

## PIERRE’S RATIONAL AND PUBLIC BELIEFS

VÍCTOR M. VERDEJO

*Universidade Santiago de Compostela*

---

**Abstract.** Paradigmatic cases of disagreement seem not to be compatible with a widespread kind of solution to Kripke’s celebrated Pierre puzzle. As a result, the classical puzzle about rational belief is shown to be also a puzzle about public disagreement/agreement phenomena. In this paper, I defend that the new public version of the puzzle is substantial and challenging and conclude that a full solution to Kripke’s considerations must offer a satisfactory account of both the rational and public character of belief attributions. I then argue that a notion of non-belief-individuating understanding is plausibly the key notion that would allow us to arrest both versions of the Pierre puzzle.

**Keywords:** Pierre puzzle; rationality; disagreement; coordination; non-belief-individuating understanding.

---

### 1. A paradigmatic case of disagreement

Consider Antoine, a respectable logician and philosopher somewhere in Paris. Antoine has never been in London but has heard beautiful things and has seen wonderful pictures and TV spots of a city called ‘Londres’. As a result, Antoine has acquired every inclination to assent to utterances of “Londres est jolie”. Antoine is just as smart and linguistically capable as his Parisian peers. Since one typically takes coherent assent/dissent manifestations as a guide to a subject’s belief and since, presumably, beliefs are translatable from one language to the other, we seem entitled to judge that Antoine believes that London is pretty. We can state Antoine’s belief in terms of a two-place belief (*B*) relation between Antoine and a proposition ‘*Pl*’ which is obtained from the predicational combination of the constant ‘*l*’ for London with the monadic predicate ‘*P*’ for the property of being pretty:

(A1)  $B(\text{Antoine}, Pl)$

Now consider Anthony, a perfectly rational and capable logician who lives and works in an unattractive and especially dirty part of London. Anthony is like his co-workers and neighbours and shows concern about his clearly improvable urban surroundings. They often speak negatively about what they call ‘London’, the name of the city they live in. As a result, Anthony has every inclination to assent to utterances of the sentence “London is not pretty”. Again, since consistent assent is typically a sufficient condition for belief attribution we seem entitled to attribute to Anthony the belief

*Principia* 16(3): 451–469 (2012).

Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

that London is not pretty. Using the same notation as before, Anthony's belief can be expressed as in (A2):

$$(A2) \quad B(\text{Anthony}, \neg Pl)$$

There are several details of the particular life-story of Antoine and Anthony that may make their case an especial or exceptional one. I would like the reader to consider, however, that this is just as mundane and ubiquitous a case as the actual lack of details suggests. Now, Antoine and Anthony need not ever interact with each other or even be coetaneous. However, the discussion I wish to bring to the fore in this paper proceeds via a fundamental assumption. The assumption is that, as described above, the case of Anthony and Antoine is a case, and indeed a paradigmatic case, of disagreement. Antoine and Anthony disagree about the beauty of London, or at least, they disagree if anyone ever did.

## 2. Rationality in, disagreement out

I doubt that this target assumption is even remotely controversial at this preliminary stage. As it happens, however, this assumption leads to an incredibly challenging complication of one of the most discussed puzzles in the philosophy of language and mind. Now, as the reader no doubt has noted, the Antoine-Anthony case is in fact a slight variation of one of Kripke's celebrated versions of his puzzle about belief (Kripke 1979, p.249–54). Let us briefly review the Pierre puzzle.

Pierre is a leading philosopher and logician in France. In ordinary interaction with his French peers, Pierre hears wonderful things of a famous distant city called 'Londres'. On this account, Pierre adopts an inclination to assent to utterances of "Londres est jolie". Thus, we are apparently forced to conclude that Pierre believes that London is pretty, if anyone ever did.<sup>1</sup>

At a later time, however, Pierre ends up living and working in what is, in fact, a Londoner suburb. The vicissitudes that led him to his final destiny do not matter for present purposes. As it happens, Pierre's actual surroundings are especially unattractive and dirty. As his Londoner neighbours, Pierre has every inclination to assent to the utterance of "London is not pretty". As a result, we seem to be forced to conclude that Pierre believes that London is not pretty. In terms of the above introduced notation, we can succinctly expose the paradoxical outcome of the Pierre case in terms of the truth of both (P1) and (P2).

$$(P1) \quad B(\text{Pierre}, Pl)$$

$$(P2) \quad B(\text{Pierre}, \neg Pl)$$

The truth of (P1) and (P2) are paradoxical of course in the light of the fact that

Pierre is, by assumption, a perfectly rational subject. And it goes without saying that rational subjects do not believe contradictory propositions. The underlying idea, as we may put it, is that in characterizing Pierre's beliefs we must respect the principle of non-doxastic contradiction (NDC):

(NDC)  $S$  is a rational subject only if, for any  $p$ , it is not the case that  $B(S, p)$  and  $B(S, \neg p)$ .

On the face of it, there is little reason to deny or doubt (NDC).<sup>2</sup> Even more, Kripke strongly suggests that (NDC) is the key principle to be respected when he urges us to tell whether Pierre believes or disbelieves that London is pretty.

To reiterate, this is the puzzle: Does Pierre, or does he not, believe that London is pretty? It is clear that our normal criteria for the attribution of belief lead, when applied to *this* question, to paradoxes and contradictions. (Kripke 1979, p.259, emphasis his)

In this passage, Kripke clearly assumes that, at least in principle one should not say that Pierre does believe and disbelieve that London is pretty. Hence the "paradoxes and contradictions" he points out.

Now, this might not be the only way of stating the puzzle introduced by the Pierre case. For instance, some authors focus on a meta-theoretical or formal reading of the case, according to which, the puzzle would be one about the consistency or correctness of our belief-reports and not, or not only, about the rationality of Pierre as threatened by the truth of those reports. For expository purposes, I will just focus on the version of the puzzle that takes (NDC) as the core of the problem although both (the theoretical and the rationality) sorts of puzzles are arguably strongly connected.<sup>3</sup> What matters for present purposes is that several solutions have been provided to the Pierre puzzle under the hypothesis that to respect (NDC) is the key for a solution.

