Ascriptions in context

An Experimental Study on the Context-sensitivity of Belief Reports

Laura Natalia Pérez Pérez

Ascriptions in context

An Experimental Study on the Context-sensitivity of Belief Reports

Dissertation

Laura Natalia Pérez Pérez

*Genoveva Marti Campillo (supervisor)

Department of Logic, history and Philosophy of science

Cognitive Sciences and Language Program

Faculty of Philosophy

University of Barcelona

Barcelona

July-2012

Abstract

In this dissertation I study and empirically test whether belief attributions are context-sensitive, i.e., whether the truth value of a belief report of the form 'A believes that S' is sensitive to contextual parameters. More specifically, I examine whether the Referential Knowledge that the Audience possesses (that is, whether the hearer of a report is familiar with the name employed in it) and/or the Stakes for the agent at the time of attribution, affect patterns of attribution in a way such that the variation in them causes that a single report be correctly made in one context but not in other while nothing in the mental state of the agent has changed. To this end, I designed original experimental material and tested it on several samples of undergraduates at the Universitat de Barcelona and Universitat Autònoma de Barcelona. I here make an analysis of methodological approaches in empiric testing of some philosophical matters and offer considerations about experimentation on belief ascriptions in particular. I conclude the data gathered supports the hypothesis that there are contextualist patterns of belief attribution as regards the Referential Knowledge of the Audience parameter, but not for the Stakes variable.

Key words: *Belief, belief reporting, belief attribution, experimental philosophy, epistemology, contextualism.*

Table of contents

Preface

The purpose of this dissertation is to study and empirically test whether belief attributions are context-sensitive, i.e., whether the truth value of a belief report of the form 'A believes that S' is sensitive to contextual parameters such as the Referential knowledge that the audience possesses (that is, whether the hearer of a report is familiar with the name employed in it) and/or the Stakes for the agent at the time of attribution. To this end, original experimental material that consists of a series of written vignettes was designed.

This dissertation might be of interest to philosophers involved in belief attribution, epistemology and related matters, as well as to those who want to develop a reliable empirical methodology for testing philosophical problems or are concerned about experimental philosophy.

The initial idea behind this research was presented to me by Professor Genoveva Marti, as well as many insightful comments and discussions that made this work possible. I must say that I'm really grateful to her, among many things, for agreeing to tutor a psychologist with a project like this and for her incredible patience with my lack of philosophical background. Without her dedicated teaching and support I'm completely sure I wouldn't have been able to present this dissertation today.

I am also thankful to Professors Albert Costa (Universitat Pompeu Fabra), Ana Gavarró (Universitat Autònoma de Barcelona), Nat Hansen (University of Reading) and Ángel Pinillos (Arizona State University) who contributed to the design of the vignettes with their very useful comments on preliminary experimental material that finally led me to the tests here presented.

The experiments were applied on several samples of undergraduates at the Universitat de Barcelona and Universitat Autònoma de Barcelona. I would like to thank Professors Laura Bosch (Universitat de Barcelona), Marta Campdelacreu (Universitat de Barcelona), Gemma Celestino (Universitat de Barcelona), Fernando Gabucio (Universitat de Barcelona) and Manuel Pérez Otero (Universitat de Barcelona), for their kind help providing me with the samples for the study.

I also have great gratitude to all the Logos people for letting me part of their stimulating and welcoming group; especially, I thank Gemma Celestino who was

not only a great help during my time as a T.A. and discussing philosophical stuff of this work with me, but also became a kind and reliable friend of mine.

Also, I thank all the amazing friends that during these years made my stay in Spain a very memorable time of my life, particularly Aleja, who helped me with the language revisions when I needed, and César, who helped me with informatics stuff when I got stuck with Word. It was also very important to me to have the love and support of Eme, Ana, Angie, Rafa and Marisol.

Finally, I would also like to thank my parents, Esperanza and Teodoro, and my beloved sister, Sonia, who had to learn to have me away all these years.

Chapter One

The issue of context-sensitivity of belief reports and its origins

1.1. The issue

Suppose the great roman orator Cicero is known in America as 'Cicero', but in the United Kingdom he is best known as 'Tully'. Now suppose I have a good friend, Lily, who is an American philosopher and a perfectly rational person, who thinks very highly of the philosopher and believes that he is a great orator. As I have travelled and know well the philosophical environment in both countries, I know that 'Cicero' and 'Tully' are both names for the same orator, but in the US, people, including Lily, commonly do not know this. They generally only use the name 'Cicero' to refer to the roman orator without associating other names for the man. A similar situation occurs in the UK with respect to the name 'Tully'.

One day I am with Daniel, a friend from London. We are having a discussion about oratory and he says: "There are not good Roman orators, Greek orators are much better". I want to argue that he is wrong because my friend Lily, whose opinion we both value and recognize, believes otherwise, so I tell him "I think you are wrong, Lily believes that Tully is a great orator". This attribution I make in this situation seems not only acceptable but quite literally true.

Now imagine I am in a different situation: One day I gather with Greg, an American colleague. We are having a discussion about oratory and he says: "There are not good Roman orators, Greek orators are much better". I want to argue that he is wrong because my friend Lily, whose opinion we both value and recognize, believes otherwise, so I tell him "I think you are wrong, Lily believes that Tully is a great orator". Perhaps, the attribution I make in the second context does not seem correct to the reader; to some, it might even seem strictly speaking false[1]. One can even imagine that if Lily had been in

---

[1] Stating that the attribution above is "quite literally true", or that the latter might seem "strictly speaking false" might sound too strong to some. As I will explain later, in the discussion of cases of this sort there is still the question as to whether they are showing a genuinely semantic phenomenon or a pragmatic one. For this reason, and to remain neutral about the treatment of the data offered, in this dissertation when

the office with me and Greg at the same time and she had overheard what I reported, she may have responded puzzled: "What are you talking about? I don't even know who that guy is!" [2]

The attribution in the latter scenario can probably be perceived as less natural. It might also seem that Lily, ignorant of the fact that 'Tully' and 'Cicero' are names for the same person, could rightly refuse to accept the attribution that she believes that Tully was a great orator. Maybe it happens that one feels impelled to judge the attribution in this scenario as inadequate, but no matter how impelled one might feel or whether it seems right to judge that the attribution is incorrect, it might also look suspicious.

In essence, it looks like what I am attributing to Lily is the same in both scenarios: that she believes that Tully is a great orator; and that nothing about her that might be relevant for the attribution has been modified from one situation to the other. Observe the fact that Lily's mental state[3] remains stable through both hypothetical situations where the report is uttered (she could have even been sleeping while I make the reports, for instance, so her mental state stays invariable). In fact, the only thing that has changed between scenarios is the audience, more specifically, the fact that Daniel is only familiar with the name 'Tully' and Greg with the name 'Cicero' to refer to the roman orator. Hence, it may seem puzzling that the same attribution of a belief to Lily appears to be correct in one context but not in other while nothing has changed in her, and only features of the context have varied.

In principle, it would seem that the truth value of the report that Lily believes certain content should only depend on her, the agent, and the relation she holds with respect to that content. Also, it is generally accepted in the philosophy of language that belief is a relation between an individual and a proposition, and that belief ascriptions are just reporting that relation. That being so, the apparent shifting phenomenon observed in the case above needs to be accounted for. Certainly, the mainstream intuitions are that one would correctly attribute in one context but not in the other; altogether, it seems

referring to cases like the one here presented I will be talking in terms of correctness and incorrectness, instead of talking about truth and falsity. Also for the sake of neutrality, 'correct' and 'incorrect' should not be read as implying a pragmatic approach to the matter.

[2] The example here presented is modelled on the basis of and is similar to the kinds of examples that have been used by philosophers to motivate the contextuality of belief reporting. Some of these examples can be found in section 1.3 of this chapter.

[3] Whether belief is something that is in the mind, as I will explain later, is quite controversial. When I here refer to an agent's mental state do not imply that it is mental in the sense that is in their head, rather, I intend to refer to her psychological, physical and other relevant attitudinal aspects.

plausible that correctness of attributions of belief can change across contexts when nothing relevant about the attributee's mental state changes. Here I will be calling intuitions that go in accord with this pattern, *parameter-sensitive intuitions*[4].

This sort of variability of intuitions has not only been identified in cases of belief attribution. In fact, many philosophers (Perry & Crimmins, (1989); Richard, (1990); Wettstein, (2004), Cappellen and LePore, (1997); DeRose, (1992); Stanley, (2005); among many others) have suggested in different manners the idea that the truth value of a report might be affected by features of the context. For instance, in areas such as indirect speech and attributions of knowledge and desires, among others, it has been proposed that a single ascription can be false in one context and true in other, even if the agent remains in the same mental state.

This idea that certain reports might be context-sensitive has become stronger in the past few years, to the point that it has become the centre of very rich discussions in philosophy. The attractive that such theories might have, consists mainly in that they seem to offer a plausible and natural explanation for certain data that allegedly represent our common linguistic practices of attribution. These data are the intuitions that a single report can be correct in one context but not in other when the agent's mental state has not varied. In order to elicit and make explicit that these intuition patterns exist and correspond with our ordinary linguistic behaviour, numerous *context-shifting cases* have been introduced in the philosophical literature, i.e. possible scenarios (like the one presented above and others that will be offered later on) where it seems evident that not only mental states of the attributee but also the variation of contextual features, can influence our attributions.

For instance, John Perry (forthcoming), in his most recent discussion of the semantic nature of names, develops an approach of the difference in cognitive value between identity statements where he suggests that the interests of the audience of a report are relevant in establishing, between two co-referential names, which one to use to make correct attributions of assertions. The example he gives to motivate his claim is the following:

---

[4] I do not want to imply here these intuitions necessarily support context- sensitivity; in fact, they occur regardless of contextualists being right or wrong about semantics of belief reporting. Actually, there are non- contextualist explanations of them which I will explain further on.

Following the energetic Ivan's lead, a group of linguists and philosophers headed towards a conference in San Sebastian have boarded a bus. Some of us begin to worry if Ivan led us to the right bus. I ask him, and he says, "This bus is headed to San Sebastian, I'm sure of it." Reporting back to the others I might report:

(*) Ivan said that this bus is headed to San Sebastian

(**) Ivan said that this bus is headed to Donostia.

Surely (*) is the most natural thing for me to say, especially if Ivan has never heard of Donostia, or has but doesn't know that Donostia is San Sebastian. However, suppose that although Ivan, on his trip to the Basque Country, doesn't realize that San Sebastian is Donostia, the rest of us do; we are veterans of these conferences, and are in the habit of using 'Donostia' once we have arrived, since the local Basques prefer that name. Given that I'm conveying what Ivan said as a way of reassuring the veterans that the bus is headed the right place, I might use (**). (Perry (forthcoming), p. 22-23)

Similarly, Cappellen and LePore (1997) propose that sensitivity to contextual variables seems to affect indirect discourse. Particularly, they propose that the acquaintance the audience has with the name used in the ascription – i.e. whether they are familiar with the name and use it competently or not- is a factor that influences reports of what an agent says. These are two of the many examples given by the authors:

Suppose George says, "John leaves for Berkeley next week". It is not incorrect to report George to someone unfamiliar with Berkeley, seeking a ride to Northern California, with:

George said that John's going to northern California next week. (Cappellen and LePore (1997), p. 285)

A second example:

Françoise: Chartreuse is Maria's favourite colour.

> Someone, knowing his audience is unfamiliar with the word 'chartreuse', employs another means for conveying what Françoise said. Demonstrating a chartreuse dress, she reports
>
> Françoise said that *the colour of that dress* is Maria's favourite colour. (Cappellen and LePore (1997), p. 283)

In addition, there has been a lot of discussion about the possibility that knowledge and belief ascriptions be context-sensitive. As regards the possible context-sensitivity of belief reports (from now on in this dissertation, BRs[5]), that is, the likelihood that the correctness of ascriptions of belief might be influenced partly by contextual features, so that a single attribution can be correct in one context but not in other, it is the main topic of this research and the central matter of this chapter. Besides, given the connection between knowledge and belief, I think that the discussion about knowledge ascriptions may be relevant for the debate around belief reporting. Therefore, I will devote the second chapter of this dissertation to recent discussions on the apparent sensitivity to context of knowledge reports.

The issue of the possible sensitivity to context of BRs has come to the forth connected to a discussion in the philosophy of language about the apparent failure of the *principle of substitutivity salva veritate* of co-referential terms (abbreviated in this dissertation as PS) in propositional attitude ascriptions[6]. In the next section of this chapter I discuss this topic, which is the background that allows me to concentrate on contextualist theories of belief attribution in Section 3. By contextualist theories I mean theories which endorse in different ways that the truth or falsity of a BR depends to some extent on the context, so that a single attribution can be true in one context and false in other while nothing relevant has changed in the attributee's mental state.

I hope the background I will give in these last two sections will make clear the motivation and allows me to formulate (in Section 4) the ultimate aim of this research which consists in, first, empirically confirm whether parameter-sensitive intuitions about attributions of belief are shared by folks, and second, focus on some specific

---

[5] BR for the singular.

[6] It could be said that historically, the attempts to find a semantic theory that accommodates to data on propositional attitude reports, allowed for the contextuality talk to arise. I don't intend to go deeper in these aspects that concern history of philosophy as I don't believe this adds to the discussion here presented.

contextual parameters to determine if they have an influence in our attributing behaviour. The main bulk of the dissertation (Chapters 3 to 6) is devoted to that aim.[7] In Chapter 3 I discuss a few approximations in experimental philosophy to testing variability of truth value of reports in epistemology, as well as other related empirical works, with the hope that the discussion around those experiments helps my purpose of offering reliable tests for belief attribution. My specific empirical proposal can be found in Chapter 4, as well as the results gathered by the implementation of it. Finally, the analysis and conclusions about the data and methodology can be found in Chapter 5.

Let me now explain how the worry about possible contextuality of BRs comes about.

1.2.*Some history*

1.2.1. *The Fregean approach to belief reports*

Frege (1892) offers one of the most influential semantic theories of *Indirect (oblique) Contexts*. Indirect contexts are the *that-clauses* of propositional attitude ascriptions, sentences which we embed in another sentence in order to talk about what someone *says, believes, or knows, etc.* Ever since Frege, propositional attitude reports have drawn the attention of many authors resulting in a vast amount of literature that aims to provide an intuitively correct account of our reporting practices.

