

This is the final peer-reviewed accepted manuscript of:

'Interworld Disagreement', *Erkenntnis*, Online First, 2019.

The final published version is available online at:

<https://doi.org/10.1007/s10670-019-00171-w>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Interworld Disagreement

Abstract. Disagreement plays an important role in several philosophical debates, with intuitions about ordinary or exotic cases of agreement and disagreement being invoked to support or undermine competing semantic, epistemological and metaphysical views. In this paper we discuss cases of (alleged) interworld doxastic disagreement, that is to say, cases of doxastic disagreement supposedly obtaining between (the beliefs of) individuals inhabiting different possible worlds, in particular between an individual inhabiting the actual world and his/her counterpart in another possible world. We draw a distinction between propositional and attitudinal disagreement, bring it to bear on the issue of the conditions of this kind of disagreement, and raise some metaphysical and epistemological worries about the claim that an individual inhabiting the actual world can disagree with an attitude or a speech act of his/her own counterpart, or of another individual, in a different possible world.

1. Introduction

In this paper we discuss ‘interworld’ doxastic disagreement, that is to say, the kind of disagreement that some philosophers allege can obtain between the beliefs of individuals inhabiting different possible worlds. Two caveats before we begin. First, our discussion deliberately ignores nondoxastic disagreement, that is to say, disagreement concerning nondoxastic attitudes like desires or likings. Second, the focus of our discussion is exclusively on the *state* of disagreement, which is something that can obtain even between individuals that do not know of each other, as opposed to the *activity* of disagreeing, which constitutively depends on the attitudes individuals have towards each other and requires some form of

interaction between them (Cappelen & Hawthorne 2009, pp. 60-61; MacFarlane 2014, pp. 119-120). We make these two caveats because we will often speak, for the sake of brevity, of two agents' 'disagreeing' with each other rather than of their 'being in a state of doxastic disagreement', and such phrases should always be construed in a doxastic, state sense.

Disagreement plays an important role in several philosophical debates. The existence of deep and apparently irresolvable forms of disagreement about moral and religious matters, or among competing comprehensive worldviews (*Weltanschauungen*), has long been offered as evidence of an alleged lack of objectivity and/or relativity of wide segments of moral, religious, and philosophical discourse. While this theme has enjoyed a long and distinguished philosophical history (from Montaigne to Mackie, to pragmatists and postmodernists of many stripes), the topic of disagreement has recently occupied centre stage also in more specialised debates in epistemology and semantics.

One issue is the epistemological problem that the phenomenon of disagreement among epistemic peers raises for the justification of our beliefs (Frances & Matheson 2018, sect. 5). How should the discovery that one of our epistemic peers – someone who possesses exactly the same evidence that we possess and who is just as good as we are in processing that evidence – rejects a propositions that we accept (or vice versa) affect our doxastic attitude towards that proposition?

Another issue concerns the choice of the correct semantic theory for certain types of discourse. Intuitions about agreement and disagreement are frequently invoked by critics of contextualist treatments of predicates of personal taste, epistemic modals, future contingents and other linguistic constructions. This is because agreement and disagreement intuitions can be used to support or undermine 'the view that there is stability of content across a variety of contexts where the contextualist is committed to thinking that content varies' (Cappelen & Hawthorne, p. 25). Thus, the charge that contextualist treatments fail to account for the fact

that disputes between speakers engaging in certain types of discourse may amount to genuine cases of disagreement has been argued to warrant a preference for a relativist semantic of such types of discourse (Egan, Hawthorne & Weatherson 2005; Lasersohn 2005; Egan 2007; Stephenson 2007). Relatedly, the idea that disputes of taste involve a form of *faultless* disagreement has been used to make a case for truth relativism, on the assumption that only truth relativism can account for the possibility that two parties may disagree without anyone's fault (Kölbel 2002; 2004).

Quite interestingly, the intuitions invoked in the debate between contextualists and relativists have come to concern not only ordinary disputes between individuals inhabiting the same possible world, but also more exotic disputes between individuals inhabiting different possible worlds. Thus, intuitions about situations presented as cases of interworld disagreement have been used to construct further arguments in support of a relativist treatment of certain kinds of discourse, arguments that depend on the idea that there is a *bona fide* sense in which individuals inhabiting different possible worlds can disagree with each other (MacFarlane 2009; 2014, chap. 6).

If clarifying the nature and conditions of disagreement in general can shed light on many traditional and recent philosophical debates, getting clear about the conditions, and indeed the very possibility, of interworld disagreement is then crucial for assessing the force of some of the arguments that are currently employed in discussions of contextualist and relativist semantics.