So understood, the kind of preferred and well-known solution to Pierre's puzzle has been to introduce senses, modes of presentation or other sorts of intensional categories in order to discriminate between Pierre's 'Londres'- and 'London'-beliefs. Variations of this line of reasoning, of both Fregean and Russellian spirit, can be found in authors of practically all stripes: Bilgrami (2012, p.119–23), Chalmers (2011, p.611–12), Crimmins (1992, p.161–63), Salmon (1986, p.129–32), Schiffer (1992, p.507–10) or Zalta (1988, p.189–96) are some examples.<sup>4</sup> Under this interpretation, the Pierre case should be understood as involving in fact two distinct but non-contradictory beliefs. More precisely, the interpretation is that Pierre should be attributed something like the beliefs specified in (P3) and (P4) where ' $l_1$ ' and ' $l_2$ ' denote the modes of presentation, senses or notions associated to the expressions 'Londres' and 'London' respectively:

- (P3)  $B(\text{Pierre}, Pl_1)$   
 (P4)  $B(\text{Pierre}, \neg Pl_2)$

Note that, *prima facie*, and having (NDC) as the primary target to be respected, it is perfectly reasonable to suppose that some analysis along the lines of (P3) and (P4) would do the job for surmounting Kripke's Pierre puzzle. According to (P3) and (P4), it turns out that Pierre does not believe after all contradictory propositions. He simply believes and disbelieves non-contradictory or directly compatible propositions. Pierre's rationality is thus completely saved.

Optimism grounded in analyses along the lines of (P3) and (P4) is undermined, however, by reflection on the Antoine-Anthony case presented at the beginning of this paper. More precisely, analyses in terms of (P3) and (P4) are actually unacceptable from the point of view of the Antoine-Anthony case. These cases are structurally related. As the reader may recall, the Antoine-Anthony case is mounted in such a way that Antoine's situation is exactly like Pierre's situation in France; whereas Anthony's situation is exactly like Pierre's situation in England. It follows from this that whatever is a good characterization of Pierre's belief about London in France must, as a matter of course, be good for the characterization of Antoine's belief about London. Likewise, whatever is judged to be appropriate for the characterization of Pierre's belief about London in his Londoner lifetime must also be judged appropriate for the characterization of Anthony's belief about London. To repeat, the cases are mounted in such a way that differing analysis for each would be utterly unjustified.<sup>5</sup> From all this it follows that, if our preferred solution to the Pierre puzzle is an analysis along the lines of (P3) and (P4), then we seem to be forced to analyse the Antoine-Anthony case along analogous lines, that is, along the lines of (A3) and (A4):

- (A3)  $B(\text{Antoine}, Pl_1)$   
 (A4)  $B(\text{Anthony}, \neg Pl_2)$

But then, please note, what begun as a puzzle about belief, apparently turns into a puzzle about disagreement. For recall that we made the apparently quite uncontroversial assumption that, as described above, the Antoine-Anthony case is in fact a paradigmatic case of disagreement. As it happens, however, many philosophers have pointed out that a necessary (albeit arguably not sufficient) condition for disagreement is the holding of contradictory beliefs or belief-contents (e.g., MacFarlane 2007, Richard 2011). In other words, by general philosophical assent, two subjects  $S$  and  $S'$  are taken to disagree only if  $B(S, p)$  and  $B(S', \neg p)$ —perhaps together with some other conditions. If Antoine and Anthony genuinely disagree, as we are assuming, then they must belief and disbelief the very same proposition or a given proposition under the same mode of presentation. But that is exactly what (P3) and

(P4) force us to deny.

In short, a well-known strategy for dealing with Kripke's rationality puzzle about Pierre's beliefs appeals to different senses, modes of presentation, notions and other intensional categories in order to save what seems to be a central principle of rationality, i.e., (NDC). However, this seems to save Pierre's rationality at an unpalatably high price, namely, the price of giving up the very possibility of an account of paradigmatic cases of disagreement. It seems to follow that the foregoing widespread strategy for dealing with Kripke's Pierre puzzle changes a puzzle about rational belief for a puzzle about disagreement or public belief.

Now, our starting assumption, i.e., the assumption that the Antoine-Anthony case is a case of disagreement may at this point begin to seem contentious. This is why I would like to offer some motivation for accepting the assumption. First, the Antoine-Anthony case seems to offer a very basic and minimal sort of intercultural disagreement. If we admit that people speaking different languages may and often do disagree with each other then it is hard to deny that Antoine and Anthony should be taken to disagree with each other. There does not seem to be anything especial about this case that would prevent considering it as a genuine and central case of disagreement.

On the other hand, the assumption is also well-motivated if we consider the Antoine-Anthony case in the specific context of Kripke's discussion about Pierre. As known, Kripke appeals to two principles in the analysis of the Pierre case. One is the Disquotational Principle (DP), according to which sincere assent that  $p$  is sufficient for believing that  $p$ . The other is the Translation Principle (TP), according to which translations of true belief-reports are true (cf. Kripke 1979, §2). These principles are not apodictic. Kripke himself dispensed with TP by introducing the monolingual Paderewski case (Kripke 1979, p.265–66). Others have argued that not even DP is required in order to generate de puzzle (e.g. Salmon 1986, Sosa 1996). Nonetheless, if, following Kripke, the principles are accepted as the source of the Pierre case, then it is hard to see what else should be needed for Antoine and Anthony to actually disagree with one another. Anthony assents to "London is not pretty" and by DP that is sufficient for his believing that London is not pretty. Antoine assents to "Londres est jolie" and by DP and TP Antoine believes that London is pretty. What else should be required for there to be disagreement? Note that the fact that there is disagreement in this case is not paradoxical in the same way as the Pierre case. The disagreement is paradoxical only in the light of the above purported (roughly Fregean) solutions to the Pierre case.