One of the reasons why this topic became an object of study for so many philosophers is the need to respond to the apparent failure of substitutivity of co-referential terms in indirect contexts[8]. Frege claims that substitution of co-referential terms in reports of propositional attitude such as belief reports can intuitively turn a true report into a falsehood[9]. Belief reports are utterances of sentences of the form '*A* believes that *S*' where *A* is an agent of whom we say that he believes the proposition expressed by the sentence *S*. To appreciate Frege's worry, consider for example

---

[7] More details about the contextual parameters to test will be discussed in section 1.4 of this chapter.

[8] In the presentation I make here, I narrow my focus to only BRs, but it is important to remark that an important part of the discussion on propositional attitude reports in general is based on the discussion about belief reporting. For this reason I will be using the verb 'believe' in the explanation of the problem and concentrating in the discussion about BRs, but the reader should bear in mind that the problem I present in this section is general for propositional attitude ascriptions. I must also say however, that puzzles of belief are different –although in some cases related- from problems that arise for other propositional attitudes, and of course theories are very diverse, as it will be clear later.Work on other propositional attitudes is extensive and I cannot afford to discuss it here, but interesting bits of those debates will be brought up when necessary for explaining the issue.

[9] Belief reporting can be very complex, but here I'll try to stick to the simplest case possible (which is already complicated enough).

(1) Nana believes that Phosphorus is a star visible in the morning

(2) Nana believes that Hesperus is a star visible in the morning.

You need to suppose, for instance, that Nana is an ancient Babylonian who does not know 'Hesperus' and 'Phosphorus' are both names to refer to the same heavenly body – the planet Venus- which is visible in the morning and in the evening as well. When she sees the planet in the morning, she calls it 'Phosphorus', but when she sees it in the evening she refers to it as 'Hesperus'. It has never occurred to Nana that *"both"* stars could actually be just one.[10]

According to PS, when two proper names share the reference, substituting one of them with the other in a statement should preserve truth. Therefore, switching 'Phosphorus' with 'Hesperus' in (1) –the result of this procedure is sentence (2) - must not alter truth value. This is a very intuitive principle[11]; in fact, it might seem that independently of the substitution what we are saying with sentences (1) and (2) is that a certain agent believes something about a certain matter which is the same in both of them. However, the example seems to show that an utterance of (1) is perfectly acceptable and true while an ascription like (2), apparently is false.

In fact, according to Frege, the embedded sentences in (1) and (2), namely,

(3)   Phosphorus is a star visible in the morning

(4)   Hesperus is a star visible in the morning

don't express the same proposition, even if they only differ in that different co-referential names occur in the same place in their structure. 'Hesperus' and 'Phosphorus' have an additional semantic feature apart from the referent, the *sense*.

Senses are modes of presentation that determine the reference of an expression and are constituents of the propositions expressed and therefore of the beliefs attributed. In Frege's account, senses capture differences in cognitive value between sentences (3)

---

[10] Here, I restrict the example to proper names, though the worry can be extended to other co-referential terms.

[11] Names seem useful for speaking about things in a way that definite descriptions are not. With definite descriptions we can mistakenly designate someone that we do not intend. Instead, the proper name seems to always select the individual I intend to speak about. Now, if one wants to construct a theory that respects this basic fact about names, the most natural way to do it is to say that names exclusively refer. Frege defends there is problem with this which is evident in that pares of co-referential names like 'Hesperus' and 'Phosphorus' differ in cognitive value, as I'll explain next.

and (4), namely, they help to explain why it is possible for competent speakers to accept (3) without accepting (4). For him, reference alone could not account for this phenomenon, a mode of presentation of that reference was necessary to explain it. Frege understands senses are fundamental parts of the meaning of proper names, and concludes that if two co-referential names occur in sentences that have the same structure but different cognitive value, those names must also have different senses. 'Hesperus' and 'Phosphorus' have different senses, and therefore, according to Frege (3) and (4) express different propositions; different propositions with different truth conditions.[12]

Now, belief, as a propositional attitude, is understood as the mental state of having a relation towards the proposition expressed by the sentence embedded under the scope of 'believes that'. For Frege belief is a binary relation between an agent -the believer-, and a proposition. Fregean propositions are thoughts, abstract structured entities constituted by senses[13]. Thus, to believe that Phosphorus is a star visible in the morning, for example, is for an individual to stand in a relation to a representation[14] that is the propositional content of 'Phosphorus is a star visible in the morning'. By the same token, BRs are just a formulation of that relation; when we make a belief ascription we are making a claim about the binary relation. In virtue of this, ascriptions (1) and (2) are claims about a relation Nana bears to the propositions expressed by the that-clauses in them.

Based on his understanding of proper names, and to account for the apparent failure of the PS observed in the intuitive difference in truth value between (1) and (2), Frege proposes that proper names change their reference when they occur in subordinate clauses of a BR. Under the Fregean view, when proper names such as 'Hesperus' And 'Phosphorus' appear in that-clauses of propositional attitude reports, they do not denote their usual reference but have an oblique one.

When the names 'Phosphorus' and 'Hesperus' occur in embedded clauses, they are only *indirectly* about the planet Venus. In indirect contexts, the proper names refer to the

---

[12] That meaning is compositional means that the meaning of a complex expression is determined by the meaning of the simple expressions that compose it, according to the way these are combined.

[13] As I will explain later, this is different from other approaches to propositions.

[14] I will stay neutral here about this issue, so I want to make clear that when I refer to belief as a mental state I don't intend to convey that it is mental in the sense that it is in the head. I leave it open that it'd be representational, exclusively in the head, or composed by other dispositions such as behaviour, emotion, etc.

mode of presentation of the planet, consequently, (1) and (2) are not bound to have the same truth value. When I utter (1) I am making a different claim than when I utter (2), because with each report I'm claiming that Nana is in a relation with different propositions. This solves the problem that the apparent failure of PS poses for semantics of proper names and propositional attitude ascriptions.

That belief is a binary relation is commonly accepted by many as linguistically plausible and fits well into Frege's semantic picture, but problems arise for other views that reject a Fregean approach to the semantics of proper names and other details of the Fregean theory. I will next present non-Fregean treatments of belief reporting to illustrate how the idea that BRs are contextual originated.

1.2.2. Direct Reference

According to theories of *Direct Reference* (hence *DR*, also called *Millianism*, in honour of John Stuart Mill who suggested a similar idea) names refer directly, with no mediation of senses or other components of any type. Proper names always refer to their bearer, even if they occur in sentences that are embedded in other sentences of any kind. 'Hesperus' and 'Phosphorus' both always refer to the planet Venus directly, with no mediation, and the semantic contribution of a proper name to the propositions expressed by utterances of the sentences where it occurs, is an individual, the referent.

According to Direct Referentialists, the Fregean explanation of the apparent failure of PS based on oblique references contradicts the intuitive fact that "Tully was a great orator" and "Cicero was a great orator", (or in the other hand "Hesperus is a star visible in the morning" and "Phosphorus is a star visible in the morning"), reflect the same thing about the world, have the same truth conditions and express the same proposition. It seems that when you talk about Cicero, you also talk about Tully for the same reason that if you hit Cicero you have also hit Tully (or, if you have landed on Hesperus, you have also landed on Phosphorus).

For theorists who hold DR, the fact that two sentences with identical structures which vary only in that they contain different co-referential names seem to talk about the same thing, is relevant in assigning propositions to sentences. For instance, utterances of sentences (3) and (4) express the same proposition, because the only contribution that the names make to the proposition expressed by (3) and (4) is the planet Venus. As the

contribution the names make is identical, and everything else in the structure of these sentences is also equal, the propositions expressed should also be the same.

DR theorist accept the Russellian conception of propositions according to which propositions are structured, complex entities (n-tuples) with individuals, relations or properties as their constituents; no senses play a role in this view[15]. Most DR theorists accept Frege's main idea that belief is a relation between an agent and a proposition, but to them the second element of the relation is not a Fregean proposition. To direct referentialists, belief is a relation between an agent and a Russellian proposition-, a structured entity that may be constituted by individuals.

To sum up, DR theorists hold that the only relevant semantic feature of a proper name is its referent, with no senses or other components to mediate between the name and the individual it refers to. This has the consequence that then sentences (3) and (4) must express the unique singular proposition that Phosphorus is a star visible in the morning. Accordingly, believing the proposition expressed by (3) or (4) is for an agent to stand in a belief relation with this single proposition indistinctively.

An important consequence of this way of regarding the matter is that if one wants to stick with the definition of belief ascriptions as being just reports of the belief relation, then one needs to account for the apparent failure of the PS in sentences such as (1) and (2). If the names 'Hesperus' and 'Phosphorus' always refer to the same entity, they should be intersubstitutable preserving truth, because they presumably maintain semantic content (namely the proposition expressed). Under this optic, truth conditions of (1) and (2) should be equal: in all possible situations in which (1) is true, (2) must also be so. According to this view, when you say that Nana believes that Hesperus is a star visible in the morning, and that is true, your report that Nana believes that Phosphorus is a star visible in the morning must also be true. Substituting the names in the report should not alter truth value.

The upshot that co-referential names ought to be intersubstitutable *salva veritate* is regarded by many as implausible and a challenge for DR. In words of Theodore Sider (1995): "Some contemporary Russellians, defenders of the view that the semantic

---

[15]   Direct Reference is a thesis about the reference of terms, while Russellianism is a view about propositions. Most defenders of Russellianism are also Direct Referentialists, however, it is important to note that there can be theorists that sustain one, but not the other or vice versa.

content of a proper name, demonstrative or indexical is simply its referent, are prepared to accept that view's most infamous apparent consequence: that co-referential names, demonstratives, indexicals, etc. are intersubstitutable *salva veritate*, even in intentional contexts." (Sider, 2005, p. 1). Therefore, if the direct referentialist wants to deal with the apparent problem of substitution, he must explain the intuitive difference in truth value that we find in sentences like (1) and (2).

A type of DR account that offers an alternative to explain away the apparent problem with intersubstitution is the *Pragmatic* view or *Naïve Russellianism*, sustained by Nathan Salmon and Scott Soames. To them, BRs express a binary relation between an agent and a proposition where propositions are not Fregean but Russellian. On this view, the fact that one intuits that (1) and (2) are different in truth value, is due to *pragmatic* features of the sentences and it is not a genuinely semantic phenomenon. It is not that the propositions expressed by these sentences actually differ in truth value, but that we are mislead in our intuitions of truth and falsity due to other non-semantic components of language. According to this proposal, the variable intuitions towards (1) and (2) show that sometimes people have difficulties assessing the truth value of sentences because there are pragmatic traits of the language that interfere with semantics.

Inspired by the Gricean idea of implicature, the authors propose that in utterances of propositional attitude ascriptions there is a difference between the proposition semantically expressed and some collateral information conveyed by the utterance of each sentence in a conversation. This collateral information has an important pragmatic conversational function that serves communication, and for that reason it makes it difficult for ordinary users of the language to determine with clarity the truth and falsity of statements.

The problem, as the Pragmatic view holders describe it, is that belief is in effect a relation between an agent and a singular proposition but it arises in connection to attitudes towards sentences (like accepting or uttering). The intentional relation of believing or not is individuated by an object, therefore if Nana believes the proposition expressed by (3) she must also believe proposition expressed by (4). However, even though propositional attitude ascriptions report relations to propositions which are exclusively responsible for their semantic value, they also entail relations to other

information which serve conversational purposes. Given its pragmatic use, that information *appears* to be a relevant part of the meaning, even if it is not.

Furthermore, Soames argues in general about indirect contexts that "Sentential attitudes are often more significant for explaining action than propositional attitudes" (Soames, 1988, p. 57) and so, given the pragmatic relevance of those sentential attitudes that appear as collateral information, to the unaware listener truth and falsity of singular propositions can mix up giving as a result anti-Russellian intuitions.

In summary, the core idea is that we confuse semantic with non-semantic information, and that this confusion is responsible for our divergent truth value assessment in pairs of sentences like (1) and (2). This view endorses that sentences (1) and (2) report a relation between the believer and a unique proposition, and that the apparent difference in truth value between (1) and (2) is (purely pragmatic) not semantic. But we can easily mistake what is pragmatically conveyed with what is semantically expressed: we are inclined to see a difference in truth value where there is actually none, the authors explain. In Soames' words: "The general thesis, then, is that the substitutivity principle […] is correct; and that resistance to it is based on a failure to properly distinguish the semantic information expressed by a sentence relative to a context from the information conveyed by an utterance of it in a given conversation" (Soames, 1988, p.220). In the end, what this view says is that (supposing report (1) is true) (2) is a true but misleading report of a belief Nana does indeed have.

Even though this explanation seems reasonable, many have criticized it. After all, we still have the intuition that both reports differ in truth value even though they only vary in that they contain different co-referential names. A common reaction towards the pragmatic strategy can be summarized in Mark Richard's words:

> …other than using bribery, threats, hypnosis, or the like, there is simply nothing you can do to get most people to say that Jones believes that Tully was an orator, once they know that Jones sincerely denies 'Tully was an orator', understands it, and acts on his denial in ways appropriate thereto. In particular, pointing out that Jones can express something he believes with 'Cicero was an orator' seems simply irrelevant to most people… The Russellian is correct when he says that our intuitions about

truth conditions are not wholly reliable. But they are certainly not to be ignored. (Richard (1990), p. 125)

If recurring to pragmatics does not serve to account for our intuitions about truth value of sentences (1) and (2), the DR theorist is required to find new strategies that explain the apparent failure of the PS.

1.2.3. From the substitutivity problem to the possible context-sensitivity of Belief Reports

Apart from the Pragmatic view, other possibilities for explaining come into the picture for DR theorists who don't want to dismiss the intuition that (1) and (2) vary in truth value. One alternative comes from views that endorse there is an important component that exceeds the language in semantics of propositional attitude reports. These are views that endorse context-sensitivity of BRs, approaches that consider that features of the context of utterance can be crucial to assessing truth value of belief attributing sentences. Theories that advocate for context-sensitivity of BR usually offer an explanation that accommodates the apparent failure of PS; however, this is not the problem they mainly intend to elucidate.