We make two points in this paper. First, we argue that *if disagreeing across possible worlds is possible*, then it is easier than some philosophers have assumed, but not so easy as others have maintained. In sects. 1 and 2 we discuss the main options on the table and emphasize the importance of distinguishing two different senses in which philosophers are apt to describe a situation as a state of doxastic disagreement. Indeed, we argue that conflicting

views on the conditions for being in such a state fall out of different construals of the notion of doxastic disagreement, namely, a propositional and an attitudinal one. When each construal is given its due the conflict disappears, and one is tempted to conclude that different views on the conditions for disagreement can peacefully coexist. However, it should not be taken for granted that disagreeing across possible worlds is possible in the first place, which is why we express the upshot of our discussion in sects. 1 and 2 by means of a conditional: *if disagreeing across possible worlds is possible*, then it is easier than some philosophers have assumed, but not so easy as others have maintained, depending on the kind of disagreement (propositional or attitudinal) that is being considered. Indeed, in sect. 3 we go on to maintain that there is in fact no such thing as interworld doxastic disagreement. More precisely, we argue that there are good metaphysical and epistemological reasons to resist the temptation to describe as genuine cases of disagreement the situations in which the doxastic attitudes or speech acts of individuals inhabiting different possible worlds cannot both be accurate, and we suggest that this casts an unfavourable light also on the possibility of propositional interworld disagreement. Needless to say, we believe that these reasons are fairly compelling. But it is worth noting that the overall structure of our argument leaves it open for readers unconvinced by the reasons offered in sect. 3 to accept the conditional conclusion of sects. 1 and 2 (and indeed to detach its consequent), as well as for readers sceptical of the conditional conclusion reached by the end of sect. 2 to endorse the moral of sect. 3.

2. The case of Jane and June

Recent discussions about doxastic disagreement frequently consider cases of disagreement involving not only inhabitants of different worlds, but an inhabitant of the actual world and his/her counterpart in another possible world. One famous case is due to John MacFarlane:

Consider Jane (who inhabits this world, the actual world) and June, her counterpart in another possible world. Jane asserts that Mars has two moons, and June denies this very proposition. Do they disagree? Not in any real way. Jane's assertion concerns our world, while June concerns her. If June lives in a world where Mars has three moons, her denial may be just as correct as Jane's assertion. (MacFarlane 2007, p. 23)

The moral that MacFarlane draws from this and similar cases is that having incompatible (doxastic) attitudes – typically, acceptance and rejection – towards one proposition is sufficient for disagreeing ‘only when the acceptance and rejection take place in the same world’ (MacFarlane 2007, p. 23). When this is not the case, for two subjects to disagree it is necessary (and sufficient) that their respective attitudes *cannot both be accurate* (MacFarlane 2007, p. 24), that is to say, that the accuracy of either of them precludes the accuracy of the other (MacFarlane 2009, pp. 9-13; 2014, pp. 125-128).

The principle suggested by MacFarlane does justice to the intuition that, while Jane and June have incompatible attitudes towards the proposition that Mars has two moons, they aren't in a state of doxastic disagreement: their assertions concern different worlds, so the accuracy of either of them does not preclude the accuracy of the other. However, MacFarlane does not maintain that having beliefs or making assertions that concern different worlds, or, more generally, different circumstances of evaluation, is *sufficient* for not disagreeing. To find suggestions of such a tenor one must look elsewhere. Thus, François Recanati maintains that agreement and disagreement always involve complete contents or *incomplete contents that concern the same circumstance of evaluation* (Recanati 2008, pp. 56-57). Ragnar Francén is even more explicit:

To really disagree, [...] [Jane and June] would have to be in the same circumstance of evaluation, or at least intend their judgments to hold for the same circumstance of evaluation. In MacFarlane's terms, the assertions would have to *concern* the same circumstance of evaluation. (Francén 2010, p. 25)

If Francén is right, disagreeing is a rather difficult affair: to avoid being in a state of doxastic disagreement, it is sufficient to have beliefs or make assertions that concern different circumstances of evaluation – which means that, in many cases, it is sufficient to inhabit different possible worlds. But is this really so? Here is a tempting argument for the opposite conclusion (to avoid misunderstanding, let it be clear from the start that it is *not* an argument that we wish to endorse).