Importantly, however, the present discussion need not establish the correctness of the assumption to go through. It only requires its initial plausibility (I would say more about the possibility of cancelling this assumption in the last section). We can therefore explicitly conditionalize the present considerations without losing

theoretical appeal: if the Antoine-Anthony case is one of disagreement, then the standard solutions to the Pierre puzzle change a puzzle about rationality for a puzzle about disagreement. In the next Section 3, I would like to defend that this is not an incidental result, but is deeply rooted into the cases under consideration.

### 3. Rationality out, disagreement out

There might seem at first glance that the theoretical possibilities of theorists wanting to resort to the widespread analysis of (P3) and (P4) are not exhausted by the above remarks. Is there a way in which one can accommodate Pierre's respecting (NDC) along the lines of (P3) and (P4) and still leave room for the paradigmatic sort of disagreement shown in the Antoine-Anthony case? At this point, a particular strategy seems to be especially promising for the defender of (P3) and (P4) and similar accounts. This is the strategy: to connect Pierre's (rational) beliefs as specified by means of (P3) and (P4) to some other sort of (public) beliefs in such a way that, even if (P3) and (P4) do credit Pierre with an impeccable and non-contradictory rationality, (P3) and (P4) can still be the ground for a disagreement of the sort is targeted in the Antoine-Anthony case. Following this train of thought, one should discriminate between attitude-reports that account for a subject's rationality and attitude-reports that account for a subject's public disagreement/agreement interaction.

What kind of link would then be appropriate to connect rational and public kinds of belief in this way? Several authors have suggested links that may be judged to be apt for this task. For instance, some authors have considered a link between beliefs in terms of representation of the same Russellian or purely referential propositions (e.g., Heck 1995) or in terms of similarity of Fregean thoughts (e.g., Forbes 1987) or by means of a notion of coordination (e.g., Fine 2007, Chalmers 2011). In this paper, I would like to consider Chalmers's (2011) coordination-account. For present purposes, Chalmers's proposal is of especial interest since it has the advantage of providing both a particular solution to the Pierre case and also a characterization of the notion of disagreement.

In order to understand the target (Chalmersian) notion of coordination it is important to note that Chalmers distinguishes between endorsing and believing (2011, p.619–21). These propositional attitudes are defined, within Chalmers theory, over what he calls enriched propositions but we can ignore this for present purposes and just think of propositions in the standard ways.<sup>6</sup>

Now, on Chalmers's view, belief is analysed in terms of endorsement and coordination. According to Chalmers, for *S* to believe that *p* is for *S* to endorse a proposition coordinate with *p*. This view leads us to a specific thesis about disagreement. In particular, Chalmers contends that

... two people disagree when one believes a proposition  $p$  and the other believes  $\neg p$ , which requires that one endorses a proposition coordinate with  $p$  and the other endorses a proposition coordinate with  $\neg p$ . (2011, p.619)

Clearly, Chalmers's is a perfect instance of the kind of view advanced at the beginning of this section, that is to say, the kind of view that distinguishes between rationality-grounding attitude-reports—viz. endorsements—and publicity- or disagreement/agreement-grounding attitude-reports—viz. coordinated beliefs. Following this train of thought, Chalmers (2011, p.611–12) observes that Kripke's case should therefore be handled by means of the two-place endorsement ( $E$ ) relation between subjects and propositions and the two-place coordination ( $C$ ) relation between propositions. With the here presented notation, we can state this proposal for analysing the Pierre case in terms of (P5) and (P6):

(P5)  $E(\text{Pierre}, Pl_1) \& C(Pl_1, Pl)$

(P6)  $E(\text{Pierre}, \neg Pl_2) \& C(\neg Pl_2, \neg Pl)$

(P5) says that Pierre is  $E$ -related to the proposition that contains the intension associated with 'Londres' ( $Pl_1$ ) and (P6) says that Pierre is  $E$ -related to the proposition that has the intension associated with 'London' as a constituent ( $\neg Pl_2$ ). In turn,  $Pl_1$  is coordinated with  $Pl$  and thus involves a belief that  $Pl$ ; whereas  $\neg Pl_2$  is coordinated with  $\neg Pl$  and, accordingly, amounts to a belief that  $\neg Pl$ .

On this account therefore, (P5) and (P6) provides us with a way of analysing Pierre's rationality-grounding (belief-)attitudes along the lines of (P3) and (P4) and, at the same time, a way of satisfactorily analysing Antoine and Anthony's disagreement-grounding beliefs along the lines of (A1) and (A2). The puzzles about belief and about disagreement are, one might think, apparently overcome.

This impression is, on reflection, clearly illusory. The main reason is that analyses along the lines of (P5) and (P6) fail to accommodate another principle that seems undeniable as a constraint on *both* genuine rationality-grounding *and* disagreement-grounding propositional attitudes. The key idea is now that rational subjects, whatever exactly their demerits and merits, do not disagree with themselves. We may state this idea in terms of the principle of non-reflexive disagreement (NRD):

(NRD)  $S$  is a rational subject only if, for any  $p$ , it is not the case that  $S$  disagrees with him/herself about whether  $p$ .

I think that (NRD) is better seen indeed as a corollary or as a consequence of (NDC), the principle of non-doxastic contradiction seen in Section 2. Thus, it may be taken to arise directly from the minimal requirements for a rational kind of belief.<sup>7</sup> Be that as it may, (NRD) seems quite uncontroversial a principle about rational belief *and* disagreement. However, this is precisely a principle that Chalmers's analysis cannot in any way accommodate. Let us make this explicit.

Recall that we may use the theoretical apparatus of (P5) and (P6) to conclude that Antoine and Anthony do disagree with each other about the beauty of London. However, if this is correct, we are forced to conclude also that Pierre actually disagrees with himself about the beauty of London. Again, what is good for the Antoine-Anthony case must be good for the Pierre case. But Pierre cannot be taken to disagree with himself. He is a remarkable logician and philosopher. How could he disagree with himself about anything? It follows that (P5) and (P6) leads us to a result that threatens both a reasonable notion of rationality and disagreement. It would seem then that there are no options available for those who defend a strategy along the lines of (P3) and (P4) in the account of Pierre's rationality-grounding and disagreement-grounding beliefs. For any theory that appeals to something like (P3) and (P4) is a theory that either keeps Pierre's rationality in—in terms of (NDC)—while leaving out the disagreement exhibited by Antoine and Anthony; or even worse, a theory that can neither account for the rationality of Pierre nor for a genuine notion of disagreement—by not respecting (NRD).