As I've been explaining, if we regard the substitution problem from a Russellian point of view, it seems to follow that Nana has contradictory beliefs. The apparent falsehood of (2) leads us to the conclusion that Nana does not believe the proposition that Phosphorus is a star visible in the morning, while from the apparent truth of (1) we derive that Nana believes the proposition that Phosphorus is a star visible in the morning. Therefore, it seems that with our reports we are saying that Nana believes, and at the same time does not believe the same singular proposition: that Phosphorus is a star visible in the morning. This presents a problem strongly connected to the apparent failure of substitutivity: that it'd be possible that an agent be reported as believing *p* and as not believing *p* at the same time. Take the following case as an example of this:

> Mutt and Jeff agree on what sentences Odile accepts. They agree about her dispositions to behavior. They agree on just about everything which seems relevant to the question, does Odile believe that Twain is dead?
>
> They don't agree on the answer. When Mutt was asked, it was because someone wanted to know whether Odile would list Twain under dead

Americans. Mutt knew she accepted 'Twain is dead' and thus said 'yes'. Jeff was asked by someone who couldn't understand why Odile, who's pointing to Twain's picture, wants to meet him. Doesn't she realize that Twain is dead? Jeff knew she rejected 'he's dead'. He answered that no, Odile didn't believe that Twain was dead. (Richard (1989), p. 317)

In this case, Odile is reported both as believing and as not believing that Twain is dead.

Also, take the case presented to introduce this chapter where the sole report that "Lily believes that Tully is a great orator" seems to be correct in one context but not in other. It is just one report, so in principle it must always express just one proposition (either Fregean or Russellian). No substitution occurs here, but still an apparent change in truth value shows; a single report is used to say something apparently true in one context and something false in other. I could as well take the example to a level similar to Odile's case before, and imagine that if someone asks whether Lily believes that Tully is a great orator in a context where it is very important that I be faithful to the agent, I may report correctly that no, she does not believe that Tully is a great orator, for the reason that Lily herself would refuse to accept this report of her belief. On the contrary I would happily report that she does believe so to a British audience that would only understand the proper name 'Tully' for the orator, just as shown in the example given above. I therefore would be explicitly reporting of her that she believes *p* and disbelieves *p* when nothing relevant to the ascription has changed about her circumstances. Again, we are therefore either forced to admit that a presumed rational agent holds contradictory beliefs, or required to find a semantic theory that offers an explanation of the apparent context-shifting phenomena, respecting our intuitions about truth value of ascribing sentences.

Cases like these seem to suggest that BRs are more problematic than Frege thought. To Frege, propositional attitude reports represented a problem to semantics because substitution of co-referential names in them apparently did not preserve truth, but cases like Odile's and Lily's seem to show there is more to be explained than apparent failures of the PS. Observe these cases are *beyond substitution*, no change in words has occurred here for the utterances of the report to apparently shift truth value. For instance, what we see in Lily's case is that, precisely, it is *not* substituting what creates the apparent change in truth value between the attributions in the two contexts.

Furthermore, there are cases in which we want to substitute *precisely* to preserve truth value, something that is not accounted by Frege. Take the following as an example of this:

> Imagine that Emily doesn't know that the famous actress, Marilyn Monroe, was none other than her childhood friend Norma Jeane Mortenson. She assumes that Norma spent her life back in the small town in California where they grew up together, raising hers and Jack's (her high school boyfriend) children. Norma Jeane, she says, was certainly no famous actress or anything; she was rather ugly, a bit dumb and couldn't really act or sing. After they left high school they never saw each other again. Years later, Emily has become a huge fan of Marilyn's acting and one of her big dreams is to have a copy of one of her films signed by her. She knows very well that Marilyn was a worldwide recognized icon.
>
> Once, by coincidence, Marilyn and Emily come across each other in their born town. Marilyn is staying at her parents' house as she is going through a rough moment. When the two old friends met, Marilyn was normally dressed, was not wearing makeup and has not dyed her hair for months, so she is just like any other brunette. Emily fails to recognize her and did not ask her for an autograph.

Consider these reports:

(3)   Emily believes that Marilyn Monroe is an amazing actress
(4)   Emily believes that Norma Jeane is an amazing actress

Intuitively, if I were to substitute 'Marilyn Monroe' by 'Norma Jeane' my report would apparently turn into a falsehood. However, imagine I am speaking to Marilyn's parents who are watching from afar the re-encounter between the two friends and feel both estranged and sad seeing that Emily has not asked Marilyn for an autograph. I want to console them and reassure them that Marilyn is still an idol for Emily. In that case my use of (6) does not seem false, but adequate and true.[16]

---

[16] This case is modelled to make evident correct intersubstitution, based on a case coined by Wettstein which I present later on in subsection 1.3.3.

These data about BRs are problematic for Naïve Russellians as much as for Fregeans: if it is just one report we are talking about, no difference in senses can be blamed for the difference in truth value, as the sense of 'Hesperus' is eternal independently of the context, and so is the sense of 'Phosphorus' or any other proper name. On the other hand, that substitution can preserve truth value is difficult to explain for the Fregean, as in indirect contexts the names involved in the examples are not co-referential. Exchanging one for the other would be a mistaken and infelicitous move. In turn, the explanation recurring to pragmatics seems more difficult yet to accept.[17]

Views that endorse context-sensitivity, instead, are attractive for the reason that they seem to deal well with the cases as they particularly endorse that different utterances of the same report, can vary in truth value depending on the context of use. This is the distinctive proposal of this strategy, and each author holding this view develops his own explanatory framework with the aim of delivering this result in different ways. Also, although it is not their main focus, each view makes its own attempt to offer a solution for the problem of substitutivity, applying in their specific ways the idea that attributions of belief are sensitive to contextual features. In general terms, a view of this sort would be appealing as regards this problem if endorsing that there are contextual features that might be relevant for establishing truth value of propositional attitude reports, would allow (3) and (4) to express the same singular proposition while keeping that (1) be true while (2) be false. The way this result is achieved though, varies from one theory to the other.

1.3. Theories that endorse context sensitivity of Belief Reports.

In this section I summarize three accounts of belief reporting that endorse in different ways the context-sensitivity of BR: Perry and Crimmins', Mark Richard's and Howard Wettstein's. As we will see in what follows, most theorist who endorse that truth value of belief attributions is sensitive to contextual parameters, sustain names are Millian, but modify the approach to belief in order to explain the evidence of the apparent semantic change of attributions through scenarios. The proposals here presented for such an account come from three directions: (A) Either they modify the accepted characterization of belief and expand the binary relation to comprise other elements that

---

[17] See more about the challenge that contextuality might represent for Fregeanism and Naïve-Russellianism at the end of this chapter.

play a semantic important role in belief reports, where the extra elements in the belief relation are sensitive to the context (Perry and Crimmins), or (B), they try to maintain a binary belief relation characterizing the second relata as composed not only by a proposition but further truth-relevant elements over which features of the context impose restrictions (Richard), or (C), they reject completely the representational picture of belief therefore objecting to the whole conception of belief as a relation between an agent and a proposition, and allowing reports to depend on the context of utterance (Wettstein).

1.3.1. Perry and Crimmins' hidden indexicality

John Perry and Mark Crimmins offer a general account for the treatment of propositional attitude reports that maintains that propositions are singular and names are Millian, while being faithful to our intuition that (1) and (2) have different truth values. To them, propositional attitude reports involve collateral information that is implicitly referential and therefore affects truth value of these sentences, so that when I claim that someone is in a relation to a certain proposition I am thereby also referring implicitly to that information. I will explain the fundamentals of this proposal.

Perry and Crimmins (1989) introduce the idea that BRs might be context-sensitive. They focus on explaining how an agent can be reported to believe and disbelieve the proposition that *p* at the same time, when both reports seem truthful. In their strategy, they reformulate the belief relation to add a third element that (at the time of utterance) is affected by contextual features. Thus, the authors suggest that context plays a significant role in defining the truth conditions of belief reports.

To understand what Perry and Crimmins propose, let's go back to the Tully/Cicero example and suppose that Lily has heard about Cicero with the name 'Cicero' in her oratory courses and has heard of him with the name 'Tully' in her political philosophy class, but it so happens that she doesn't relate them at all, she doesn't realize that Tully and Cicero are the same individual. She doesn't (apparently) believe, therefore, that Tully is a great orator. Following the story, we can say that Lily has two beliefs with different origins and manifestations: i) that Cicero was a great orator, caused by the acceptance of information in the oratory classes she has attended, which causes her to affirm "Cicero is a great orator". ii). That Cicero was not a great orator, caused by her

acceptance of the information in the political philosophy classes she has attended and which might cause her to affirm "Tully is not a great orator".

In this example, a single attribution -namely "Lily believes that Cicero is a good orator"- can be used to say something true in determined contexts but also to say something false in others, while Lily's state of mind is kept unaltered (she has both beliefs (i) and (ii) constantly and nothing relevant about her circumstances has changed). Consequently, something that does not concern the agent's mental state must vary for the attribution to change truth value, claim the authors.[18] The proposal they make, roughly, is that what changes between (1) and (2) is what we are talking about: Lily has two different *notions* of Cicero and each plays a distinctive role in her beliefs. For a more complete explanation of the issue, I will offer some definitions that are relevant for grasping the characterization of BRs the authors make.

 To Perry and Crimmins, beliefs are concrete cognitive structures (things in the head), particulars that hold complex causal relations to an agent's perceptions, actions, abilities, and other cognitive structures. According to the authors, beliefs have *ideas* and *notions* as their constituents. Ideas are representations[19] of properties and relations, and notions are representations of individuals; they are concrete cognitive particulars whose content helps determine the content of beliefs. The content of an idea or notion is determined by its external properties such as facts about its origin, nonetheless, "The content of an idea is not always fixed once for all by facts about the circumstances of the idea's origin. *Some ideas are context-sensitive, in that their contents change with changes in the agent's circumstances*". (Perry and Crimmins, 1989, p. 691. Emphasis added).

---

[18] Perry and Crimmins' (1989) original example they present is the following: In Mark Twain's The Prince and the Pauper, Tom Canty and Edward Tudor decide to exchange lives for a day, but for reasons beyond the boys' power, the interchange lasts much longer than planned. An intelligent reader of Twain's tale could be able to describe the important aspects of it with certain confidence. Such reader might explain why Miles Hendon -a nobleman who encounters the boy poorly dressed- does not bow to the prince, by saying

> Miles Hendon did not believe he was of royal blood.

And the reader could continue saying that Miles Hendon was foolish by noting

> Miles Hendon believed Eduard Tudor was of royal blood.

The authors want to argue both claims by the reader are true.

[19] I want to mention that the authors do not commit themselves to any specific characterization of mental representations.

This context-sensitivity of ideas is said to hold a close relation to indexicality[20] in that there is a semantic role of ideas that is a function determined by the agent's circumstances. The content of an idea can be determined by external conditions. This can happen in two ways: Origin-sensitivity and context-sensitivity. When the facts surrounding the origin of an idea fix its content once and for all, we talk about origin-sensitivity. This is the case of Lily's idea of red, for instance. When otherwise, the semantic role[21] of the idea is sensitive to changes in the agent's circumstances making its content vary from occasion to occasion; here we talk about context sensitivity. An example of this kind of variation is Lily's idea of past, which stands for different properties as Lily's life progresses. An idea may exhibit one type of sensitivity, the other, or both.

A notion, like an idea, is something in the mind that stands for things in the world, particularly, it stands for individuals. That two beliefs are about the same thing means they share notions or contain notions that are linked. In occasions, an agent can have several notions of a single individual. One reason this may happen is misrecognition, i.e. the agent doesn't link or connect the two notions she has of the same individual so she fails at identifying that both are notions of the same individual.

Perry and Crimmins' core idea is that belief reports are a claim that an agent has a belief with certain content, but there is also a further condition that the belief must meet when we report it. This view endorses that when we make a BR we affirm that an agent a has a belief with a certain content *p*, but we are saying more than that, there are constituents of the report that are not manifest but that contribute to the meaning and therefore take part in the truth of the report. The third element that is part of what is said is not evident but tacit, doesn't appear explicitly mentioned in the report but is an important constituent of its meaning: is an unarticulated constituent of the report. We are talking about the notion the belief report concerns, in the authors' words:

> Our view is that, in reporting beliefs, we quite often are talking about such notions, although our belief reports do not explicitly mention them. The general solution to the puzzles is to allow a condition on particular beliefs, over and above a content condition, to be part of the claim made. The

---

[20] More about the relation this theory holds with indexicality, further on.
[21] An idea's semantic role is the function that determines the idea's content based on the agent's circumstances.

version of this strategy we shall pursue here is to take this further condition always to be a specification of the notions that are supposed to be involved in the ascribed belief. (Perry and Crimmins (1989), p. 697)

As a result, we have that the belief relation invoked by BRs is a three place relation involving an agent, a proposition, and a sequence of notions which are *unarticulated constituents* of the content of the report at a time. When we make a report we are implicitly referring to these unarticulated constituents, the notions; for this reason even if the proposition that the agent is claimed to believe is the same, substitution of co-referential names in a report can affect the way of grasping that proposition to which a specific utterance of the report is implicitly referring. As a consequence, intersubstitution of co-referential names can result in a change of truth value of a report for the reason that the expressions themselves, the proper names ('Hesperus' and 'Phosphorus'/ 'Cicero' and 'Tully'), are relevant for determining which way of grasping is being referred to.

It is important to stress that there is a suggestion that the third element in the belief relation might vary with the context and the wording in which the belief is presented. That someone be truthfully reported to hold a certain attitude towards a proposition depends also on the extra-linguistic context, which establishes which notion is implicitly intended by an utterance of the sentence.

As said by the authors, unarticulated constituency is very similar to indexicality, and is particularly related to it in that both phenomena rely on the context. They are different in that for unarticulated constituency the context alone does not suffice to identify the constituent, rather the whole utterance (the context plus the words uttered) is necessary. However, context plays the same role in indexicality and unarticulated constituency: that of supplying us with implicit contents of what we are talking about.

Many describe this as a "hidden-indexical theory" because it sustains a proposal close to indexicality, but there are no syntactical elements of the belief attributing sentences that explicitly refer to the notions. In sentences containing indexicals like "It is three o' clock here" we can think of the reference as being achieved by the context-sensitive expression "here", that is, when spelled out the sentence would be 'It is three o'clock at p', where p ranges over places. In unarticulated constituency there is no such expression

to determine the constituent in this way, instead, the notions appear tacit but the reference they help achieve is obvious in the context.