Darius asserts the proposition that Port Stanley is in the Falklands, and Daria accepts it; so they do not disagree about that proposition. But suppose that Darius' counterpart in M_1 – call him Marius – asserts, and that Daria's counterpart in M_1 – call her Maria – rejects, the proposition that Port Stanley is in the Falklands. It seems clear that Marius and Maria disagree on that proposition: having incompatible attitudes towards one and the same proposition is sufficient for disagreeing *if what is said or believed concerns the same circumstance of evaluation*. Do Daria and Maria disagree about the proposition that Port Stanley is in the Falklands? If we admit, as it seems inevitable to admit, that Marius and Darius do not disagree, we might be tempted to answer this question in the affirmative. For assume for *reductio* that Daria and Maria do *not* disagree. Now, absence of disagreement is obviously symmetric (if A fails to disagree with B about proposition P, then, necessarily, B will fail to disagree with A about proposition P), and apparently also transitive (if A fails to disagree with B about proposition P and B fails to disagree with C about proposition P, then, necessarily, A will fail to disagree with C about proposition P). So, if Marius and Darius, Darius and Daria,

and Daria and Maria do not disagree about the proposition that Port Stanley is in the Falklands, neither will Marius and Maria. Contradiction! Conclusion: Daria and Maria disagree.

This piece of reasoning is actually far from compelling. Since the proposition concerned is straightforwardly factual, we would defend the argument's (anti-relativist) presupposition, which is in fact shared by all parties to the debate, that its protagonists aren't 'normatively insulated' from one another in the sense of Rovane (2013).¹ But the impression that the argument is sound crucially depends on the tacit assumption that M_1 isn't that distant from the actual world: if in M_1 Port Stanley were *not* in the Falklands, there would be nothing for Maria and Daria to disagree about. The assumption that M_1 isn't that distant from the actual world is crucial because the claim that absence of disagreement is transitive is actually *false*: absence of disagreement is transitive only across *relevantly similar* circumstances of evaluation – and a world in which Port Stanley is not in the Falklands is a world which isn't sufficiently similar, in the relevant respects, to the actual world.

Even so, the case of Daria and Maria does suggest that *under certain conditions* it is possible to disagree even if what is believed or said does not concern the same circumstance of evaluation. The moral to be drawn from the case thus seems to be that disagreeing is significantly easier than MacFarlane's case has suggested to some interpreters: inhabiting different worlds – more precisely, making assertions or having beliefs that concern different circumstances of evaluation – is *not* sufficient for failing to disagree. For there to be a disagreement, the beliefs or assertions of the relevant subjects need not necessarily concern one and the same circumstance of evaluation; rather, they must concern circumstances of evaluation that are *sufficiently similar in all relevant respects*. This admittedly vague formulation may be made more precise by invoking a theory of subject matters.² One such theory has been recently proposed by Stephen Yablo, who takes a leaf from David Lewis'

book and treats subject matters as partitions – i.e., as decompositions into mutually disjoint subsets – of the set of worlds (Yablo 2014, pp. 26-28; see Lewis 1988, pp. 11-14). One might then say that, for there to be a disagreement, the relevant beliefs or assertions *must concern* (in Yablo’s terms, *be about*) *the same partition of the set of worlds* – which is the case with Daria and Maria if M_1 is reasonably close to the actual world, but not with Jane and June, as the former’s assertion concerns the set of possible worlds in which Mars has as many moons as in Jane’s world, while the latter’s assertion concerns the set of possible worlds in which Mars has as many moons as in June’s world.

3. *Propositional vs attitudinal disagreement*

Having beliefs or making assertions that concern different circumstances of evaluation is thus insufficient for failing to disagree. But Herman Cappelen and John Hawthorne have claimed that situations of interworld disagreement are actually more frequent than MacFarlane’s case appears to suggest, even when it isn’t taken to support the radical conclusions that have just been rejected. Their view is indeed that the case of Jane and June *is* a case of doxastic disagreement (Cappelen and Hawthorne 2009, pp. 64-66).

Looking with suspicion at MacFarlane’s realist talk of possible worlds and counterparts, Cappelen and Hawthorne redescribe his case in plain subjunctive terms as a situation in which Jane actually says *that Mars has two moons*, but would have said *that it is not the case that Mars has two moons* if it had been the case that Mars did not have two moons. They then go on to argue that states of interworld disagreement are just situations in which someone would have rejected (or accepted a proposition incompatible with) the *proposition* that they, or someone else, actually asserted or believed.³ And they conclude that MacFarlane’s case is a clear example of such a situation, for if Jane were to deny the proposition that Mars has two moons, she would reject the proposition that she accepts in the

actual world, finding herself in blatant disagreement with what she actually said (Cappelen and Hawthorne 2009, p. 65).