We have arrived at this conclusion by paying special attention to Chalmers (2011) developments. However, it seems quite plausible that the conclusion can be generalized to any theory that resorts to a version of (P3) and (P4). If we claim that Pierre's rationality is saved by appealing to attitudes that have the propositional contents specified in (P3) and (P4) then, no matter how exactly we understand those contents (e.g. in terms of Fregean thoughts, enriched propositions or other intensional notions), the paradigmatic disagreement of Antoine and Anthony is clearly left out. And if we try to capture the disagreement-agreement case of Antoine and Anthony by establishing a relation between allegedly rational attitudes—viz. the ones specified along the lines of (P3) and (P4)—and allegedly public beliefs—viz. the ones specified by means of (A1) and (A2)—then, no matter how exactly we understand that relation, a problem will arise about the attribution of irrational disagreement to Pierre. Quite generally then, once we have approached the issue by means of (P3) and (P4) we would seem forced to miss a satisfactory notion of rational *and* public belief. A substantial and challenging puzzle about rationality and disagreement is with us.<sup>8</sup>

#### **4. Bringing rationality and disagreement back together**

Now, is there a solution to the Pierre puzzle that respects both our intuitions about rationality and our intuitions about public disagreement/agreement phenomena? I would say that this question leads us into quite unexplored territory. It may very well be that, at this stage, we are not in the best position to offer a fully satisfactory answer to the question. There is however one preliminary point that seems plausible

and is worth emphasizing. The preliminary point is that, if the foregoing considerations are sound, it would seem that we really need to provide a solution to both sides of the Pierre puzzle at once. Many authors—among which we should count, I believe, Kripke himself—have taken it that the Pierre puzzle is mainly or fundamentally a puzzle about rational belief. A way of summarizing the main points of our discussion is to note that Pierre's puzzle is actually a puzzle about the rationality and also about the publicity of our beliefs. In other words, the target puzzle is a puzzle about how belief reports can be compatible with the rationality-grounding and public disagreement-grounding character of our beliefs. The presumption is therefore that theorists are well-advised to consider both the rational and public character of their belief-analyses for a fully satisfactory solution. As I see it, the here presented problem regarding disagreement comes largely from neglecting the necessary public aspect of our rational beliefs.

Once this preliminary point is made clear, there are, I would say, two general strategies that seem to be open for a theorist faced with this new version of the puzzle. One alternative is to reject the assumption by which I begun this paper. This would amount to claiming that, in fact, Antoine and Anthony do not really disagree with each other or that, in case they actually disagree, this is an exceptional outcome. The guiding idea for this line of reasoning would be that cases of genuine disagreement are extremely rare. For instance, it may be argued that genuine disagreement occurs only among epistemic peers and epistemic peers are themselves rare (see King 2012 for a view along these lines). Under this interpretation, we could happily safeguard the analyses of (P3) and (P4) and resolutely maintain that Antoine and Anthony do not disagree after all. One such view would call of course for a deep revision of our pretheoretical notion of disagreement.

There is however an alternative strategy that seems to me more faithful to our pretheoretical intuitions. This alternative strategy consists of maintaining that the Antoine-Anthony case is a paradigmatic case of disagreement. But to secure this claim would, however, have us abandon highly intuitive principles such as (NDC) and (NRD), at least, as they are formulated above. In a sense, therefore, this alternative would also involve the abandonment of important intuitions about rationality and disagreement.

But how exactly these plausible principles must be modified in order to offer a way out, or at least the promissory notes for a way out, of the puzzle that concerns us here. In the remainder of this paper, I would like to offer the motivation and the formal statement of such modification. The modification expresses what I take to be the correct way of approaching the issue once we have decided to preserve the intuition that Antoine and Anthony actually disagree. Let me then first briefly motivate the target notion of understanding on which my proposal relies.

The notion of understanding of a given propositional content or concept may be

approached in a number of ways. Here I wish to be liberal about what exactly this notion would amount to. A baseline characterization says that understanding of the propositional content (and constituents of)  $p$  involves a kind of knowledge of  $p$  (cf. Burge 1979, Brown 2000, Wikforss 2008). Crucially for present purposes, rationality has taken to be dependent upon understanding in a very particular and especial way. The idea, common among both Fregean and non-Fregean views, has been that differing understandings relevant for rationality are in fact content or belief-individuating. This idea, in my view, is what is behind the attempt to save Pierre's rationality by means of (P3) and (P4) above. More precisely, (P3) and (P4) involve the postulation of different beliefs or belief contents in accordance with the senses or modes of presentation associated with 'Londres' and 'London' respectively. This gives credit to the view that different understandings associated with 'Londres' and 'London' turn on different belief contents or modes of presentation, and thus, in different but compatible individuated beliefs. (P3) is in this way taken to be compatible with (P4) to the extent that they express different understandings Pierre has of London in his French and English lifetimes. Roughly then, on this view, a subject's having a different understanding regarding the belief that  $p$  results in this subject actually having another different belief that  $p'$ .

The idea that different understandings of  $p$  automatically turn on different  $p$ -beliefs is probably so unquestioned that it would appear almost trivial. How else, one might think, could understanding affect our beliefs if it is not by giving rise to different beliefs? As it happens however, the idea that different understandings of  $p$  turn on different individuation of  $p$ -beliefs is, at least in the simple terms just exposed, profoundly misguided. On the one hand, differing understandings are not rightly taken to affect a subject's individuated beliefs when the differing understandings in question are, in fact, a kind of misunderstanding. Intuitively, the fact that  $S$  misunderstands  $p$  should not count as automatically affecting the identity of  $S$ 's  $p$ -beliefs. This is, for instance, one of the lessons of the celebrated 'arthritis' case (Burge 1979). Tyler Burge's anti-individualism precisely receives support from scenarios in which a certain substantial misunderstanding—e.g. the misunderstanding involved in believing in the actual world that someone has arthritis in her thighs—does not affect the individuation of one's beliefs—e.g. does not amount to a claim that the subject has a kind of 'arthritis'-beliefs different from those of her peers.