That there might be sentences that behave as if they contained indexicals, even if they do not have in their syntactic structure any element that holds responsible for the variation in the proposition expressed by utterances of that sentence, is not that strange an idea. Take the case of "it is raining" or "it's five o'clock", which according to some express a proposition in which the place of utterance is a constituent, much as if there was an explicit 'here' as in the sentence above.

In addition, the same way it happens with indexicals, when we need to clarify what is tacitly being said with a BR we can easily evoke the notions we mean to talk about and make them explicit. In the case of our example at the beginning of this subsection, we could for instance reveal for our audience that Lily believes Cicero is a great orator in one way, the way related to her oratory class, but not in the other that she associates with her political philosophy class. We can explain ourselves in whatever way we choose when required to give information about our reports, making clear what are the notions intended by an utterance of a sentence.

In this line, under this view it is foremost important for the proposition expressed by a belief ascribing sentence not only what Lily believes but *how* she believes it, that is, the notions that might be involved in her belief. It is also essential to note that the notions that are relevant are the attributee's, Lily's, not those of the reporter.

This applies as well to the treatment of the case introduced in Section 1, in which attributing to Lily that she believed that Tully was a great orator was correct in one context but didn't seem so in another. This can be explained away saying that the report in the two situations presented involves different and disconnected notions, even though they are not explicitly mentioned. The idea is that Lily has different notions of Cicero due to misrecognition. Perry and Crimmins would say that depending on the context we might be talking about different things when we utter the single report "Lily believes that Tully is a good orator", because different notions can be implicated. As notions are an implicit but fundamental part of propositions (and thus are indispensable for establishing semantic value of reports), and they vary across contexts, sensitivity of BRs should not be surprising. Moreover, to the authors it is perfectly coherent to report a person as believing p and as not believing p at the same time if the notions involved are

21

unlinked. A report denying a belief and other affirming it can both be true depending on the context. Problem solved.

To recap, what we report when we talk about an agent's beliefs at a certain time is a relation between and agent, a proposition and a third element that is a series of notions. Notions are cognitive particulars that are unarticulated constituents of ascriptions and vary with the wording and the context like indexicals. This means that belief reports often exhibit a hidden indexicality, i.e. their truth value can vary across contexts because they are constituted by elements that, although they don't have a syntactic place, are relevant for semantics. It is important to remark that the notions relevant for the variability in truth value of BR across contexts are those of the agent; that is, it is the way the agent believes -in relation with the familiarity the agent has with the name used in the report- what influences its truth value depending of restrictions in different contexts.

Now, there is the question as to which parameters of the context might be relevant for establishing the notions in play of an utterance. Unfortunately, there is no explicit answer to this question; as the authors recognize: "We have claimed that belief reports are context-sensitive, that they invoke unarticulated constituents, without offering any general method for determining what the relevant contextual factors are, and how they give rise to these unarticulated constituents of belief reports." (Perry and Crimmins 1989, p. 711).

However it is possible to grasp a suggestion on the matter from their treatment of several puzzles in the literature apart of the ones here presented. Principally, it seems that in their explanation of Kripke's famous puzzle[22], they suggest that the notions intended by a report need to be "germane", to match the current conversational context. In their words: "The speaker of the former report is claiming that Pierre [the agent] has a belief involving some notion germane to the current conversation about the stories, with the content that London is pretty." (Perry & Crimmins, 1989, p. 707). This could

---

[22] The Puzzle is the following: a Frenchman, Pierre, first hears of London referred to with French name 'Londres', and he comes to believe it is pretty. Then he moves to London, and, unable to connect it to the city he has heard about wit the French name, comes to believe it is not pretty. He does not change his mind about the city he has heard of, but simply does not make the link that the city he knows as 'Londres' is London. Given this case, with sentence

(7) Pierre believes that London is pretty

we seem to be able to say something true and to say something false when explaining different parts of the story.

suggest that interests of the audience and the speaker and the familiarity or unfamiliarity the hearers have with the words involved in the report might be contextual features that are of importance to determine truth value of BRs. Still, I highlight again that this is only a suggestion extracted from the treatment of the certain dilemmas, and it is not an explicit proposal of the authors.

This view has been criticized by other philosophers who consider a three part relation of belief is implausible because linguistically there is no place for a third element in propositional attitude reports.[23] But more importantly, the "claim that belief reporters make reference to ways of taking propositions is not plausible psychologically" (Bach, 1997, p.5), that is, according to some this view is asking too much from an agent who reports and judges truth value in an automatic and often unconscious way.

In spite of the criticisms made to the view, the proposal by Perry and Crimmins that context might play a greater role than generally thought in establishing truth conditions of BRs, also takes shape in other theories. Between them, I consider especially important Mark Richard's theory which I will present below and shares with Perry and Crimmins' not only the contextualist approach to belief reporting but also a representational conception of belief. As we will see afterwards, other contextualist theories will reject this characterization in favour of other approximations to belief.

1.3.2. Richard's representational contextualism

The central idea of Richard's view is that when a person uses a belief ascription sentence such as 'Nana believes that Phosphorus is a star visible in the morning', she is making a claim about what sentences Nana has in her *representational system*, or would accept. Particularly, he is claiming that she accepts a sentence that correctly corresponds to 'Phosphorus is a star visible in the morning'. In our example, we see that Nana accepts "Phosphorus is a star visible in the morning", but if this sentence cannot be correctly represented by 'Hesperus is a star visible in the morning', then the report of her belief is false.[24] Then, it would be important to understand what exactly counts as a correct representation or in virtue of what sentences correspond or fail to correspond. As

---

[23] See Stanley (2002) for a criticism of this sort.
[24] Note that this idea is extremely different from Naive Russellianism, which concludes that the believer's dispositions to accept the proposition that is being reported don't contribute to the truth value of the report.

I will explain in this section, Richard's proposal is that it is restrictions of the context that govern correct reporting.

Richard (1989; 1990) argues that appropriateness of reports varies from context to context of utterance. Consequently, truth value of utterances of reports like (1) or (2) may change across contexts without any variation whatsoever in Nana's belief state. According to this view we cannot inter-substitute co-referential names in belief reports preserving truth, just as our common intuitions predict. Richard explains this is due to contextual boundaries and properties that restrict (correct) representation. The suggestion is that among the contexts there are differences in what (for instance) 'Phosphorus is a star visible in the morning' may correctly represent. The case is analogous with that of comparative adjectives like 'tall', that when applied to a subject, have a value that can vary across contexts because what counts as tall fluctuates in different contexts depending on the standards of comparison. For instance, a sentence like "Michael is tall" is true if Michael is a 1.20 meters 5 year old and we are discussing if Michael have chances of winning a height competition held between kids his age, but it would be false if we are talking about the importance of height for the selection of players for the local basketball team. It depends on the context whether Michael can be described correctly as tall or not. What makes the difference is the standard of comparison to which 'tall' is subject in each context, that standard can be seen as a contextual restriction that contributes to semantics.

Before, I have illustrated a clear situation of how this happens in belief reporting with Lily's Cicero/Tully case (at the very beginning of this chapter), but it is easy to imagine other scenarios in which this occurs: in court, for instance, because the attributor has little confidence that the belief holder herself would assent to what he is declaring. It is clear here how the standards present in each context restrict correct representation and thus affect truth value of belief reports. According to Richard, the way to understand cases like these (but also other cases like the Hesperus/Phosphorus) is saying that in each context where one utters the report ascribing to her the belief that Tully is a good orator, there is a set of restrictions imposed by that context to the report; "these restrictions are typically the result of the shared intentions of those involved in the conversation" (Sider 1995, p. 4).

So under Richard's view restrictions imposed by the context are determined by intensions and interests of the speaker and the audience. I want to highlight here this interesting suggestion about these parameters of the context that intuitively might affect truth value of our attributions. Sadly, it is not further elaborated or discussed in depth. On the contrary, the piece of our attributions that is sensitive to context is central to this theory. According to Richard, a belief report is true, as uttered in a context, if the pairing of the sentence in the language and the proposition it expresses (determined by the subordinate clause) represents some related pairing in Lily's representational system under some correlation that obeys the restrictions in the context relevant to Lily, the believer.

Richard's account relies on detailed machinery for explaining how can he keep DR about names and still be faithful to our intuition that co-referential names don't seem to be intersubstitutable *salva veritate* in cases of propositional attitude reports. I will not give an in depth explanation of the details of Richard's theory, but in essence, the view is that the 'that S' clause of a BR determines an entity that he calls a RAM (abbreviation of 'Russellian Annotated Matrix'). A RAM is a pairing of a natural language sentence with the Russellian proposition that it expresses. The core of Richard's idea is that each agent has a Representational System consisting of RAMs of all the sentences she accepts as true, and for ascribing correctly it is necessary that the correlation between RAMs in the agent Representational System and RAMs determined by the 'that S clause' in the ascription be acceptable. This is something that is determined by the context of utterance of the BR. It is important to note that RAMs are psychological entities.

In Richard's theory, saying that belief reporting is contextual means that in different contexts where the restrictions imposed for the correctness of representation vary, truth conditions also diverge, while the agent whose belief we are reporting remains in the same state. We do, under this picture, report the agent's belief, so our ascription needs to capture something about the believer's representational system, but the appropriateness of our representation depends on how well it fits the restrictions of the context. Thus, context can turn a true report into a false one. Here, an element of the belief relation agent-proposition-representation is context-sensitive when we make a report of it; therefore, as each context presents its own restrictions, the relation may change in every context.

As we noted earlier, it is distinctive of Richard's contextualism that it comes along with a representational view of belief. However, some philosophers[25] reject the idea that propositional attitudes as constituted by representational states, are implied by or are necessary for explaining our linguistic behaviour and mental life. Also, in philosophy of mind[26], that beliefs are psychological representations has been questioned in favour of more dispositional accounts. According to authors who hold this position, the representational depiction of belief doesn't seem to be faithful to our current linguistic and psychological practices which they regard as probably far more complex than that in the sense that they involve more than merely representing; they endorse that belief is a compound of dispositions that include mental and emotional states as well as behaviour. Philosophers working on belief reports also started to think that the Fregean characterization of belief might not be as good a suggestion as it was agreed by the majority, and argue for a non-representational understanding of belief. But if it is not representational contents, then what are we making a claim about when we attribute a belief to an agent? I present next Howard Wettstein's account, which offers a radical answer and an alternative for the treatment of belief and BRs.

1.3.3. Wettstein's conjuring trick

Howard Wettstein (2004) provides us with a striking suggestion pointing towards a radically different, non-representational account of belief reporting, that aims to honour our intuitions on cases like those presented in Section 1 while rejecting classical approaches that understand belief reporting as an appraisal of someone else's psychological state or attitude towards a content. In his account, names are directly referential and context plays a central role in semantics of propositional attitude ascriptions. He also takes the data revealed by cases like Lily's as suggesting that belief reporting is something different from a formulation of a relation between a person and a propositional content.

In order to build his account, Wettstein takes inspiration on Quine's suggestions about treatment of reports of sayings and the division he makes between Direct (Lily said "p")

---

[25] See Denett, D. (1987), or Churchland, P. M. (1981) for views that support this approach to mental representations.

[26] See Schwitzgebel (2002) and Schwitzgebel, E. and McGeer, V. (2006) for a defence of this sort.

and Indirect Discourse ("Lily said that p")[27]. According to Quine when we quote directly there is no need for interpretations of any kind and therefore there is not much implication of the reporter whose report is true just in case she gets the sentence uttered just right. On the contrary, an indirect discourse report is a complicated practice that consists in putting a sentence in the agent's mouth. In indirect discourse the reporter is active and needs to choose a sentence that conveys the original speaker's point to the audience in a specific context. In view of that, what we do when we report in indirect discourse is very similar to translation, that is, we have two important tasks: (i) try to get the original discourse in touch with a new audience while (ii) maintaining the integrity of her discourse. As it is evident, there is a tension here because keeping faithful to the original discourse encourages the reporter to stick to the speaker's words as much as possible, which might not be a useful strategy when, for instance, speaking to an audience of a different culture, as this last task might require that one adjusts the original words employed. The example of translation is very clear on this matter: if we translate a writing word by word, the odds are high that the new audience will receive an incomprehensible text that really doesn't reflect the original author's ideas. For that reason indirect discourse, just as translation, is not a matter of capturing literal meaning, but a matter of paraphrasing.

The reporter granted such latitude, it becomes manifest that what counts as making a truthful report is a lot more difficult to define for indirect than for direct discourse. It is clear that with translation we don't even attempt to judge truth and falsity, we say of translated works that they are better or worse according to certain standards, but with indirect discourse we don't seem to regard the issue in such a lax way. An indirect discourse report counts as true if and only if the embedded sentence paraphrases the original remark satisfactorily; all substitutions are allowed, but with certain limits set by contextual considerations. Again, just as we see in Lily's example in the introduction, some contexts seem to impose restrictions to a report such that substitutions of co-referential names can turn a true report into a false one. In this line of thought, what counts as good paraphrase in one context can no longer be so in another, and there are contexts in which there is a very delicate line that infringes especial restrictions on reports. These are contexts where "it is all important how the speaker was thinking of

---

[27] I'm talking here about 'say' and 'believe' indistinctively, grouping them as phenomena that may require a contextual explanation.

the referent" (Wettstein, 2004, p.199). Suppose for example in our Cicero/Tully case in Section 1, that Daniel asks me whether Lily believes Tully is a great orator because he is interested in knowing if Lily knows that Cicero and Tully were the same orator: here it is clearer that my report can turn into a falsehood if I substituted the co-referential names.

 In the case of indirect speech, the basic semantic idea is that when we report someone else's discourse the 'said that' creates a context in which it is clear that the following lines are a paraphrase rather than an exact quotation. Therefore, the embedded sentence is used to express something, but not to assert. As Wettstein expresses: "I don't take the embedded sentence of an indirect discourse report as a device of reference, nor do I take an indirect discourse sentence to be relational. "Says that" rather creates the sort of context just described –display or quasidisplay- and signals that what follows is a contextually appropriate paraphrase" (Wettstein, 2004, p.202). With belief reports the case is similar, but there are certain particular difficulties that arise from the fact that indirect discourse is a more observable phenomenon that is anchored on outer speech, while belief reporting responds to a wide variety of considerations because it is concerned with an internal experience. When you report what someone said your ascription is bound to someone else's audible words, there is not linking required between them an unobservable (inaudible) behaviour. Instead, for reporting beliefs it is possible that no openly visible behaviour grounds the ascription. In fact, we often attribute belief on the basis of our observation of actions by an agent, without the agent having uttered any sentence explicitly about the matter.