The plausibility of Cappelen and Hawthorne's conclusion clearly depends on their rejection of the view that (interworld) doxastic disagreement is disagreement with an agent's (doxastic) *attitude* or *speech act*. And they are certainly right to emphasize that "the claim that each of two individuals in different worlds accepts some proposition P is not akin to the claim that two individuals in different countries accept that proposition" (Cappelen and Hawthorne 2009, p. 64). Indeed, much of what we shall have to say in sect. 3 about the possibility of interworld disagreement is an elaboration on this crucial point. However, the point is not in itself sufficient reason to endorse the conclusion that doxastic disagreement is *always* disagreement with a proposition – or the attendant view that disagreeing with someone's belief or assertion that *p* boils down to having beliefs whose contents are jointly incompatible with the proposition that *p*. For disagreement with an agent's attitude or speech act resists being reduced to disagreement with a proposition even when the focus is exclusively on *intraworld* disagreement.⁴ The main argument against the 'Simple View' of disagreement (as MacFarlane 2014, p. 121 calls it) is that it gives the wrong verdicts when applied to situations featuring beliefs with centred or tensed propositions as their contents. Andy can believe the centred proposition *I am eating a sandwich* without disagreeing with David, who believes the (complementary) centred proposition *I am not eating a sandwich*. And Jack can believe the tensed proposition *It is raining in Boston* on Monday without disagreeing with Helen, who disbelieves the tensed proposition *It is raining in Boston* on Tuesday (MacFarlane 2014, pp. 125-128). Of course someone might offer semantic reasons for resisting the introduction of centred and temporal propositions; our point (which is actually MacFarlane's point) is mainly a methodological one: it would do a bad service to semantic theory to rule out from the start

such views as temporalism by offering, in place of independent semantic grounds, an argument that crucially depends on the assumption of a very narrow notion of disagreement.

Cappelen and Hawthorne's argument can thus be resisted by insisting that the case at hand be construed in terms of disagreement with an *attitude* or *speech act*. When disagreement is conceived as a relation that can obtain (or fail to obtain) between attitudes or speech acts, the moral that can be drawn from the case of Jane and June is again that, if a situation sufficiently dissimilar from the actual one had occurred, a person could legitimately have denied (or asserted a proposition incompatible with) what they (or someone else) actually said without disagreeing with their own act of *asserting* it (and likewise for *belief* and other doxastic attitudes). This is indeed how MacFarlane now unambiguously sees the matter, having disentangled the claim that "to disagree with someone's attitude [...] is to have attitudes the accuracy of which would preclude its accuracy" (MacFarlane 2014, p. 126) from any suggestion that to disagree with an attitude (or a speech act) of someone inhabiting a different possible world is to disagree with a proposition.⁵ And indeed it seems plausible to say that, if in a situation sufficiently dissimilar from the actual one Jane were to deny the proposition that Mars has two moons, she might perhaps find herself to disagree with what she actually said in the sense that she would reject the *content* of a speech act she actually performed, but she would not disagree with the speech act of asserting it, that is to say, with the *assertion* she made. Thus it is tempting to conclude that, if states of interworld disagreement are construed as states of disagreement with a proposition, they are as frequent as Cappelen and Hawthorne maintain, but if they are construed as states of disagreement with an attitude or speech act, they aren't.

4. *An irenic attitude?*

In light of the above, we might perhaps end this discussion on an irenic note by saying that, even when used in the state sense, phrases like ‘A disagrees with B’s belief’ and ‘A disagrees with B’s assertion’ lend themselves to being interpreted in (at least) two different ways, and that (most of) the problems one faces when analysing the conditions of interworld disagreement arise from an ambiguity of such phrases, which allow both a ‘propositional’ and an ‘attitudinal’ reading – that is to say, they can be used either to say that A disagrees with the *propositional content* of an attitude or speech act of B, or to say that A disagrees with B’s *believing* or *asserting* that content (to avoid misunderstanding, please recall that in this paper we are only concerned with *doxastic* disagreement and *doxastic* attitudes, so ‘propositional’ and ‘attitudinal’ contrast with each other, and neither of them contrasts with ‘doxastic’; propositional and attitudinal disagreement are both species of doxastic disagreement).

However, this irenic conclusion is not one we really want to stand by. To repeat, what we are prepared to defend is just the conditional claim that *if disagreeing across possible worlds is possible*, then the conditions for being in such a state will depend, in the way that we have attempted to describe, on whether it is propositional or attitudinal disagreement that is at stake. But we believe there are fairly compelling reasons to reject the antecedent of the conditional.

