⁶

Insofar as desires can stand in relations of agreement and disagreement, that suggests that [Thesis 1] may be extended to desires. Moreover, desires provide just one example of mental states that stand in relations of agreement and disagreement. For instance, in addition to desires, Stevenson (1944, 3) takes the notion of disagreement in attitude to cover purposes, aspirations, and preferences.

However, there are mental states that do not appear to stand in relations of agreement and disagreement. For instance, if *A* and *B* are both in pain, there is no sense in which *A*'s pain agrees or disagrees with *B*'s pain. There is also no sense in which we would want to say that having this pain is sufficient for *A* and *B* to agree or disagree. That is true regardless of how similar or different the feeling of pain is. The upshot is that [Thesis 1] cannot plausibly be extended to mental states like pain.

It is perhaps unsurprising that [Thesis 1] cannot be extended to mental states like pain. Such mental states are commonly individuated in terms of their phenomenology, what it is like to be in the state. However, there are other mental states that do not seem to stand in relations of agreement and disagreement. That includes mental states that are commonly taken to have content. For instance, it is at best unclear whether there are any case of agreement involving imagination. Let us suppose that both *A* and *B* imagine that it is sunny. That does not appear to be sufficient for them to agree. If we assume [Thesis 2], that also means that their mental states do not agree.

A similar point can be made regarding perception. For instance, let us suppose that both *A* and *B* have a visual experience as of it being sunny. Again, that does not appear to be sufficient for them to agree. To see that, we need to make sure that *A* and *B* do not have any beliefs that agree. Let us therefore suppose that while *A* takes her experience to be veridical and forms the belief that it is sunny, *B* takes herself to be hallucinating and believes that it is not sunny. In that case, it does not appear that *A* and *B* agree, even though they both have a visual experience as of it being sunny.

²⁵It is for instance possible that *A* and *B* both believe that Elizabeth Warren ought to run for President. However, that need not be the case. Someone can want something to happen even if they do not believe that it ought to happen. See e.g. Huvenes (2017, 276) for relevant discussion.

²⁶See e.g. Huvenes (2017, 276) for a similar example involving disagreement. See also Marques (2015) for relevant discussion about what it takes for desires to disagree.

It is worth noting that the previous remarks assume [Thesis 2]. If one is willing to reject [Thesis 2], matters become more complicated. If we set aside the question of whether *A* agrees with *B*, it does not seem altogether unnatural to say that *A*'s visual experience agrees with *B*'s visual experience. Perhaps one could even say that *A*'s imagining agrees with *B*'s imagining. But if that is true, it would have to be a kind of agreement that can obtain between mental states without making it the case that the bearers of those mental states agree. That again raises the worry that we are extending the notion of agreement too far.

Where does this leave us? The tentative conclusion is that there are good prospects of extending [Thesis 1] to desires. However, it is also clear that it cannot be extended to all mental states, as demonstrated by mental states like pain. It is also questionable whether it can be extended to imaginings or perceptual experiences, especially if we assume [Thesis 2].

6. Concluding remarks

Let us take stock. It is common to explain agreement and disagreement in terms of relations among mental states. The main purpose of the discussion has been to present an alternative way of thinking about the relationship between mental states and agreement and disagreement. The idea is to use agreement and disagreement to give identity-conditions for at least certain kinds of mental states. That is what [Thesis 1] does.

We have seen that [Thesis 1] allows us to derive interesting conclusions about *de se* beliefs. Furthermore, it does that while relying on assumptions that are interestingly different from the standard assumptions in the literature. There is for instance no need to assume that agreement consists in believing the same content. Instead, we can make do with fairly minimal assumptions about the formal properties of the agreement relation.

An important question is how far this framework can be extended. Most of the discussion has focused on beliefs. However, we have seen that there are other mental states, such as desires, that also seem to stand in relations of agreement and disagreement. That suggests that there are good prospects of extending the framework to those mental states. That said, there are other mental states, such as imaginings and perceptual experiences, that raise difficult questions. This is something that merits further discussion.

Looking ahead, more can also be said about how the present framework relates to other ways of individuating mental states. There is for instance much to say about how the framework relates to accounts that individuate mental states in terms of their causal roles. While we briefly touched on that as a part of the discussion of *de se* beliefs, a more comprehensive discussion is beyond the scope of the present discussion.

Acknowledgements

Earlier versions of this paper were presented at the Higher Seminar in Philosophy at Umeå University in February 2021 and the SAD seminar in March 2021. I am grateful to the participants for helpful discussion and comments. I would also like to thank an anonymous reviewer for helpful comments. Research for this paper was conducted at Umeå University and supported by Torfinn Huvenes' Vetenskapsrådet Research Project Grant (2017-03222).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Vetenskapsrådet [grant number 2017-03222].