Inspired by Quine's *reporting as theatre* Wettstein thinks of the reporter as someone who plays the role of the agent, taking on her view. The reporter theatrically acts the part of the agent speaking as her and pretending to be in her state of belief, then putting words in her mouth. It s not even necessary that there be an utterance from the agent on the concerned topic as it is for indirect discourse, instead, the reporter must take into account other various behavioural, emotional, dispositional and contextual features like agent's remarks on related matters, affective reactions, particularities of his culture, etc. Belief reporting seems to be a finer procedure than paraphrasing: a process the author calls the *conjuring trick*. That the speaker makes a conjuring trick when attributing belief is a metaphor to express that "Attributing belief often involves distillation, summing up, an "all things considered" judgment" (Wettstein, 2004, p. 207). The

reporter takes into account for his attribution past behaviour of the agent and assertions about the matter or related topics, together with dispositions of a wide range.

In belief reporting the reporter is engaging in a conjuring trick, a task that involves the collection of factors that, together, allow the formation of the report. From a series of conditions plus factoring the reporting context we obtain a sentence to embed. Even when the agent has pronounced herself about the topic, the reporter who wants to be faithful to her belief must take in more than a single utterance, he must make a verdict, draw a conclusion based on numerous facts, that results in the ascription. When reporting beliefs, then, we make a judgement that considers a wide range of factors instead of doing the paraphrasing that was sufficient in indirect discourse, but both phenomena are comparable in that the sentence embedded needs to fulfil the same requirements, namely, it must be faithful to the original content and must be appropriate to the current context. Therefore, the embedded sentence in the belief ascription -as in reporting of sayings- expresses without asserting, and when she utters it, the reporter is playing the agent's part, speaking as if he was him.[28]

Under this account, patterns of intuitions like those seen in the Cicero/Tully case are not a problem, as it is explained that adequacy of substitution depends on the context. Whether it is accurate to report Lily as believing that Tully is a good orator, would depend not only on whether she herself has said so or recognize the attribution, but on many other aspects of the context including the audience, past behaviour and the knowledge of the reporter. To see this clearer, let's take a look at a case presented by Wettstein:

> Imagine that Sam doesn't know that the famous actor John Wayne was none other than his boyhood friend, Marion Morrison. He assumes that Marion spent his life back in Winterset, Iowa, pumping gas or some such thing. Marion Morrison, he tells us, was certainly no famous actor; he even failed to make the cut in various high school productions. […] Consider "Sam believes that Marion Morrison is a great actor". If we were reflecting on Sam's remarks about famous actors and his subsequent response to queries about his high school classmates, we will judge this sentence false. Clearly

---

[28] As radical as it may seem, Wettstein's idea that BR do not report what an agent believes is also endorsed by other authors, such as Kent Bach (see Bach, 1997) who offers an account of BR where reports merely describe what an agent believes, not report it directly.

he would deny Morrison was a great actor. If we were speaking to Morrison's relatives in Iowa, trying to encourage pride in Morrison's achievement, it will be natural and correct to quote the famous Sam as having said, or as believing, that Marion made the grade. (Wettstein, 2004, P.164)

A second case, which is parallel to the one I presented to introduce this chapter, is the following:

Imagine that the famous orator is called "Cicero" in America but in England he is called "Tully." Sam, an American, says "Cicero was an orator"; let's assume he is unfamiliar with the name "Tully." To an English audience I can report Sam as having said that Tully was an orator [...] By contrast, consider this case: I report Sam's remark to other Americans who are interested in whether Sam knows that Cicero and Tully are one. In such a context "Sam said that Tully was an orator" would be false[29]. (Wettstein, 2004, p. 199)

Notice that in the first case it seems that what matters for the attribution to be correct are the interests of the audience and the reporter, whether they are interested in different bits of information regarding the belief of the agent; in the second case, it is knowledge of the audience that makes the report false in one context and false in other. Although not explicitly, apparently Wettstein suggests these might be parameters of the context that govern over truth value of our attributing sentences.

In Wettstein's view the agent can remain in the same state while her belief is being reported in different contexts truly or falsely, because we are actually not reporting something that is in her head or a relation of any kind but making an all things considered judgement, a conclusion that takes into account multiple features of the context of utterance such as the parameters mentioned before. Given that the reporter expresses without asserting, and that contextual features affect truth value of belief reports, names can be directly referential with no need to recur to modes of presentation, senses or any third element mediating meaning. Furthermore, Wettstein rejects Frege's relational conception of believing; in his account "the verb 'to believe' does not refer to

---

[29] It is important to remark that Wettstein thinks of reports of sayings and reports of belief as fairly analogous, as explained.

a relation between an agent and a content, rather it indicates that what follows is the agent's take on the question at hand" (Wettstein, 2004, p. 211).

Also, dispositionality seems appealing, and clearly Wettstein's account treats belief as something very similar to a cluster of dispositions[30], taking into account speech, behaviour, and other states. However, he will also refuse to accept a dispositional account of belief, as he finds the dispositional picture problematic in that it gives dispositions a too central role that they don't seem to have when we regard belief reporting practices of common users of language. In particular, he argues, we don't seem to talk about them when we report belief; we merely speak for the agent.

We are not reporting a relational, nor a dispositional phenomenon, so what if not these are we making a report of? What is belief according to this view? Well that is a question whose answer is pretty much left open. The hint that the author gives us to solve it is, in his own words, the following: "Here is my radical suggestion: Taking our cue from the reports, perhaps we have been looking in vain for some sort of state or process of believing. We refer to no such state or process when we report belief. We just speak for the person." (Wettstein, 2004, p. 214). We, then, in belief reporting do not make claims about belief, or refer to dispositions, or to relations, and the nature of belief might be quite far from what other philosophers have thought if belief reporting is a practice such that doesn't concern an agent's belief.

1.4. The motivation and aim of this research

As I have shown, there is a collection of literature that endorses the idea that belief reporting might be context-sensitive. This possibility affects equally Fregean and not Fregean approaches to belief reporting.

For the Fregean, cases such as the Cicero/Tully explained in the introduction where contextual features seem to be influencing truth value of reports, are problematic. A Fregean characterization of propositions and meaning appealing to senses serves well to

---

[30] To make the contrast, Erick Schwitzgebel's (1999; 2002) phenomenal-dispositional account of gradual belief defends that believing is nothing more than being disposed to do and experience certain kind of things. Based on the observation of children's development where intermediate states of belief are present, Schiwitzgebel (1999) proposes that states of this kind occur also later in life and depend on the dispositions that constitute the belief. Belief here can have in-between, gradable states ranging from states of full belief where the individual possesses the belief completely, to states of empty belief where she doesn't have it at all. He presents several cases where it seems difficult to say with certainty if in determined conditions an agent possesses or not the belief.

approach problems with substitutivity, but does not contemplate that the established binary relation between the agent and a proposition might change across contexts. For the Fregean, the binary belief relation holds eternally, it contains senses, but they do not solve the question as to why a single attribution might change truth value across contexts. Furthermore, Frege lacks an explanation that would account for why depending on the context (1) and (2) could in fact be interchangeable; for him this would amount to ascribing the wrong proposition to an agent, a serious mistake. If context-sensitivity was the case, it would represent a challenge for the Fregean.

But what contextualists show seems to be incompatible with Fregean semantics almost as much as with Naïve Russellianism. If contextuality of BRs turns out to be true, Naïve Russellians also suffer as they share with fregeanism one important feature in their semantics: that belief is a binary relation that holds *eternally* between the agent and a proposition (however this last component is typified). For Frege, as well as for Naïve Russellians as Soames and Salmon, it is a yes/no question whether an individual has a relation to a proposition. This relation is eternal and does not depend on where you are or features of the context of any kind. In this sense, the issue of contextuality is a more general one than the intrinsic problem of the alleged failure of substitutivity of co-referential terms, because it does not only restrict to offering a solution to the problem Frege elucidates for indirect contexts and in fact it can be regarded as entirely independent of that.

The problem pointed out in apparently context-shifting cases, is that they seem to show that if we want to accept that BRs are about a belief relation, this relation cannot be eternal, instead, we need to allow that it varies across contexts. The evidence that an attribution that is right in one setting, seems incorrect when we move to another situation, is the root of the idea of context sensitivity. Based on this, theorists endorse that parameters of the context might influence truth value of BRs, which presents an issue that goes beyond intersubstitutivity for the reason that even if we grant that either Fregean or Naïve Russellian depiction of propositions and meaning is right, and even if we admit that their treatment of the apparent failure of PS is correct, there is still the question as to whether a report can be sensitive to context. This is an independent question that neither Naïve Russellianism nor Frege can explain.

On the contrary, contextualist ways of analyzing the matter making truth of reports relative to a context look plausible and match well with our intuitions about apparently context-shifting cases. However, the burden of proof is on the contextualist to show that in actuality truth value of BRs can be affected by contextual features, that a single report that an agent *A* believes that *S* can correctly be reported in one context and not in other where the mental state of the agent of whom we are claiming that she believes that *S*, is kept stable.

On the basis of this suggestion, my purpose in this dissertation is to develop an experimental plan to test whether parameter-sensitive intuitions towards cases of belief attribution arise in non-philosophers, aiming to offer empirical evidence regarding the philosophical discussion around this issue. Experimentation will consist in designing adequate vignettes that portray context-shifting cases, and test them with non-philosophers under controlled conditions. I expect this to provide insightful data about intuition patterns as regards those cases.

I do not intend here to solve theoretical disputes by means of experiments, as I do not want to uphold the idea the philosophical theories in all their complexity can be proved or disproved on the sole basis of the data I here gather. Nonetheless I hope that testing intuitions about our belief reporting practices can give us valid inputs that could translate on theoretical elaboration and analysis. Details of the procedure and methodological considerations can be consulted in the Methodology (Chapter 4).

This follows what I consider is a common practice in contextualist theories: to hold the assumption that intuitions are a reliable source of evidence of our linguistic behaviour. In classical literature on different kinds of contextualism, authors make usage of the method of eliciting intuitions by giving their readers thought experiments that intend to make evident relevant aspects of our usage of language that support the theories proposed. In this line, I intend to determine whether there is empirical evidence that would constitute a reason to be inclined towards either a contextualist or a non-contextualist approach of BRs. In other words, I aim to test whether contextual features play a role in our attributions of belief by evaluating if a belief report can be right in one context and incorrect in other while the believer to whom the belief is attributed remains in the same mental state.

In the case of other propositional attitudes, examination of possible sensitivity to contextual features has revolved around the evaluation of specific parameters. For instance, as I will present later, in epistemological literature we find an important discussion that has been centred on whether low or high stakes can contribute to variation of truth value of knowledge attributions. Also, recent work in this field concerns whether the apparent variation is to be explained in terms of contextual features of the context of the utterer, or if, on the contrary, it should be regarded as a matter that involves the context of the agent.[31] No such analogous debate has been carried out by proponents of contextualist theories of belief reporting.

For designing quality experiments that collect the right data, it would be important to first be clear about which variables in the context might affect attributions of belief and whether it is the attributor's or the attributtee's context that affects truth of BRs. There are not many explicit proposals in the literature about exactly which parameters of the context might affect our judgment and attribution behaviour. This is a theoretical gap that even some of the proponents of these theories themselves have recognized (for instance, see Perry and Crimmins quote about the matter in previous pages).

However, as I have highlighted in the presentation of the approaches, the analysis of cases that contextualists offer, together with certain aspects of their theories, gives us hints on which would be candidates for a variable that affects truth value of belief ascriptions. As explained previously, Perry and Crimmins state that the agent's notions - related with the familiarity the agent has with the name used in the report- influence the truth value a BR has in each context, and there is a slight suggestion that interests of the audience and the speaker as well as the *referential knowledge* (by this I mean the knowledge -the familiarity or unfamiliarity- that a certain name denotes a specific individual) the hearers have with the words involved in the report, might be relevant parameters. On the other hand, Richard seems to propose that reports are constrained in different contexts depending on the intensions and interests of the speaker and the audience. Finally, for Wettstein, interests of the hearer and the speaker, the referential knowledge that the audience possesses, and several other undetermined parameters, are responsible for variation.

---

[31] Chapter 2 is devoted to this discussion.

This brief summary of patterns of variability extracted from the contextualists theories here presented can show that suggestions about this matter point in many different directions and still, no explicit proposal has been endorsed by any of these authors. There is no consensus or even discussion around precise elements of the context that might be responsible for the alleged truth-shifting. In addition, even though the discussion around stakes in epistemology has been very important, and there is a recognized strong connection between belief and knowledge, no author has made a concrete proposal either in favour or against the likelihood that stakes be important for context dependency of belief ascriptions.

The same happens with respect to the question as to whose context matters for variation of the attribution of belief. It looks as if in general, all three theories I here present consider that it is the context of the reporter that is important. Whether the conjuring trick suggested by Wettstein results in a truthful attribution, as well as the correctness of representation of a belief as Richard understands it, and the notions intended by the speaker with an utterance of a report as Perry and Crimmins portray them, seem to depend on the context of the speaker. However, the possibility that it'd be the situation agent is in which affect truth value has not been argued for or against.

It seems to me that there is an open field for elaboration on our understanding of BR concerning the discussion about what in the context might be responsible for the alleged variability of truth value of our attributions and whether it is parameters in the context of utterance or in the context of the agent.

On the other hand, none of the variables previously mentioned, have been put to empirical test to observe if intuitions about correctness of BRs covary with them in the actual practice of users of the language. I certainly believe experimental research gathering data about the intuitions on cases testing specific parameters, and checking whether contextualist patterns of attribution are actually present in folks, can enrich the discussion about belief reporting. For this reason, in this dissertation I want to focus specifically on exploring empirically two of the previously mentioned contextual parameters that might influence truth values of BRs. I propose that what is at stake in the context and the referential knowledge -i.e. the familiarity/unfamiliarity- of the audience are variables worth testing for the hypothesis that people have parameter-

sensitive intuitions on belief attribution cases where the state of the attributee is kept fixed. [32]

Referential Knowledge of the Audience (from here on in this dissertation RKotA) seems like an interesting parameter to test not only because Wettstein – and probably also Perry and Crimmins- suggest it is relevant for determining truth of belief reports, but also because, as introduced previously, other theories in the vicinity of propositional attitudes seem to regard reports as depending on it. That's the case of Perry's theory of semantics of names and also Cappelen and LePore's account of 'said that', which regard acquaintance of the audience with a name as a parameter of the context that affects truth value of reports.