On the other hand, it is simply implausible for a minimally public notion of belief that it should always vary with different understandings. At least sometimes—i.e. when the understandings in questions are sufficiently close to one another—differing understandings (of a certain belief-content) need not involve differing beliefs. If minimal variations of understanding  $p$  automatically involved different  $p$ -beliefs the very possibility of public beliefs among subjects (exhibiting slight understanding variations) would be seriously undermined.

The foregoing remarks should be sufficient to motivate the serious consideration of a non-belief-individuating notion of understanding. Now, once the theorist has decided to preserve our intuitions about disagreement, the key movement for a satisfactory solution of the here discussed Pierre puzzle is precisely the appeal to this notion of difference in understanding, namely, a notion of difference in understanding that does not automatically affect the individuation of one's beliefs—as opposed to the (epistemic) relation one stands towards such beliefs.<sup>9</sup>

This is why a particular modification of the principles (NDC) and (NRD) above might be justified. The modification crystallizes in a new qualifying condition for (NDC) and (NRD). The new qualifying condition is that the beliefs involved in those principles are understood in the same way. As noted, I wish to be liberal about the notion of understanding so that theorists are free to offer their preferred detailed interpretation. Let us then explicitly state this general modification in terms of (NDC') and (NRD') where ' $\mu$ ' denotes an arbitrary monadic function on propositional contents:

(NDC')  $S$  is a rational subject only if, for any  $p$ , it is not the case that  $B(S, p)$  under understanding  $\mu_1(p)$  and  $B(S, \neg p)$  under understanding  $\mu_2(p)$  and  $\mu_1(p) = \mu_2(p)$ .

(NRD')  $S$  is a rational subject only if, for any  $p$  and understanding  $\mu(p)$ , it is not the case that  $S$  disagrees with him/herself about whether  $p$  under  $\mu(p)$ .

Equipped with (NDC') and (NRD') we may be able to provide an analysis of the Pierre case that is compatible with our intuitions about rationality and disagreement. Following this, we should interpret Pierre's case in terms of (P7) and (P8).

(P7)  $B(\text{Pierre}, Pl)$  under  $\mu_1(Pl)$

(P8)  $B(\text{Pierre}, \neg Pl)$  under  $\mu_2(Pl)$

For (P7) and (P8) to actually be of use in the elucidation of the Pierre case we need to assume that, in fact,  $\mu_1(Pl) \neq \mu_2(Pl)$ . But this assumption is well grounded. The different understandings in question can be seen as arising from Pierre's ignorance about the fact that 'Londres' and 'London' are names of the same city. In other words, Pierre understands the belief contents in (P7) and (P8) in different ways. This is why he in fact holds contradictory beliefs. Because of the different understandings involved, Pierre does not realize they are contradictory beliefs! This is plausible: Pierre, with all his rationality and logical capacity, need not always fully understand what he believes. In short, it would seem that (NDC') is utterly respected.<sup>10</sup>

Similarly, note further that, since there are different understandings involved, there is no single understanding of  $Pl$  under which Pierre could be said to disagree about whether  $Pl$ . It follows that (NRD') is also respected.

Now, how would this analysis deal with the case of Antoine and Anthony? Can

we account for the disagreement they intuitively exhibit? The propounded analysis would be something along the lines of (A5) and (A6):

(A5)  $B(\text{Antoine}, Pl)$  under  $\mu_1(Pl)$

(A6)  $B(\text{Anthony}, \neg Pl)$  under  $\mu_2(Pl)$

If the analysis is to be consistent regarding the foregoing considerations however, note that, since by assumption Antoine and Anthony are, in the relevant respects, exactly in the same situation as Pierre, also in this case  $\mu_1(Pl) \neq \mu_2(Pl)$ . It follows that Antoine and Anthony disagree in spite of arguably having a sort of misunderstanding about the target beliefs. They also lack the knowledge that the city called 'Londres' and the city called 'London' are the same city. But note that this is a perfectly plausible situation. Just imagine that Antoine does not know a word of English and that Anthony does not know a word of French. Still, their situation is plausibly a situation in which there is genuine disagreement. They understand differently two contradictory beliefs but this does not prevent them from holding contradictory beliefs. It follows that the Antoine-Anthony case is still plausibly a case of genuine disagreement. Disagreement seems to be finally rescued.

It may seem questionable that, according to the just introduced analysis, we are attributing significant sorts of misunderstandings regarding the contents of their beliefs to subjects that are, by assumption, perfectly rational and linguistically capable. How can a rational subject misunderstand the contents of her own beliefs in such a way? This is not the place to fully address this worry. It seems to me, however, that the source of the uneasiness at this point has to do with the fact that misunderstanding might be thought to threaten the intuitive view that rational subjects have privileged or authoritative knowledge of her own beliefs. In reply, it is crucial to note that we must distinguish sharply between having (privileged or authoritative) knowledge of and (fully) understanding our own beliefs. I had said that the view that understanding does not always individuate beliefs is in part supported by cases of partial (mis)understanding of one's own concepts and belief contents, such as those pointed out in anti-individualistic scenarios. It does not follow that these cases involve lack of authoritative knowledge of our own beliefs and thoughts. The possibility of combining misunderstanding with self-knowledge has been a major point of anti-individualism at least since Burge (1988). The here presented view is therefore still compatible with Pierre, Antony and Antoine's having a kind of authoritative knowledge about their beliefs even if lacking a full understanding of them. The authoritative knowledge they have would have to do with the self-verifying character of their second order beliefs and is quite independent of their understanding (see Burge 1988).