First note that if states of interworld disagreement were construed as mere states of disagreement with a proposition, irrespective of any specific attitude or speech act of which that proposition is or might be the content (if the proposition, that is, were considered *per se*, and not as the content of any actual or possible attitude or speech act in context), they would be ‘interworld’ only in a very weak and uninteresting sense. After all, propositions do not exist in this or that world, but in all worlds – or in none.⁶ But if something more is needed by way of context to get *real* interworld disagreement, the moral to be drawn from the case of

Jane and June seems to be not just that one can disagree with the proposition that *p* without thereby disagreeing with an attitude or speech act in context that has that proposition as its content, but that (with the obvious exception of necessary truths) the Simple View of disagreement cannot be applied to interworld disagreements about ‘classical’ propositions any more than it can be applied to intraworld disagreements about centred or temporal propositions. And if this is correct, it seems natural to conclude that there cannot be any genuine state of interworld *propositional* disagreement without a corresponding state of interworld *attitudinal* disagreement. For how could having beliefs or making assertions whose contents are jointly incompatible with the proposition that *p* amount to a case of genuine interworld disagreement if (i) the proposition that *p* were not thought of as the content of an attitude or speech act of an individual inhabiting a different possible world or (ii) the attitudes or speech acts of the relevant individuals could all be accurate? Apparently, then, the claim that a certain situation amounts to a state of interworld propositional disagreement can have some plausibility only if the situation in question can *also* be plausibly described as a state of disagreement with an attitude or speech act of an individual inhabiting a different possible world. To put it into a slogan: *no interworld propositional disagreement without interworld attitudinal disagreement.*⁷

But now comes the problem. For there is reason to think that invoking the attitudinal construal to make sense of the sort of interworld disagreement that MacFarlane argues may obtain between individuals inhabiting different possible worlds, and indeed between an individual and his/her counterpart in another possible world,⁸ is ultimately unsuccessful.

It is far from clear that, when one considers the possibility that the attitudes of any two such individuals may preclude the accuracy of one another, the situation in question is one that would ordinarily be described as a state of (doxastic) disagreement. Surely, it makes sense to say that a person *might have rejected* what *another person* actually asserted or

believed, or that a person *rejects* what another person might have asserted or believed. It also makes sense to say that a person *might have disagreed* with another person's believing or asserting a certain proposition (had the latter person believed or asserted something different from what he or she actually believed or asserted, the former person would have disagreed with *that* attitude): these would be cases of *intra-world* attitudinal disagreement. But does it really make sense to say, in addition, that a person *might have disagreed* with another person's *actual* act of asserting or state of believing a proposition, or that a person *disagrees* with the act of assertion another person might have made or with the state of believing another person might have been in (i.e., with a merely *possible* act of assertion or state of believing)? Is it really appropriate to describe cases in which the speech acts or doxastic states of individuals inhabiting different worlds cannot both be accurate as cases of *bona fide* disagreement?

Again, when it comes to cases featuring an individual and his/her counterpart in another possible world, it makes sense to say that someone might have rejected what *they* actually asserted or believed, or that someone rejects what they might have asserted or believed. But does it also make sense to say that someone *might have disagreed* with their own actual act of asserting or with their own actual state of believing a proposition, or that someone *disagrees* with the act of assertion they might have made or with the state of believing they might have been in? Is it appropriate to describe cases in which the speech acts or doxastic states of an individual and his/her counterpart in another possible world cannot both be accurate as cases of *bona fide* disagreement?

The answer we are inclined to give to such questions is that it does not really make sense to say that a relation of doxastic disagreement may obtain (or fail to obtain) between an individual and his/her counterpart in another possible world, or, more generally, between two individuals inhabiting different possible worlds: the acts of assertion made by two such

individuals, or the states of believing two such individuals are in, may well be such that they cannot both be accurate, but that does not entail that a state of disagreement (in any ordinary sense of ‘disagreement’) obtains between them.

There are metaphysical and epistemological reasons for doubting that such situations can be correctly described as genuine cases of (interworld) disagreement. To introduce the former, we begin with a ‘softening-up’ analogy. Consider Edvard’s case (assume that he lives in a country where same-sex marriage is legal). Edvard has proposed to Andreas because he loves him more than any other person in the world, including himself; had he loved Gert more than Andreas, he would have proposed to Gert. But had he loved *himself* more than any other person in the world, would he have proposed to... himself? Of course not. He could not have proposed to himself, because it takes two to make a marriage – and they must inhabit the same world. By the same token, we submit, when it is attitudinal disagreement that is at stake, it takes two to disagree – and they must inhabit the same world. This of course is just an analogy, but our intuition is that to say that someone might have disagreed with their own actual asserting or believing a proposition, or that someone disagrees with an act of assertion they might have made or a state of believing they might have been in does not make any more sense than to say that Edvard might have proposed to himself rather than to Andreas.