Likewise, the choice of stakes as a parameter to test for BRs was basically based on the fact that it has been fundamental in the debate around attributions of knowledge. Since knowledge is traditionally characterized as justified and true belief, I think that findings concerning sensitivity of BRs may have an impact on our approach to knowledge attributions. As we will see in Chapter 2, theories that in some way endorse sensitivity of attributions of knowledge have been centred on whether parameters affect the justification. The discussion in epistemology has revolved around whether one's true belief that p is justified enough to count as knowledge, depends partly on contextual features such as stakes so that in some contexts the agent is justified and in other contexts he is not justified. None of these theories address the possibility that attributions of belief be stake-sensitive. My purpose here is to study variability of BRs in high and low stakes situations, with the hope that results on this matter can reflect on the discussion around epistemological sensitivity. Specific details about the idea that knowledge attributions might be stake sensitive are going to be clearer after the presentation of issues in epistemology in next chapter.

---

[32] As I explain here, the choice of variables to tests is based on the lines of various proposals in belief reporting theories and other approaches to propositional attitude reports. However, it can be a letdown that I cannot find a theory to oppose to, or that would endorse the same position I'm holding. There are a few accounts that seem to rely on stake-invariantism of belief (not belief reporting), such as Ramsey's and de Finetti's, (see Armendt, 2008) who sustain a traditional account of gradable belief that do not allow for internal variation depending on stakes. But these are not theories on belief reporting and probably not relevant to the purposes of this research. Therefore, I am left with the unattractive outcome that there is not visible contender for my proposal. Nonetheless, I count this as a sign of how is novel the issue I'm treating here, and how useful can be a little light on this issue for the development of philosophy of propositional attitude attributions.

Chapter Two

Epistemic Parameter-Sensitivity

For the purposes of this chapter, let us define *Epistemic Parameter-sensitivity* (henceforth from here on, EPS) as the idea that the truth value of sentences of the form '*S* knows that *p*' might be affected by contextual parameters, where by 'contextual parameters' I mean either features of the context of utterance in which someone attributes knowledge to someone else, or factors in the situation the agent or attributee finds herself in. This idea came about as an attempt to offer a solution to several epistemic worries and as an effort to account for our common linguistic usage of knowledge attributions. For this reason, it has been encouraged on the basis of every day cases that aim to show some aspects of our attributing practices which seem relevant for the semantics of knowledge attribution. These cases are everyday situations conveyed as short stories that apparently show variability of truth value of a single knowledge attribution depending on the context of utterance[33]. For instance, the following pair of cases has been typically taken as motivation for EPS:

> *Bank A*: My wife and I are driving home on Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines inside are very long, as they often are on Friday afternoons. Although we generally like to deposit our paychecks as soon as possible, it is not especially important in this case that they be deposited right away, so I suggest that we drive straight home and deposit our paychecks on Saturday morning. My wife says, "Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays." I reply, "No, I know it'll be open. I was just here two weeks ago on Saturday. It's open until noon."

> *Bank B*: My wife and I drive past the bank on Friday afternoon, as in Case A, and notice the long lines. I again suggest that we deposit our paychecks

---

[33] Parallel to the situation with BR, these are supposed to be *context-shifting* cases.

on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case, we have just written a very large and very important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a very bad situation. And, of course, the bank is not open on Sunday. My wife reminds me of these facts. She then says, "Banks do change their hours. Do you know the bank will be open tomorrow?" Remaining as confident as I was before that the bank will be open then, still I reply, "Well, no, I'd better go in and make sure". (DeRose, 1992, p. 913)

In both cases we need to suppose that the bank does indeed open on Saturdays.

The contribution of these cases as a motivation for EPS lies in our intuitions with respect to the truth value of the knowledge attributions in each scenario where only contextual features have varied. Most people agree, and EPS theories defend, that in the first case the attribution is true but not in the second, showing that apparently the change in a contextual parameter can affect the truth value of knowledge ascriptions. What varies from one scenario to the other is one specific parameter: what is at stake in each situation, that is, the costs it might have to be wrong about the matter in question. The first situation is a low stakes one, one in which it is not particularly important that the attribution be right. On the contrary, the second story portrays a high stakes situation, a very risky one where the costs of being wrong are elevated.

EPS theorists stress what they understand as a correspondence between their portrayal of knowledge attributions and our ordinary usage of language, a connexion that makes their proposal a grounded answer to scepticism which reveals how people outside philosophy make use of knowledge sentences. They regard intuitions elicited by cases like the ones above as important evidence that has to be accounted for and that appears to support parameter-sensitivity. An abundant quantity of cases has been described in the literature[34], which serve the theorists' aim of providing evidence for an account that explains our natural behaviour in conversation when we make knowledge attributions.

---

[34] Under the interpretation of some (see Shaffer, 2006), in cases like these here presented it is not only stakes that vary, but also other variables such as salience of error seem to be relevant. For this reason

Throughout the discussion about the best way of explaining our intuitions regarding ordinary cases, the basic idea of epistemic parameter-sensitivity has developed into two contrary views in epistemology: *Contextualism* and *Interest-relative Invariantism.* I will next describe the two approaches briefly in broad terms[35].

1.) *Contextualism*, proposed by Keith DeRose and Stewart Cohen among others, is "the position that the truth-conditions of knowledge ascribing and knowledge-denying sentences (sentences of the form "*S* knows that *p*" and "*S* doesn't know that *p*" and related variants of such sentences) vary in certain ways according to the context in which they are uttered" (DeRose, 1999, p. 1). According to these authors, an adequate way of regarding knowledge attributions is to "suppose that attributions (or denials) of knowledge are indexical or context-sensitive. The standards that apply are determined by the context of attribution" (Cohen, 1986, p. 578). In this view, the proposition expressed by an utterance of a knowledge ascribing sentence can vary depending on certain standards about what counts as 'knowing' imposed by the context of use. An ascription that is true uttered in a context with lower standards can turn to be false if constrained by a higher standards context.

Different lines within this approach have considered different contextual features to be relevant to the semantics of knowledge attributions, but independently of the account, it is general to all forms of Contextualism that these aspects of the context are not to do with the attributee (i.e. matters about the evidence she has or her physical or mental state), but rather they are factors related to *the attributor's* psychological state or practical situation. As a result of such dependence on the context, a single knowledge attribution expresses different propositions when uttered in different contexts, and thus may differ in truth value according to the situation.

2.) On the opposite side, we have *Interest-Relevant Invariantism* (henceforth IRI, also called, Subject-Sensitive Invariantism), put forward by Jason Stanley (2002; 2005; 2007; 2008) and John Hawthorne (2004; 2008) among others. *Invariantism* is the thesis that what is expressed by an utterance of a knowledge ascribing sentence does not vary with the context of use. *Interest-relativism* is the thesis that whether a person knows a proposition *p* at a time, partly depends on her practical interests at that time, so that the

---

many thought experiments have followed these original cases, maintaining the same structure but making several modifications in order to support certain theoretical preferences.

[35] In later sections I will give more details about each view.

higher the costs of being wrong about *p*, the harder it is to attain knowledge. IRI, then, is a view that combines these two theses.

The different authors that sustain IRI do it in very different ways, but it is common to all IRI accounts that they reject Contextualism, as they deny that factors in the context of utterance (such as practical features in the attributor's context) are relevant for establishing the content of the proposition expressed by knowledge ascription sentences. Instead, these are views that endorse subject-sensitivity, that is, they hold that the truth of knowledge attributions might partly depend on the context *the agent* or *attributee* is in. In the words of Hawthorne (2004):

> Suppose instead that the kinds of factors that the contextualist adverts to as making for ascriber dependence – attention, interests, stakes and so on – had bearing on the truth of knowledge claims only insofar as they were the attention, interests, stakes and so on of the subject. Then the relevance of attention, interests, and stakes to the truth of knowledge ascriptions would not, in itself, force the thesis of semantic context dependence. Here is the picture. Restricting ourselves to extensional matters, the verb 'know' picks out the same ordered triples of subject, time, and proposition in the mouths of any ascriber. However, whether a particular subject/time/proposition triple is included in the extension of 'know' depends not merely upon the kinds of factors traditionally adverted to in accounts of knowledge…but also upon the kinds of factors that in the contextualist's hands make for ascriber dependence. These factors will thus include (some or all of) the attention, interests, and stakes of that subject at that time. (Hawthorne 2004, p. 157-58)

According to IRI, knowledge is in part dependent on the practical interests of the agent and each context imposes different standards that have to be met for a subject to know. Consequently, whether we attribute knowledge rightly will also depend on those practical matters; on this view it is understood that we are disposed to be more lax on our attributions when stakes for the attributee are lower, and vice versa.

To be fair, it is important that I mention that some theorists that prefer the second view would probably refuse to regard theirs as EPS accounts, for the reason that their focus is on the agent's circumstances and that they believe that what is expressed by utterances of knowledge attribution sentences are complete propositions that do not depend on

other factors alien to the subject of the report or the evidence she possesses. As I stated before, all forms of invariantism are opposed to contextualist approaches in this sense. However, as Crispin Wright (2007) notes, this view involves "a modest element of contextuality of truth-conditions" (Wright, 2007, p. 19) and is generally considered to include parameter-sensitivity in its explanatory framework. For this reason I here present them together.

Generally for both views, it can be defended that our use of knowledge attributions seems to be affected by what is at stake in a context, but the relevant context we are talking about in each of the two approaches is different. More precisely, on the former view, the truth value of an utterance of 'S knows that p' may vary with standards determined by the context the attributor is in. Following this approach, the proposition expressed by the utterance does not have a truth value previous to the context of utterance. It is understood, however, that whether an agent's evidence suffices to constitute knowledge is dependent on her practical interests at a time, which determine what are the standards that govern correct ascription.

A way of summarizing in simple terms the difference between these two ways of regarding the presumed context-sensitivity of knowledge ascriptions is to say that the first kind is "attributor sensitivity", while the second kind would fit more under the label "attributee/subject/agent context-sensitivity", where what counts as knowing depends on depends on features of the situation the agent is in, such as her practical interests. As we will see more clearly later on, these particularities of each view amount to important differences between them as regards the treatment of apparently context-shifting cases[36].

Both Contextualism and IRI agree in that what is at stake is a key factor in explaining the apparent variation observed in the cases. Let me highlight here that, as I stated before, the prediction both EPS theories make about the Bank cases is the same, i.e. they both recognize a difference in truth value of the knowledge sentences in the low and high stakes situations: they conceive the ascription in the high stakes context is incorrect, while in the low stakes is correct, nonetheless, the explanation they give for

---

[36] One very central criterion that I will briefly introduce further on and that separates Contextualism from IRI is their position towards Intellectualism -the thesis that the epistemic relation one bears to a proposition does not depend on practical matters. This is a thesis that Anti-intellectualists oppose radically and that according to some (See Stanley, 2007 and Pinillos, 2011), Contextualism needs to embrace. This is a point around which there has been a lot of discussion and that is important to mention.

this result is quite different. Particularities of this explanation are saved for later, each view in turn, but it is important to notice here that the Bank cases are first person attribution situations in which the attributee and the attributor happen to be the same person. But third-person cases make more evident the intrinsic differences between these views.

It is not my purpose here to examine these differences in depth, or to favour any of these epistemological approaches. My motivation for discussing these theories in this dissertation comes down to two reasons: First, to provide a background for the discussion in experimental philosophy in Chapter 3, which has revolved around EPS and is central to the comprehension of the empirical proposal I make in this dissertation. And second, because given the close relation that exists between knowledge and belief, the results of this dissertation may have an impact on the discussion on knowledge attributions, something that I will explain at the end of this chapter.

Together with these views that consider the phenomena observed in the Bank cases as a semantic one, I would like to present a suggestion recently made by Jennifer Nagel according to which there is a cognitive explanation for our intuitions on the cases. Nagel believes that the apparent shift in intuitions is to do with particularities of our cognitive processing, and therefore they do not provide evidence for a pragmatic nor a semantic phenomenon but are simply confirming that we engage in certain reasoning processes which differ for high and low stakes situations. This seems like a very interesting alternative to the discussion in epistemology, but also it might be quite illuminating as regards cases of BR where stakes vary.

In the rest of this chapter I will explain how philosophers have come to regard the truth value of knowledge attributions as parameter-sensitive, how the idea came about and which are the cases used by the authors to encourage it. To this end, I will describe the idea of a social component of knowledge, which is essential to understanding the appeal of EPS. Then, I will discuss how the standards one must meet for one's belief to count as knowledge vary according to features of the context, and how alleged variability in knowledge attributions in certain cases can be explained according to authors that support EPS. I will present attributor context-sensitivity relying on Cohen's view and DeRose's brand of Contextualism. Later, I will also Introduce Hawthorne's and Stanley's IRI in order to make clearer the contrast between attributor's and attributee's

context-sensitivity. Finally, I will be introducing the recent idea by Jennifer Nagel, who intends to explain our intuitions about apparently context-shifting cases in terms of cognitive processes, proposing that no semantic or pragmatic phenomenon underlies them.

2.1. *The social aspect of knowledge*

The first piece to come to understand the argument for the parameter-sensitivity of knowledge attributions is presented by Stewart Cohen, who offers an alternative to scepticism claiming that knowledge has a social component that determines whether the evidence one possesses undermines one's knowledge. The sceptic argument, roughly, recreates possible scenarios where the experiences we consider real are not, such as being a brain in a vat that is being stimulated by scientists to recreate sensorial experience, or being dreaming, or believing that what one sees in the zoo is a zebra and not a cleverly painted mule, or believing that the table is red because it looks red but in reality it is a white table under red lighting, etc. If p is a proposition knowable through the senses in an immediate way such as "I have hands" or "there is a table in front of me", according to the sceptic, the fact that one can never rule out the possibility that what one perceives is not as one takes it to be makes knowledge of p impossible.