Summing up, once we have decided to respect the intuition that Antoine and Anthony possibly disagree with each other, a plausible strategy for dealing with the

target puzzle is to appeal to a notion of understanding that does not individuate one's beliefs. If this is correct, Pierre may exhibit a kind of understanding that prevents him from realizing that he holds contradictory beliefs. A non-belief-individuating sort of understanding is what allows us to maintain Pierre's rationality. Pierre is perfectly rational but he is also human and, sometimes, he may fail to understand fully or understand in the same way his own beliefs. On the other hand, our intuitions about disagreement are also in good standing: on the one hand, Pierre does not disagree with himself in an unpalatable way. He may be taken to disagree with himself about whether London is pretty but just because there are different understandings involved of the belief and disbelief that London is pretty. Finally, on the assumption that people can disagree about things they do not understand in the same way, Antoine and Anthony should be counted as exhibiting a genuine sort of disagreement.

One may complain at this point that the propounded solution to the Pierre puzzle just offered proceeds via a questionable assumption. The questionable assumption is that we are actually allowed to reformulate the principles (NDC) and (NRD) to get (NDC') and (NRD'). The assumption is questionable because the puzzle was originated precisely by assuming the need to respect (NDC) and (NRD). The above considered strategies, namely, the general strategy expressed in terms of (P3) and (P4) and strategies along the lines of the Chalmersian (P5) and (P6) would perhaps be benefitted if they actually appealed to this assumption. These strategies would not have to worry about their failure to respect both (NDC) and (NRD) at the same time. They could argue that they have the apparatus sufficient to deal with (NDC') and (NRD') instead.

In order to take this suggestion seriously we need to distinguish between the case of the simple Fregean version expounded in Section 2 and the sort of coordinated account discussed in Section 3. First, as regards the simple Fregean version, we need to reassess (P3) and (P4) in the light of (NDC') and (NRD'). One may argue that (P3) and (P4) correspond to different belief-individuating understandings of Pierre regarding London. On this reading, (NDC') would be trivially maintained. Clearly, if (P3) and (P4) each involve different kinds of understanding then Pierre would very clearly not even believe contradictory propositions. However, note that, just as before, in this case the differing beliefs must be kept different in the case of Antoine and Anthony and this, just as before, prevents them from satisfying the necessary condition on disagreement, namely, that they have to believe contradictory propositions. Alas, to repeat, we are assuming that there *is* disagreement between Antoine and Anthony regarding the beauty of London.

Remarkably then, the simple Fregean strategy does not really change his dialectical situation once (NDC') and (NRD') enter into the picture. Rationality by means of (NDC') and (NRD') in the case of Pierre is, just as in the case of (NDC) and (NRD), saved only at the price of giving up the potential disagreement grounded in Pierre's

beliefs. So the simple Fregean account is still no match for providing us with a public and rational sort of belief.

What about accounts along the lines of (P5) and (P6)? Well, let us assume that the target principles to be respected are therefore (NDC') and (NRD'). I would like to concentrate on two different questions. On the one hand, do (P5) and (P6) respect these principles? Secondly, do (P5) and (P6) respect these principles in a coherent way?

Let us begin with the first question. The answer is clearly yes. (P5) and (P6) respect (NDC') because they claim that Pierre believes and disbelieves that London is pretty but by exhibiting what we may take as different understandings. According to (P5) and (P6), we may plausibly say, the different sorts of understandings are expressed in terms of different propositions Pierre endorses (in the technical sense of endorsing). It follows that (P5) and (P6) are perfectly compatible with (NDC'). What about (NRD')? Well, for analogous reasons, the answer here is that (P5) and (P6) actually provide a solution compatible with (NRD') too. They assert that Pierre does believe contradictory propositions and thus disagrees (or eventually disagrees) with himself but this is acceptable according to (NRD') because Pierre exhibits different sorts of understandings or endorsements. So apparently, (P5) and (P6) are a satisfactory solution to the Pierre puzzle after all.

It is important to note at this point that the way in which (P5) and (P6) successfully deal with (NDC') and (NRD') squares very well with the propounded diagnosis when I introduced (P7) and (P8). More precisely, both pairs (P5) and (P6), on the one hand, and (P7) and (P8) on the other, involve the postulation of understandings that do *not* individuate a subject's beliefs. Indeed, the notion of belief appealed to in (P5) and (P6) is, as noted, a notion that allows different endorsements (or understandings) to correspond to one and the same belief.

Now, let us be clear. The fact that (P5) and (P6) respect (NDC') and (NRD') means that they are, *prima facie*, up for grabs for theorists interested in so respecting (NDC') and (NRD') and are, to that extent at least, a respectable theoretical option for the solution of the Pierre puzzle. However, let us now address the second question about whether (P5) and (P6) offer a solution that coheres with our previous discussion.

Well, there are two ways of seeing that the solution offered by (P5) and (P6) is not really coherent in the light of the foregoing considerations. First, the attitude of endorsement, central in the analysis of (P5) and (P6), seems unwarranted. The reason is that it is not clear at all what role exactly does this attitude play, and, more importantly, it is clear that it does not play the role it was supposed to play. Let me elaborate.

Endorsement was supposed to be a sort of rationality-grounding attitude. This Chalmers makes explicit when he claims that endorsements are "more fine-grained

characterizations [of a subject's psychological state] that [unlike beliefs] are constitutively connected to rationality" (Chalmers 2011, p.612). Note then that endorsements arise in exactly the same context of accounts such as (P3) and (P4), namely, accounts that assume that the key for explaining rationality in the Pierre case involves the postulation of different attitude-individuating (in their case belief-individuating) understandings. This is exactly what endorsements were supposed to do in the first place: to preserve Pierre's rationality in terms of non-contradictory attitudes regarding London. But this is wrong. Pierre's rationality is not preserved in this way. On the contrary, Pierre's rationality is preserved once it is acknowledged that Pierre may rationally hold contradictory (belief) attitudes for which he has different understandings. In other words, endorsements make it seem that it is a necessary condition on rational attitudes to somehow respect (NDC), but this is precisely what we are denying when considering (NDC'). Indeed, endorsements in (P5) and (P6) are just another way of stating the attitude-individuating understandings of (P3) and (P4). But we have just seen that the only way in which attitude-individuating understandings can accommodate both (NDC') and (NRD') is precisely by ceasing to individuate a subject's target rational *beliefs*. (P5) and (P6) therefore obtain the correct result but for what is just the wrong reason. They obtain the correct result—that is, they respect (NDC') and (NRD')—using endorsements as the explanation of Pierre's rationality whereas, ironically, what gives us the correct result is precisely the fact that endorsements play no role in the individuation of Pierre's *rational* beliefs.