However, if the reason why this is so is just that the parties to a relation of attitudinal disagreement must inhabit the same world, then, by the same token, to say that a person might have disagreed (or failed to disagree) with *another person’s* actually asserting or believing a proposition, or that a person disagrees (or fails to disagree) with an act of assertion that *another person* might have made or a state of believing that *another person* might have been in, will not make any sense either. Here intuitions may be a little shakier, but this can be explained by a plausible ‘confusion hypothesis’. The hypothesis is that we tend to confuse the cases in question either with putative cases of *propositional* disagreement (or failure thereof),

or with superficially similar cases of *bona fide* intraworld attitudinal disagreement (or failure thereof). How the first kind of confusion can arise should be clear by now. As for the second kind, the way it can arise can be grasped by considering a small variant to the case of Jane and June. This variant features Jane and Michael, and it differs from the original only in that its second character (Michael) is not Jane's counterpart in a different possible world. So, to recap: Mars has two moons and Jane asserts that Mars has two moons, but if Mars had had three moons, Michael would have asserted that it has three moons. This (again) is a situation in which the speech acts of the relevant agents – Jane's actual assertion and Michael's counterfactual assertion – can be, and indeed are, both accurate, but it is not, we contend, a genuine case of failure of interworld disagreement. The second part of our confusion hypothesis is in fact that, if we are tempted to (mis)construe this situation as an *interworld state* in which *Jane does not disagree with Michael's possible assertion that Mars has three moons*, this is just a consequence of our inadvertently confusing it with an *intraworld state* in which *Jane would not have disagreed with Michael's possible assertion that Mars has three moons* because she too would have asserted that it has three moons (and both assertions would have been accurate).

Someone might wish to object that our confusion hypothesis is a little too 'baroque' to dispel any residual temptation to treat the situations in which the attitudes or speech acts of individuals inhabiting different possible worlds cannot both be accurate as genuine cases of interworld disagreement.⁹ But it seems clear to us that any alleged 'intuition' to the effect that such situations amount to genuine cases of interworld disagreement is best explained as a product of philosophical training and consequent habituation to realist-sounding possible worlds talk. The convoluted turns of phrase we had to use to avoid resort to possible worlds talk in describing the situations in question are evidence that expressing these situations in ordinary language is extremely difficult – so difficult, in fact, to make it seem very

implausible that considering such cases may yield any genuinely pretheoretical intuition that can be summoned in support of the possibility of interworld disagreement.¹⁰ So we think that our confusion hypothesis effectively tips the balance in favour of the view that such situations cannot be appropriately described as cases of interworld disagreement. While there is a sense in which it is always possible for an agent to have beliefs or make assertions whose contents are jointly incompatible with this or that *proposition* (recall that propositions exist in all worlds or in none), for an agent to be able to *disagree* with someone's *attitude* or *speech act* an appropriate object for her disagreement must exist in his/her own world, and not just in any other possible world.¹¹

Of course, one might just *decide* to apply the word 'disagreement' to interworld states that would not be ordinarily described by that term. Our point is simply that stipulating a wider use for the term will do nothing to alleviate the worry that the 'intuitions' *about disagreement* we might hope to gather from such interworld states will actually be parasitic upon the (*bona fide*) intuitions we have concerning related cases of intraworld disagreement.

The idea that the relation of attitudinal disagreement can be instantiated by an individual and his/her counterpart in another possible world also raises serious epistemological worries. Consider a basic case of attitudinal disagreement supposedly involving an individual and his/her counterpart in another possible world:

Basic AD. A accepts P, and if A were to accept not-P in circumstances sufficiently similar to the actual ones, A would disagree with his/her own (actual) attitude.

This situation can be described in (at least) two ways:

(1) $\text{Acc}(A, P) \ \& \ [\text{Acc}(A, \text{not-}P) \rightarrow \text{Dis}(A, \text{Acc}(A, P))];$

(2) $\text{Acc}(A, P, @) \ \& \ [\text{Acc}(A^*, \text{not-}P, M_1) \supset \text{Dis}(A^*, \text{Acc}(A, P, @))]$.

‘Acc’ expresses the relation of accepting – binary in (1), ternary in (2). ‘Dis’ expresses the relation of attitudinal disagreement. ‘ \supset ’ is the material conditional, ‘ \rightarrow ’ the counterfactual conditional, ‘@’ denotes the actual world, ‘ M_1 ’ a possible world sufficiently similar to the actual one. We assume the usual semantics for counterfactual conditionals.

First, then, consider (1). As the antecedent of its conditional is, *ex hypothesis*, false in the actual world, (1) cannot be used to get to know its consequent by deductive inference. For by definition counterfactual conditionals cannot be employed to acquire knowledge of their consequents via modus ponens reasoning. In fact, no counterfactual conditional can feature as the conditional premise of a *sound* modus ponens – a valid argument with true premises whose major premise is the counterfactual and whose minor premise is the counterfactual’s antecedent.¹² This means that the conditional in (1) does not licence any conclusion concerning A’s disagreement (or lack of disagreement) with A’s own actual attitude, which might go some way towards explaining our uneasiness in admitting (for instance) that an individual could have disagreed with his/her own actual attitude.¹³ The uneasiness disappears when the situation is described by means of (2), which has a material conditional in place of (1)’s counterfactual conditional; for in this case it is perfectly legitimate to derive the conditional’s consequent on the assumption of its antecedent. But the price to be paid to secure this result is of course the adoption of a realist interpretation of possible worlds talk.