Cohen (1986) explains the social nature of knowledge attributions based on the notion of justification or, as he calls it, *having good reasons for believing* a proposition. According to Cohen, epistemic reasons have a prima facie structure, they are defeasible, and for that reason where there are reasons to believe that p there is also a *defeater* that makes those reasons insufficient to constitute knowledge. For example, one can have what one thinks are good reasons for believing that a table is red (namely that it looks so) but there is a defeater (that in reality it is a white table under a red light) that turns those reasons into insufficient ones. If this is so, one can only have good reasons simpliciter for believing p iff one has prima facie reasons for which one possesses no defeater. This makes evident an ambiguity in the concept of having good reasons which implies that if a subject S knows r (a prima facie good reason for believing that p) and also knows d (a defeater), S will either stop believing that p or refrain from believing it.

This is how ideally knowledge should happen, but Cohen observes that subjects fail to accomplish this basically because the relevance of a defeater can range from obvious to very subtle and therefore one might fail to appreciate the effect of a defeater and believe

43

that *p* inadequately. Still, one may think that the subject had good reasons for believing that *p* because he was not aware of any obvious defeaters, in Cohen's words: "S can have good reasons even if those reasons are defeated (by evidence he possesses) provided it is still (epistemically) permissible for *S* to believe for those reasons, i.e., provided the relevance of *d* is not obvious." (Cohen, 1986, p. 575) So, the suggestion is that knowledge might depend on the *obviousness* of a defeater, and whether one attributes knowledge truly to someone, is relative to how obvious the defeater is. But obviousness is a psychological notion, so one needs to clarify: obvious for whom?

The answer Cohen gives us is that when we say something is obvious simpliciter we presuppose a level of reasoning ability that is *intersubjectively determined*, and in addition we take good reasons to be dependent on the reasoning level of the subject of attribution. In this sense, knowledge attributions are dependent on the social context; they are sensitive to parameters that exceed the agent. To explain this in depth, Cohen characterizes four kinds of defeaters which describe different interactions between contexts and how they affect knowledge attributions. The categorization depends, both on obviousness of the defeaters –therefore they can be evident or opaque-, and on the context in which those defeaters operate –thus, they are either intersubjective or subjective-. As we'll see below with the explanation of the kinds of defeaters, Cohen is introducing sensitivity of knowledge ascriptions to parameters aiming to finally argue for the claim that the context that matters for attributions is that of the speaker.

Consider a case where Jim believes the true proposition that the table is red (*p*) on the base of knowing that the table looks red. Jim also comes to know by a reliable source that there are red lights shining on the table, but unknown to him, the source is mistaken. In spite of the defeating effect of this information, Jim keeps on believing that the table is red because of some strong bias he has in favour of this proposition. If we take it that Jim does not possess any further evidence, we must say that he fails to know that the table is red. This is an example of a subjectively opaque defeater. According to Cohen, this suggests that if Jim possesses a subjectively evident defeater of his knowledge that the table looks red (as a reason to believe the table is red), then he fails to know *p* on the basis of this knowledge. Therefore, knowledge entails *subjectively* good reasons (namely, knowledge is dependent on the agent alone thus far).

Now suppose Jim fails to recognize the relevance of the fact that there are red lights shining on the table. He does not care about it and he thinks this does not make any difference to his perception of the table, and therefore continues to believe the table is red on the basis of his knowledge that it looks red. This defeater would be *intersubjectively* obvious but subjectively opaque. Certainly, Jim fails again to know, as the defeater is subjectively opaque to him, he cannot discern the relevance of an obvious defeating evidence.

Each kind of context may affect differently whether an individual has good reasons, and given that having knowledge entails having good reasons we need to ask ourselves which sense of 'good reasons' preserve that entailment, which defeaters undermine our knowledge. According to Cohen, a subject does not know *p* if the defeater is subjectively opaque, but instead, one would be able to grant knowledge to a subject when he knows a defeater that is intersubjectively opaque because its relevance is too difficult to appreciate. This makes evident that knowledge entails intersubjectively good reasons and thus, that knowledge has a social component. Attribution of knowledge is affected by features that are beyond the agent himself and pertain to the context. If the subject possesses a defeater whose relevance is obvious by social standards, and the subject fails to appreciate it, then the subject fails to know.

An example that shows clearly the underlying reasoning is the following:

> "Suppose *S* possesses a defeater of *r* (the table looks red) as a reason to believe *q* (the table is red), the relevance of which escapes the closest scrutiny of extremely acute subjects. We can imagine that only a super genius would see it[37]. Does *S* know that *q* on the basis of r? When considering this case, we must remember that *S* believes the table is red on the basis of it looking red, and the table looks red because it is red. It seems that, given this, *S* does not fail to know the table is red simply because he possesses a misleading defeater, the relevance of which is discernible only to an inscrutable line of reasoning." (Cohen, 1986, p. 578).

What seems to be happening is that in this case there is no violation of the intersubjective standards, and that definitely these are which contribute to make a

---

[37] Imagine, for instance someone who has a hyper- developed visual cortex, and can discern colours and shades people normally would not. It would then be defeaters that are only obvious to him.

knowledge attribution true. Now it will be important to determine how exactly this intersubjectivity is characterized, namely, which group is going to count for establishing the standard that can make a knowledge attribution true.

One might be tempted to say that the standards that affect knowledge attribution are established by the social group to which the agent belongs, but according to Cohen this option will give us an inappropriate account of knowledge because if S were among a group to which a certain parameter is particularly obvious, they would judge S as failing to know even if he belongs to a society that is completely blinded to features to do with that parameter, and wouldn't consider him at fault. According to Cohen, the parameters that matter to knowledge correspond, instead, to the context of utterance. In line with this, attributions of knowledge seem to be context-sensitive in that their truth changes according to the intersubjective standards imposed by the context of attribution[38]. In other words, "the truth value of a knowledge attribution will depend on the status of the defeater the subject possesses, relative to the standards that apply in the context of attribution" (Cohen, 1986, p. 579).

In our reports, it is sometimes made explicit by the speaker which standards apply, and it is possible that the he intends standards that are in accord with some social group to which he belongs, but he may arbitrarily decide to use whatever standards. He might even use standards that are not necessarily intersubjective, however -Cohen defends-, his attribution will be true only depending on the standards of evidentness he intends and the standards of evidentness of the defeaters possessed by the subject of attribution.

So far I have presented the argument for the social contribution to knowledge attributions, which already gives us a hint of what might motivate contextualist proposals. In the following section, I will discuss further the idea that the semantics of knowledge ascriptions depends on what is at stake in the context of utterance of the attribution.

2.2. Contextualism of knowledge attributions.

It seems plausible that knowledge attributions are affected by social components, but how exactly is that specific parameters come into the picture of EPS is something that

---

[38] This discussion about defeaters allows Cohen to arrive to the conclusion that the relevant context is that of the speaker. I want to highlight that, as briefly explained in the previous chapter, no parallel discussion has taken place in belief reporting.

requires more attention. Other contextualists apart from Cohen have followed this line of thought and have provided evidence that knowledge attributions are context-sensitive. They have also accounted for which aspects of the context are relevant for determining truth value of knowledge ascriptions. This is central to the debate around EPS in epistemology and experimental philosophy, and also an important element for my research project here.

As stated earlier, EPS is the general idea that knowledge attributing sentences can vary in truth value relative to parameters in different contexts. Within this framework, Contextualism is the position that the proposition expressed by an utterance of a knowledge attribution sentence, vary in certain ways according to the context of use. Therefore, truth value of this kind of ascriptions is context dependent.

To argue in favour of such approach, DeRose (1992) presents the famous Bank cases which I introduced at the beginning of this chapter. With them, he aimed to illustrate his view, eliciting contextualist intuitions that afterwards he uses to defend his view against invariantism and scepticism. These cases are originally designed as a contextualist attempt to establish the point that '*S* knows that *p*' and '*S* does not know that *p*' can both be true simultaneously relative to the context of utterance, while nothing relevant has changed in the epistemic status of the attributee. More precisely, the truth value of utterances of those sentences depends on what is at stake in the context of use. For this reason the cases involve scenarios in which one's intuitions about whether o not to ascribe knowledge seem to vary according to contextual changes. The intuitions elicited by these cases suggest that if a person is in two situations that are exactly similar except for some contextual features independent of the attributee, and it is possible to say in one that he has knowledge but not the other, then one cannot maintain that the difference between knowledge and true belief relies exclusively on epistemic features of the agent.

DeRose argues that it is natural to say that when he declares that he knows the bank will be open in *Bank A* -low stakes- he is saying something true, but it also seems acceptable that when he replies that he does not know the bank will be open in *Bank B* –high stakes-, he is also declaring something true. At the same time, given that there is essentially no difference in the position he is in for knowing in each case (the beliefs involved and the evidence he possesses are the same), it would follow that the sentence

'I know that the bank will be open on Saturday' ought to vary in semantic content in these two contexts of use.

It is clear that DeRose wants to defend that there is no contradiction between declaring one knows in *Bank A* and denying it in *Bank B*, that both of these claims are true even though the evidence and generally the same truth-related components concerning the relevant proposition is the same. However, it looks like we have an incompatibility between the knowledge utterances. If conversations like those pictured in the cases seem to make sense and we want to preserve the possibility of actually knowing something, then we need a good explanation that makes sense of the apparent shift in knowledge attributions, one that allows the knowledge denial in *Bank B* to be true without affecting the truth of the claim of knowledge in *Bank A*.

More precisely, DeRose states there are at least three contextual differences between *Bank A* and *Bank B*: 1) the importance of being right or what is at stake in the context of utterance. The higher the stakes, the higher the requirements a knowledge attribution must meet for being true. 2) The mentioning of a possibility. In order to say one knows something when facing a possibility one needs to rule out that possibility. 3.) Actually considering a possibility (i.e. that someone mentions it, might make one consider the possibility and have it into account for ruling it out before one says truthfully that one knows something). All three factors appear in *Bank B* but not in *Bank A*.

The position that factors 1-3 don't affect the truth conditions of knowledge attributions, as it was introduced at the beginning of this chapter, is called Invariantism. Instead, contextualists hold that contextual standards vary according to characteristics 1-3 or similar. Different types of Contextualism diverge in what features of the context affect the truth value and to what extent they do so. Most of them, nonetheless, favour more objective factors [39] such as the importance of being right for the attributor.

Another case that illustrates how, depending on what is at stake, the truth value of knowledge attributions might change is offered by Cohen (1999):

> *Airport Case*: Mary and John are at the L.A. airport contemplating taking a
> certain flight to New York. They want to know whether the flight has a

---

[39] Objective in the sense that they are not internal of the attributer, however, other less popular approaches endorse more subjective features, as the possibilities considered by the speaker in a determined context of use.

layover in Chicago. They overhear someone ask a passenger Smith if he knows whether the flight stops in Chicago. Smith looks at the flight itinerary he got from the travel agent and respond, 'Yes I know — it does stop in Chicago.' It turns out that Mary and John have a very important business contact they have to make at the Chicago airport. Mary says, 'How reliable is that itinerary? It could contain a misprint. They could have changed the schedule at the last minute.' Mary and John agree that Smith doesn't really know that the plane will stop in Chicago. They decide to check with the airline agent. (Cohen, 1999, p. 58)

In this case it is evident how what is at stake can apparently change truth value of knowledge attributions in the same situation, where nothing else has changed. Mary and John, who are in a high stakes situation, deny knowledge allegedly truthfully to Smith, while he himself claims to know in his low stakes position. Apparently there is an incompatibility between the attribution that Smith knows that the plane stops in Chicago and the report that Smith does not know that the plain stops in Chicago, but our intuitions tell us that they are equally correct and true. The explanation that contextualists offer for why this shift occurs is that practical importance (the need of getting it right) raises the stakes, and when the stakes are raised the standard one should meet for one's belief to count as knowledge is higher. Therefore, both the attributions by John and Mary and by Smith can be true, as they have been uttered in different contexts (high and low stakes ones). In each one, the knowledge attribution acquires a different truth value, and for that reason the utterances are not inconsistent.

To DeRose, expecting that both claims be evaluated by the same standards does not conform to the reality of our linguistic practices. According to him, an evaluation of the report that does not take into account the context of the attributor can lead to dialogues as weird as the following: "While standing in a bright yellow room, I said, 'This room is yellow.' The lawyer then dragged me by the ear into a room in which all was grey and got me to say 'This room is grey', and now he is jumping all over me: 'First he says, 'this room is yellow,' then he says 'this room is grey.' Which is it?'" (DeRose, 1992, p. 926). In this fragment, the fact that two assessments made in different contexts are judged under a single standard is made in to a parody to make obvious the importance of a change of contexts for assessment.

Still, even if we grant that contextual features are relevant to determining truth value of knowledge attributions, the contextualist who wants to explain the truth of utterances appealing to a change of context will need to elucidate what aspects of our knowledge sentences make their truth value variable through contexts. Many contextualists hold that the key answer here lays in the way the semantics of the verb 'to know' is defined. According to them, the verb 'to know' is semantically similar to indexical expressions. To most of them, a sentence expresses different propositions in different contexts, according to this way of interpreting the matter, when I utter "it is raining here", this expresses a different proposition if I utter it in New York (namely that *it is raining in New York*) than if I utter it in London (that *it is raining in London*).

Summarizing, the proposal is that 'knows' is comparable with indexicals in that it stands for different things depending on the context. Therefore, knowledge ascribing sentences that contain it, express different propositions in different contexts, with the result that one can attribute or deny knowledge to a subject whose epistemic state has remained stable, without being contradictory. It is stated that truth value of knowledge attributions is true relative to certain standards which are determined at least partially by what is at stake in the context of utterance. This is a view which is criticised and opposed by IRI, as we will see in next section.

2.3. Interest-Relevant Invariantism

Anti-Intellectualism about knowledge is the position that knowledge is not exclusively an intellectual concept but also, at least partially, a practical one. [40] What this means is that whether someone who believes *p* counts as well as knowing *p* can, in certain cases, depend on practical features of the agent's situation. In this sense, knowledge is understood as being sensitive to practical interests, such as the stakes the attributee is in. Together with *Invariantism* and *Interest-relativism (*as explained in the introduction), Anti-Intellectualism completes the picture of the fundamental theses that constitute Stanley's *Interest-Relevant Invariantism (IRI)*.