Following this, it is remarkable that, at least for present concerns, we could leave endorsements and other related attitudes entirely out of the analysis. All that we need is to postulate different understandings for Pierre's contradictory beliefs. We do not need in addition to consider these understandings in terms of a particular propositional attitude. This shows that such doxastic or quasi doxastic states are really useless in this particular context. They are certainly an option for the resolution of the Pierre puzzle and their postulation might be justified on other grounds. They are, nonetheless, entirely optional and unwarranted for the specific Pierre puzzle under consideration. One may postulate differing understanding of a proposition *p*, without assuming that it involves a sort of doxastic or quasi doxastic attitude regarding another suitably related or coordinated proposition *p'*. Once the real problem is uncovered, the need for the usual characterization of Pierre's French and English beliefs as involving different sorts of belief- or attitude-individuating contents simply vanishes.

There is, I believe, one last and deeper reason to cast doubts on the coherency of analyses along the lines of (P5) and (P6) in this context. The reason is that (P5) and (P6) make it seem that rationality issues and publicity issues affecting Pierre's beliefs are distinct and quite unrelated kinds of issues (see Section 3 above). By contrast, it is notable that a satisfactory solution to the Pierre puzzle requires respecting, not

only (NDC'), but also (NRD'). And it is only plausible to suppose that (NRD') is a principle about disagreement *and* rationality! To advocate (P5) and (P6) is, however, to advocate a view according to which there are, on the one hand, attitudes that ground rationality—the ones associated with a subject's endorsements—and there are, on the other hand, attitudes that ground public interaction—the ones associated with a subject's coordinated beliefs. But this seems to be just wrong. How could one tell apart those allegedly different kinds of attitudes in the case of (NRD')? Publicity and rationality, we have seen, are quite clearly two sides of a single doxastic coin. Coherency with these points seem to require rejection of the assumption, which underlies the analysis of (P5) and (P6), that we have to postulate different attitudes to account for the rationality and the publicity of Pierre's beliefs.

## 5. Conclusion

In this paper I have defended two main points. The first is negative. It consists of uncovering the connection between the classical Pierre puzzle about rational belief and a challenging puzzle about disagreement regarding the Antoine-Anthony case. I have argued that standard solutions to the Pierre case—either simple or more elaborated—fall short of providing a solution to both versions of the puzzle. It seems to follow that a substantial challenge still awaits the proponent of a fully satisfactory solution to Kripke's seminal puzzle.

I am more confident about the intriguing character and importance of the uncovered puzzle about rational and public belief than I am of the particular general positive answer here provided. Nonetheless, the other main point of this paper is positive. It amounts to the claim that, once the very possibility of paradigmatic disagreement is assumed, a satisfactory resolution of the puzzle can only go through via the appeal to a notion of non-belief-individuating understanding. As the remarks in the last section suggest, there might be a number of specific theories capable of providing the requisite notion of understanding. They all however would seem to appeal to the propounded non-belief-individuating notion. This paper leaves many theoretical possibilities entirely open even if, for the reasons seen in Section 4, it shows preference for accounts that are not like (P5) and (P6).

Patently, many substantial issues remain. Do the foregoing considerations encourage the revision of the analysis of other puzzles such as Frege puzzles? How exactly should the target notion of non-belief-individuating understanding be articulated? Is the target notion available to any kind of theory of belief-reports? What does the propounded kind of solution reveal about the transparency or self-knowledge of our rational beliefs? No doubt, these are all excellent questions that will have to await further research. The suggestion is that this research is worth

pursuing once the intrinsic connection between rational and public belief has been brought to the fore.

## Acknowledgements

I am grateful to an anonymous referee for this journal, for useful comments and suggestions. This work has received financial support from the Ministerio de Economía y Competitividad (Spanish government), via the research projects FFI2009-08828/FISO and FFI2012-35153, and from the Generalitat de Catalunya (Catalan government), through the consolidated research group GRECC (SGR2009-1528).

## References

- Bilgrami, A. 2012. Why meaning intentions are degenerate? In: A. Coliva (ed.) *Wittgenstein, Mind, Meaning and Knowledge. Themes from the Philosophy of Crispin Wright*. Oxford: Oxford University Press, p.96–126.
- Brown, J. 2000. Critical reasoning, understanding, and self-knowledge. *Philosophy and Phenomenological Research* **61**: 659–76.
- Burge, T. 1979. Individualism and the mental. Reprinted in *Foundations of Mind*. Oxford: Oxford University Press, 2007, p.100–50.
- . 1988. Individualism and self-knowledge. *Journal of Philosophy* **67**: 649–63.
- Chalmers, D. 2011. Propositions and attitude ascriptions: a Fregean account. *Noûs* **45**: 595–639.
- Crimmins, M. 1992. *Talk about Beliefs*. Cambridge, Mass.: MIT Press.
- Feldman, R. 2006. Epistemological puzzles about disagreement. In: S. Hetherington (ed.) *Epistemology Futures*. New York: OUP, 216–36.
- Fine, K. 2007. *Semantic Relationism*. Oxford: Blackwell.
- Forbes, G. 1987. A dichotomy sustained. *Philosophical Studies* **51**: 187–211.
- Heck, R. 1995. The sense of communication. *Mind* **104**: 79–106.
- King, N. 2012. Disagreement: What's the problem? or a good peer is hard to find. *Philosophy and Phenomenological Research* **85**: 249–72.
- Kripke, S. 1979. A puzzle about belief. In A. Margalit (ed.) *Meaning and Use*. Dordrecht: Reidel, p.239–83.
- MacFarlane, J. 2007. Relativism and disagreement. *Philosophical Studies* **132**: 17–31.
- Richard, M. 2011. Relativistic content and disagreement. *Philosophical Studies* **156**: 421–31.
- Salmon, N. 1986. *Frege's Puzzle*. Cambridge, Mass.: MIT Press.
- Schiffer, S. 1992. Belief ascription. *Journal of Philosophy* **89**: 499–521.
- Sosa, D. 1996. The import of the puzzle about belief. *Philosophical Review* **105**: 373–402.
- Verdejo, V. M. (in press.) A puzzle about disagreement. *Disputatio*.
- Wikforss, A. 2008. Self-knowledge and knowledge of content. *Canadian Journal of Philosophy* **38**: 399–424.
- Zalta, E. 1988. *Intensional Logic and the Metaphysics of Intentionality*. Cambridge, Mass.: MIT/Bradford.