ORCID

Torfinn Thomesen Huvenes  <http://orcid.org/0000-0001-7368-2397>

References

- Armstrong, D. 1968. *A Materialist Theory of the Mind*. London: Routledge.
- Björnsson, G., and S. Finlay. 2010. "Metaethical Contextualism Defended." *Ethics* 121: 7–36.
- Blackburn, S. 1998. *Ruling Passions*. Oxford: Oxford University Press.
- Cappelen, H., and J. Hawthorne. 2009. *Relativism and Monadic Truth*. Oxford: Oxford University Press.
- Dreier, J. 2009. "Relativism (and Expressivism) and the Problem of Disagreement." *Philosophical Perspectives* 23: 79–110.
- Egan, A. 2006. "Secondary Qualities and Self-Location." *Philosophy and Phenomenological Research* 72: 97–119.
- Egan, A. 2014. "There's Something Funny about Comedy: A Case Study in Faultless Disagreement." *Erkenntnis* 79: 73–100.

- Gibbard, A. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- Huvenes, T. 2012. "Varieties of Disagreement and Predicates of Taste." *Australasian Journal of Philosophy* 90: 167–181.
- Huvenes, T. 2017. "On Disagreement." In *The Routledge Handbook of Epistemic Contextualism*, edited by J. Ichikawa, 272–281. London: Routledge.
- Jackson, F. 2008. "The Argument from the Persistence of Moral Disagreement." In *Oxford Studies in Metaethics: Volume 3*, edited by R. Shafer-Landau, 75–86. Oxford: Oxford University Press.
- Jackson, F., and P. Pettit. 1998. "A Problem for Expressivism." *Analysis* 58: 239–251.
- Kindermann, D. 2019. "Coordinating Perspectives: De Se and Taste Attitude in Communication." *Inquiry* 62: 912–955.
- Kölbel, M. 2014. "Agreement and Communication." *Erkenntnis* 79: 101–120.
- Lewis, D. 1972. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50: 249–258.
- Lewis, D. 1979. "Attitudes De Dicto and De Se." *The Philosophical Review* 88: 513–543.
- MacFarlane, J. 2007. "Relativism and Disagreement." *Philosophical Studies* 132: 17–31.
- MacFarlane, J. 2014. *Assessment Sensitivity: Relative Truth and its Applications*. Oxford: Oxford University Press.
- Marques, T. 2014. "Doxastic Disagreement." *Erkenntnis* 79: 121–142.
- Marques, T. 2015. "Disagreeing in Context." *Frontiers in Psychology* 6: 1–12.
- Marques, T., and G. García-Carpintero. 2014. "Disagreement about Taste: Commonality Presuppositions and Coordination." *Australasian Journal of Philosophy* 92: 701–723.
- McKenna, R. 2014. "Shifting Targets and Disagreements." *Australasian Journal of Philosophy* 92: 725–742.
- Moss, S. 2012. "Updating as Communication." *Philosophy and Phenomenological Research* 85: 225–248.
- Ninan, D. 2013. "Self-Location and Other-Location." *Philosophy and Phenomenological Research* 87: 301–331.
- Ninan, D. 2016. "What is the Problem of De Se Attitudes?" In *About Oneself: De Se Attitudes and Communication*, edited by M. García-Carpintero, and S. Torre, 86–120. Oxford: Oxford University Press.
- Perry, J. 1979. "The Problem of the Essential Indexical." *Noûs* 13: 3–21.
- Recanati, F. 2007. *Perspectival Thought*. Oxford: Oxford University Press.
- Richard, M. 2015. "What is Disagreement?" In *Truth and Truth Bearers*, 82–114. Oxford: Oxford University Press.
- Sandgren, A. forthcoming. "Turning Aboutness About." *Pacific Philosophical Quarterly*.
- Stalnaker, R. 1984. *Inquiry*. Cambridge, MA: MIT Press.
- Stevenson, C. L. 1944. *Ethics and Language*. New Haven: Yale University Press.
- Stevenson, C. L. 1963. "The Nature of Ethical Disagreement." In *Facts and Values*, 1–9. New Haven: Yale University Press.
- Sundell, T. 2011. "Disagreements about Taste." *Philosophical Studies* 155: 267–288.
- Zeman, D. 2020. "Minimal Disagreement." *Philosophia* 48: 1649–1670.