Despite MacFarlane’s denial, the impression that those situations in which the speech acts and/or attitudes of individuals inhabiting different possible worlds cannot both be accurate may be described as situations of (interworld) disagreement in any ordinary sense of the word would thus seem to be just an artefact of the adoption of a realist construal of possible worlds talk.

5. Conclusion

In this paper we have considered situations of (alleged) interworld doxastic disagreement, discussing both cases involving individuals inhabiting different possible worlds and cases involving an individual inhabiting the actual world and his/her counterpart in another possible world. We have made two main points. First, we have argued that disagreeing across possible worlds, if at all possible, is easier than some philosophers (Recanati, Franc en) have assumed, but not necessarily as easy as other philosophers (Cappelen, Hawthorne) have maintained. Our discussion has brought out that metaphysical assumptions concerning the *relata* of disagreement play an important role in shaping accounts of the conditions for disagreeing, and we have considered the option of concluding that apparently conflicting views on the issue can peacefully coexist as long as they are taken to concern different types of disagreement relations, namely, a propositional and an attitudinal one. However, our discussion of the conditions for disagreeing in sects. 1 and 2 has only been meant to show that *if disagreeing across possible worlds is possible*, then it is easier than some philosophers have assumed, but not so easy as others have maintained. And our second point has been that the antecedent of this conditional is likely false: indeed, we have argued that there are serious metaphysical and epistemological reasons that militate against assuming that it makes sense to describe certain situations as cases of interworld doxastic disagreement in the first place. The cases for both points are substantially independent, so readers unconvinced by the reasons offered in sect. 3 are free to detach the consequent of the conditional conclusion of sects. 1 and 2, while readers sceptical of this conditional conclusion may still find some merit in the line of reasoning articulated in sect. 3. But of course, we believe both cases are strong enough to be worthy of serious consideration.

Notes

¹ If they were, the argument would involve a fallacy of equivocation, for failing to disagree out of normative insulation is crucially different from failing to disagree on a shared normative background. We are grateful to Filippo Ferrari for drawing our attention to this point.

² We owe this suggestion to Matteo Plebani.

³ Disagreement being a symmetric relation, any situation in which someone would have disagreed with the proposition that they, or someone else, actually asserted or believed will be a situation in which someone disagrees with the proposition that they, or someone else, would have asserted or believed. This, however, need not detain us here.

⁴ Cappelen & Hawthorne (2009, pp. 54-66) would likely resist describing their view of disagreement as the view that disagreeing is essentially disagreeing with a proposition: they say that disagreement is primarily a relation between *persons*. However, they expressly subscribe to the Simple View of disagreement, which is what really matters for our purposes.

⁵ MacFarlane initially defended his usage of possible worlds talk in the description of the case of Jane and June by arguing that, “[h]owever you think of modality, it makes sense to ask whether in saying what one *would have* said, in some counterfactual situation, one would have disagreed with what one actually *did* say” (MacFarlane 2007, p. 23), exposing his flank to Cappelen and Hawthorne’s criticism by suggesting that disagreement is a relation to a proposition. But he is now fully explicit that the logical form of states of (intra-world and inter-world) disagreement should be articulated by schemas such as

x is in disagreement with ϕ -ing-in-context- c ,

the underlying thought being that the accuracy of a belief or assertion depends not only on its content, but, crucially, also on its context, which includes the relevant agent (MacFarlane 2014, p. 120). For a defence of interworld disagreement along roughly the same lines, see Marques 2014, pp. 129-130. Incidentally, discussions of Conciliationism in the epistemology of disagreement have recently focused on cases of merely possible (as opposed to actual) disagreement (see, e.g., Kelly 2005; Carey 2011; Barnett & Li 2016); but the cases considered are typically construed as cases of either propositional or intraworld disagreement.

⁶ That propositions exist in all worlds or in none is a presupposition of everything that has been said so far concerning propositional disagreement: to challenge it would be to challenge the integrity of the very notion of propositional disagreement.

⁷ Note that we are not claiming, here, that the notion of attitudinal disagreement is in any sense explanatorily prior to the notion of propositional disagreement. The claim is just that whenever an attribution of disagreement is correct, it must be possible to offer an attitudinal reading of it.

⁸ It is interesting to note that MacFarlane's discussion of the case of Jane and June in *Assessment and Sensitivity* no longer assumes that the latter is the former's counterpart in another possible world: Jane and June are introduced as just two individuals inhabiting different worlds. MacFarlane does not mention the change, but cautiously concludes his discussion by saying that, since 'it is difficult to have any stable intuitions about the case', he does not want 'to rest too much weight on this argument' (MacFarlane 2014, p. 128).