V́CTOR M. VERDEJO  
 Departamento de Ĺxica e Filosofia Moral  
 Faculdade de Filosofia  
 Praza de Mazarelos s/n  
 15782 - Santiago de Compostela  
 SPAIN  
 vmverdejo@gmail.com

**Resumo.** Os casos paradigmáticos de desacordo parecem não ser compatíveis com uma espécie generalizada de solução para o célebre problema de Pierre, de Kripke. Como resultado, mostra-se que o problema clássico sobre a crença racional é também um problema sobre fenômenos de desacordo/acordo públicos. Neste artigo, defendo que a nova versão pública do problema é considerável e desafiadora, e concluo que uma solução plena às considerações de Kripke tem que oferecer uma explicação satisfatória tanto do caráter racional quanto do caráter público de atribuições de crença. Argumento então que uma noção de entendimento não individualizante de crenças é, plausivelmente, a noção central que nos permitiria evitar ambas as versões do problema de Pierre.

**Palavras-chave:** Problema de Pierre; racionalidade; desacordo; coordenação; entendimento não individualizante de crenças.

## Notes

<sup>1</sup> Just as before, we can take the belief reports to be based upon a disquotational principle—for the derivation of a subject’s beliefs from a subject’s linguistic assent—and a translation principle—for the French-to-English translation of belief reports (Kripke 1979, §2). More about these principles in short.

<sup>2</sup> The assumption that (NDC) needs to be respected is, however, one I will actually cancel when I offer a substantial reformulation of (NDC) (see Section IV below). For now, I wish to focus on the *prima facie* unobjectionable and widely accepted (NDC) in order to examine the consequences of trying to live up to in the standard ways.

<sup>3</sup> This is for instance the view of Kit Fine (2007, p.89), for whom it does not matter whether we characterize the puzzle in terms of the (theoretical) correctness or (directly) the truth of (P1) and (P2). The puzzle arises because we seem forced to accept the incompatible truth or correctness of both (P1) and (P2) as reports of Pierre’s beliefs.

<sup>4</sup> The present considerations affect both Russellian attempts to respect (NDC) by means of modes of presentation that are not part of belief-contents (i.e., analyses in terms of triadic belief-relations between subjects, Russellian propositions and modes of presentation) and the strictly Fregean theories that take modes of presentation to be part of belief-contents. The reason is that both Russellian and Fregean views agree that modes of presentation, however exactly understood, are a constitutive part and thus individuate a subject’s beliefs or belief states. In the main text, and for simplicity’s sake, I will be referring only to strictly Fregean analyses along the lines of (P3) and (P4). It goes without saying that, in any event, the propounded intensional categories or senses are typically *not* of the essentially descriptivist sort considered and rejected by Kripke himself (1979, p.259–63).

<sup>5</sup> This is not to claim that the Antoine-Anthony case is exactly like the Pierre case. Both cases are distinct even if structurally related. What is claimed is that, as far as the statement of their story goes, which is exactly the same for Antoine and Pierre-in-France and Anthony and Pierre-in-England, we are not justified in attributing differing beliefs in each case.

<sup>6</sup> Enriched propositions are a theoretical construct belonging to two-dimensional semantics. They are complex structures constituted by primary intensions (i.e., Fregean thoughts and constituent senses for present purposes) and secondary intensions (i.e., in this context, Russellian or purely referential propositions or propositional-constituents). On this account, therefore, propositional contents (and thus the objects of belief) are hybrids made out of pairs of Fregean thoughts and Russellian propositions. Nothing in the discussion of the main text hinges on Chalmers's appeal to this kind of propositions.

<sup>7</sup> This claim does not strictly speaking affect the argument in the main text but it gives a plausible rationale of what the Pierre-puzzle is really a puzzle about.

<sup>8</sup> See Verdejo (in press) for a similar argument to the same conclusion. Note that although I understand the new (version of the) Kripkean puzzle here introduced to be a puzzle about rationality and disagreement, the nature of this puzzle is quite different from the puzzles about rationality and disagreement considered in the literature on the epistemology of disagreement. The latter puzzles concern the epistemic status of our beliefs in the light of peer disagreement (see e.g. Feldman 2006 for a presentation of some of these other central kind of puzzles).

<sup>9</sup> On the assumption that cognitive value or significance is belief-individuating, the target notion of non-belief-individuating understanding must be sharply distinguished from the notion of cognitive value or significance. On this reading, understanding does not individuate cognitive significance or value either. I owe this point to an anonymous referee for this journal.

<sup>10</sup> The reader may worry at this point about the implications of (P7) and (P8) regarding a thesis about the transparency of sameness/distinctness of a subject's beliefs. How is it, according to (P7) and (P8), that Pierre believes and disbelieves that London is pretty without actually realizing this. Is it that he fails to know the sameness or contradictoriness of his own beliefs? A full answer to this question would take us too far afield. Suffice it to say however that even if (P7) and (P8) are, as I think they are, committed to a sort of failure of knowledge or transparency (of sameness or contradictoriness) of a subject's own beliefs, this may be in effect relativized and explained in terms of a subject's understandings. The key idea underlying (P7) and (P8) would be that transparency regarding one's beliefs should only be respected in cases in which the target beliefs are understood in the same way. Nonetheless, note further that one should distinguish between issues about belief-understanding and issues surrounding the privileged or authoritative knowledge of our own beliefs. More on this below.