⁹ We are grateful to an anonymous referee for *Erkenntnis* for pressing us on this point.

¹⁰ It is perhaps worth recalling here that MacFarlane has himself come very close to accepting this conclusion in the passage quoted in n. 6 above.

¹¹ Here we are presupposing that attitudinal disagreement cannot be reduced to propositional disagreement: our argument would not go through if disagreeing with someone's believing that *p* boiled down to having beliefs whose contents are jointly incompatible with the proposition that *p* (thanks to an anonymous referee for *Erkenntnis* for calling our attention to this point). However, we have already emphasised in sect. 2 that there are strong reasons for resisting the proposed reduction.

¹² To be clear: we are not claiming that counterfactuals violate modus ponens, but only that an instance of modus ponens featuring a true counterfactual conditional among its premises will necessarily have a false (minor) premise.

¹³ The same epistemological difficulty affects the counterfactual associated with *propositional* disagreement:

Basic PD. A accepts P, and if A were to accept not-P in circumstances sufficiently similar to the actual ones, A would disagree with P.

In this case, too, the conditional's antecedent is not available for running a sound *modus ponens*. One might think that here the situation is different: the counterfactual conditional associated with propositional disagreement is a tautology, because in the case of disagreement with a proposition 'accepting not-P' is actually equivalent to 'disagreeing with P'. But again, there is reason to think that there cannot be any interesting case of interworld propositional disagreement without a corresponding case of interworld attitudinal disagreement.

Acknowledgements

An earlier version of this paper was presented at the Disagreement: Current Topics workshop held in Bologna in September 2017. Thanks to all present for discussion. We are especially grateful to Filippo Ferrari for reading and commenting a draft of the paper and to Matteo Plebani for some very helpful suggestions.

References

- Barnett, Z. & Li, H. (2016), Conciliationism and Merely Possible Disagreement. *Synthese*, 193, 2973-2985.
- Cappelen, H. & Hawthorne, J. (2009), *Relativism and Monadic Truth*. (Oxford: Oxford University Press).
- Carey, B. (2011), Possible Disagreements and Defeat. *Philosophical Studies*, 155, 371-381.
- Egan, A. (2007), Epistemic Modals, Relativism, and Assertion. *Philosophical Studies*, 133, 1-22.
- Egan, A. & Weatherson, B. (2005), Epistemic Modals and Relative Truth. (In G. Peter & G. Preyer (Eds.), *Contextualism in Philosophy* (pp. 131-169). Oxford: Clarendon Press.)
- Francén, R. (2010), No Deep Disagreement for New Relativists. *Philosophical Studies*, 151, 19-37.
- Frances, B. & Matheson, J. (2018), Disagreement. (In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), URL = <https://plato.stanford.edu/archives/spr2018/entries/disagreement/>.)
- Kelly, T. (2005), The Epistemic Significance of Disagreement. *Oxford Studies in Epistemology*, 1, 167–196.
- Kölbel, M. (2002), *Truth without Objectivity*. (London: Routledge.)

- Kölbel, M. (2004), Indexical Relativism versus Genuine Relativism. *International Journal of Philosophical Studies*, 12, 297-313.
- Lasersohn, P. (2005), Context Dependence, Disagreement and Predicates of Personal Taste. *Linguistics and Philosophy*, 28, 643-686.
- Lewis, D. (1988), Statements Partly about Observation. *Philosophical Papers*, 17(1), 1-31.
- MacFarlane, J. (2007), Relativism and Disagreement. *Philosophical Studies*, 132, 17-31.
- MacFarlane, J. (2009), Varieties of Disagreement. Retrieved 20 February 2018, from <http://johnmacfarlane.net/varieties.pdf>.
- MacFarlane, J. (2014), *Assessment Sensitivity. Relative Truth and its Applications*. (Oxford: Oxford University Press)
- Marques, T. (2014), Doxastic Disagreement. *Erkenntnis*, 79, 121-142.
- Recanati, F. (2008), Moderate Relativism. (In M. García-Carpintero & M. Kölbel (Eds.), *Relative Truth* (pp. 41-62). Oxford: Oxford University Press.)
- Rovane, C. (2013), *The Metaphysics and Ethics of Relativism*. (Cambridge, MA and London: Harvard University Press.)
- Stephenson, T. (2007), Judge Dependence, Epistemic Modals, and Predicates of Personal Taste, *Linguistics and Philosophy*, 30, 487-525.
- Yablo, S. (2014), *Aboutness*. (Princeton: Princeton University Press.)