This account denies the semantic thesis that 'knows' works as an indexical, and argues that the intuitions in the stakes-shifting cases like the Bank scenarios can be adequately explained without resorting to Contextualism. According to IRI "whether a subject

---

[40] As it is self-evident, and was explained in a previous footnote, Intellectualism is the thesis that knowledge does not depend on practical facts.

knows -though not what is expressed by the relevant knowledge sentences themselves- depends upon facts about his/her practical interests" (Rysiew, 2011, http://plato.stanford.edu/entries/contextualism-epistemology/). [41]

IRI endorses that the standards for one's belief to constitute knowledge vary according to the *evidence the attributee possesses* and the *practical interests* she has in the situation. Regarding the Bank cases, for instance, IRI grants that there is a difference in truth value of 'I know that the bank will be open on Saturday' in both scenarios, but it explains that this is so not because it does not have the same content across the cases. On the contrary, this happens because there are practical facts which are relevant for establishing truth that differ between on scenario and the other, namely those facts that bear on the costs of a person's course of action such as the importance of getting the matter right.

The following is a case presented by IRI to illustrate the view that knowledge depends on practical matters of the subject, against Contextualist depiction:

> *Low Attributor-High Subject Stakes.* Hanna and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Two weeks earlier, on a Saturday, Hannah went to the bank, where Jill saw her. Sarah points out to Hannah that banks do change their hours. Hannah utters, 'That's a good point. I guess I don't really know that the bank will be open on Saturday.' Coincidentally, Jill is thinking on going to the bank on Saturday, just for fun, to see if she meets Hannah there. Nothing is at stake for Jill, and she knows nothing of Hanna's situation. Wondering whether Hannah will be there, Jill utters to a friend, 'Well, Hannah was at the bank two weeks ago on Saturday. So she knows the bank will be open on Saturday.' (Stanley, 2005, p.5)

And we suppose that the bank will be open on Saturday.

---

[41] It is important to state that for most Anti-intellectualist accounts, the claim that practical interests matter for knowledge is not eternal, that is, it doesn't entail "that for every difference in practical interests, there will be a corresponding difference in knowledge facts. Rather, it just claims that *there are cases* in which differences in practical interests facts correspond to differences in knowledge facts" (Pinillos, 2011, pg. 5-6. Emphasis added.)

Here we see a case in which the differences between IRI and Contextualism are salient. The mainstream intuition regarding the case seems to be that Jill's utterance of 'She knows the bank will be open on Saturday.' is false. In this case, Jill is ignorant that Hannah is in a high stakes situation, but she herself is in a low stakes one and in virtue of this the contextualist seems to predict that Jill's assertion is true, so allegedly here is a problem to solve for him. On the contrary, Stanley states the intuition that Jill's assertion is false is better explained by his subject-sensitive account.

IRI's idea, in more depth, is that knowing is a matter of having good evidence (having *serious epistemic possibilities* for our knowledge) under each particular situational standards, where having good evidence is understood as being a partially practical notion influenced by the agent's circumstances. In Stanley's words, "whether or not a proposition is a serious epistemic possibility [...] is determined by facts about the subject of the knowledge attribution and the time of knowing" (Stanley, 2005, p. 90).

For Stanley, knowledge ascriptions are about whether an agent bears epistemic relation to a proposition that is a *serious practical question*. A proposition is a serious practical question for an agent if there are alternatives to that proposition that the individual must rationally consider for decision making. An example in the author's words:

> Human beings are not ideal rational agents; we have limited capacities. Given this fact, there are certain questions we are rationally required to resolve before acting, and other questions we are not rationally required to resolve. We would view someone as neurotic who checked to see if there was a banana in her tailpipe whenever she went to start her car. But we would view her actions as perfectly reasonable if there were neighborhood boys who regularly stuck bananas in car tailpipes in her neighborhood, and she was aware of this fact. In the former situation, the proposition that there is a banana in the tailpipe of her car is not a serious practical question for her at the time of her checking, whereas in the latter situation, it is a serious practical question for her. (Stanley, 2007, p. 19)

Stanley believes that his account of knowledge integrates the intuitive notion that what is of practical relevance has an impact in our epistemic states. He advocates for a

general thesis according to which knowledge is constitutively connected to action[42] and thus can be affected by what is at stake for the subject of attribution precisely because stakes are a crucial factor for decision making and action. Consequently, whether someone counts as knowing *p* depends also on practical features of the situation he is in, such as what is at stake for him, the agent. The standards relevant for saying truthfully that an agent knows, vary accordingly to what is at stake for the agent, so if the subject of the attribution meets those standards, then no matter the context of the attributor the ascription must be true if it states that the agent knows and false if states the contrary. Whether the attributor is in a high or low stakes situation has no relevance, under this approach, for semantics of knowledge attributions.

Henceforth, what is at stake affects whether one counts with enough evidence for one's true belief to constitute knowledge. The reason why practical affairs of the agent's circumstances might affect our ascription behaviour is because if the agent is in a risky position, as speakers we tend to be less lax or permissive about our attributions than when there is nothing at risk for her, because we feel that a high stakes context demands her to be in a very good epistemic position to know that *p*, argues Stanley (2007). If the stakes are high, the evidence an agent must possess for her true belief to be knowledge is higher than in cases where the stakes are low. There is a requirement of the evidence that goes in accord with what is practically at stake for an agent in certain time *t*. This is a plausible explanation of the apparent context-shifting phenomenon observed in the cases offered; however, Jennifer Nagel believes it might not be the correct one. For her, as I will present in section 2.5., both Contextualism and IRI are misinterpreting a cognitive fact as a semantic one.

2.4.   *Some remarks on Contextualism and Interest-Relevant Invariantism*

So far we have seen two accounts that in different ways endorse sensitivity to parameters: Contextualism and IRI. The difference between the two views lies in how that sensitivity is characterized and what context is relevant for the supposed variation of truth value of knowledge attributions. But there is yet a deeper difference between the theories. It is important that we note that Contextualism is a theory about knowledge attributions, more than a theory about knowledge (contextualists mostly sustain some

---

[42] Other forms of IRI, like Weatherson's (2011) would diverge from this connection between action and knowledge.

primary epistemological claims such as one only knows what is true, and others of the like, but not complete views on the concept of knowledge). Contrary to this, IRI is a theory that explains the data collected on the basis of an Anti-intellectualist theory of knowledge. This entails a very important difference between the views; as Stanley says:

> […] contextualists about knowledge ascriptions have discovered many of the examples that suggest that whether a true belief is knowledge depends not just about truth-conducive features of a situation, but on what is practically at stake. However, contextualists generally share the widely held assumption that knowledge is not a matter of practical interests. So they have used these examples, together with the assumption, to argue for the thesis that a predicate such as 'knows that penguins waddle' denotes different knowledge properties on different occasions of use. Each of the resulting semantic contents is a property, possession which does not depend on practical interests. (Stanley, 2005, p. 2-3).

Defenders of Anti-intellectualism understand that the cases serve their claim that practical interests affect whether a true belief can constitute knowledge. While Contextualists like DeRose and Cohen would deny that this is a correct characterization and even consider it damaging to the very concept of knowledge. Therefore, they consider the preservation of Intellectualism (the thesis that knowledge is not affected by practical facts) as a point in favour of their account. This is a very important and discussed matter which is deeply related to the evidence the cases seem to show and the explanation of it each account supports.

As I have said before, it is not my purpose here to favour any of these accounts. What is relevant for my purposes here is that at least for IRI, it is the *justification* of knowledge what seems that is put into question in the cases, that is, it looks as if what stakes influence is whether one is justified enough for one's true belief to count as knowledge. Given that contextualism is not a theory about knowledge, it would be risky to affirm the same about it, but it can certainly be said that the question as to whether belief or belief attribution has an import for the phenomenon accounted seems to have no room in EPS approaches.

If we take a close look at the cases mentioned, we see that in all of them truth of the belief is granted, and a fortiori and what is most relevant for my purposes here, the truth

value of the belief is not supposed to change with the context. In the Bank cases, we see that DeRose points out that knowledge attributions in both scenarios are made "remaining as confident" that the bank will be open, that is, with no doubt that the belief possessed is true. The story is similar with the Airport and Low attributor- High subject stakes cases. Possibly, this is due to an intended design based on theoretical stands as regards justification, namely, the aim of the authors was to purposely disregard the possibility of other elements in the knowledge relation to be affected by contextual features as they wanted to focus on the justification. It is assumed that high stakes or low stakes do not affect truth value of our beliefs or our belief attributions. In fact, according to Brad Armendt "The idea that beliefs are stake-sensitive is an interesting one, but there is little evidence to indicate that philosophers (especially epistemologists and logicians) have considered or taken it very seriously". (Armendt, 2010, p.2).

That belief is not affected by stakes in cases like the ones shown is not so entirely obvious. Brad Armendt (2008; 2010) claims that "There are any number of things we do in the neighbourhood of believing that are properly sensitive to contexts and to what is at stake in those contexts. Choosing, acting, accepting, asserting, having justification, and knowing are examples of the sorts of things I have in mind, and the list can no doubt be extended" (Armendt, 2008, p. 2). For some of these attitudes, sensitivity to stakes has actually been taken seriously, and the fact that they have might be a motivation for us to think that belief might also be dependent upon contextual parameters. Armendt recognizes that it is difficult to determine whether beliefs are stake-sensitive, however, he suggests it is an idea worth exploring. Here I intend to take Armendt's suggestion seriously and try to test the sensitivity of belief to stakes by testing whether our practices of belief reporting are sensitive to stakes.

Taking into account that attributing knowledge presupposes the attribution of belief, it would be interesting to explore whether patterns of sensitivity of belief attribution are affected by what is at stake in cases similar in relevant features to the ones presented. I suggest that if the variability of belief reports across contexts turns out to be the case and if it is affected by this particular parameter, this might have implications on the way variation in cases of knowledge is estimated. If no evidence of variation of judgements towards BRs is found in empirical research, the result does not affect epistemological claims. However, I consider that if through the experimental work that this dissertation presents, such pattern of variation in attributions of belief is confirmed in non-

philosophers, this might be significant in further developments of our theory of knowledge. I hope, this can be an input for future theoretical elaboration.

2.5. Stakes and cognitive processes

In recent works, Jennifer Nagel (2008; 2011) has proposed that the variations in judgement we observe in cases of knowledge attribution where stakes change, such as the Bank cases, might be explained by psychological facts. Nagel proposes different psychological mechanisms to explain epistemological problems, adducing that independently of pragmatics and semantics, there is a cognitive base in our reasoning about puzzling epistemological cases.

I won't discuss the details of Nagels' argument for adjusting cognitive models to the explanation of epistemological puzzles, but I will make emphasis instead on the mechanisms themselves and their relation with evaluation of situations where stakes are altered, showing that the conditions for reasoning in knowledge attributions might be similar enough to those in belief attribution

Nagel proposes that the conflicting judgements we observe in the Harman-Vogel paradox are naturally explained by Dual Process Theory, an approach that defends that we use different cognitive modules to treat heuristic, low-stake cases on one hand, and novel or high-stake cases on the other. According to this view, we reason about situations in life making use of two cognitive modules referred to as *System 1* and *System 2*. System 1 is an automatic cognitive mode that operates swiftly and without effort, used to judge familiar cases where stakes are low. On the other hand, System 2 is a more controlled mode of cognition that engages on a serial character processing. We use this module to reason about complex situations where stakes are high or that are novel to us. Also negating is a process we engage with System 2.

Apparently, our way of reasoning is completely different when we are in high and low stakes situations. High-stakes subjects try harder than low-stakes ones when asked to solve the same problem; they think more systematically, do not rely on heuristic or automatic responses and need a greater amount of information/evidence to make a distinction confidentially. Nagel provides empirical evidence from diverse sources of how raising a subject's perceived stakes tends to activate System 2, even for problems which would ordinarily be handled by System 1. This means that even if we are judging

a common situation that normally we would asses with System 1, the more controlled cognitive mode of System 2 is activated when stakes are increased. As a consequence of this cognitive shift, we use System 1 to evaluate lax sentences in the paradox, and use System 2 instead when judging stringent ones.

Our natural way of evaluating situations in high and low stakes has as a result that a given proposition can manifest knowledge when judged under one manner but not the other. This shift in judgements we see in paradoxes and some problematic cases in epistemology, seems to obey to the change in systems of evaluation due to several mechanisms of our cognition that are activated depending on stakes. It is clear that context triggers different cognitive modules, but, what is the mechanism that impulses this shift? A possible answer comes from the research of psychologist Arie Kruglanski (cited by Nagel, 2008), who introduces the concept of need-for-closure.

Closure is understood as arrival at a settled belief, a judgemental commitment that is firm instead of hesitant. Kruglanski explains that when facing a task or a problem, subjects hold different degrees of need-for-closure, ranging from a pressing need to a very low need and avoidance of closure. A series of experiments conducted by Kruglanski show what he calls the "Unfunded Confidence Paradox"; this is the phenomenon that the higher the need-for-closure, the greater the confidence on less evidence. Interestingly, high-stake situations tend to induce low need-for-closure, while low stakes subjects experience a more neutral need-for-closure, in Nagel's words: "When cognition seems costly, we want to get it over with; we think up fewer hypothesis and feel better about the one chosen" (Nagel, 2008, p. 289).

Following this reasoning, Nagel concludes that the reactions towards the Bank cases and similar thought experiments where there is variation in stakes, is better explained by the way our cognition is structured. In this sense, what has been typically studied from a semantic perspective, argues Nagel, is probably exclusively correspondent with the way we psychologically process high and low stakes situations. If this turns out to be the case, Contextualism and IRI would be required to find stronger evidence for their claims, as in words of DeRose, "The best grounds for accepting Contextualism concerning knowledge attributions come from how knowledge-attributing (and knowledge denying) sentences are used in ordinary, non-philosophical talk: What ordinary speakers will count as "knowledge" in some non philosophical contexts they

will deny is such in others" (DeRose, 2006, p. 316). It can be said that IRI relies on the intuitions elicited by the cases just as much as Contextualism.

It is not my aim to pursue Nagel's strategy here or to support her claim against Contextualism or IRI. All I'm interested in is to show this alternative which, in further chapters I intend to argue, can also be adequately followed for belief reporting. I won't claim that the empirical evidence here collected is in any way supportive or disconfirming of this view, but it certainly offers an attractive input for discussion.

As I will show next, some experimental proposals have been put forward in epistemology testing some of the cases mentioned or different vignettes designed to test whether what is at stake can influence truth value of our knowledge attributions. These are tested on non-philosophers to see whether the alleged variation in intuitions towards low and high stakes contexts is shared by them. This has generated great controversy in the field, and literature that reviews the methodology of these experiments and explores new ways of approaching the matter has rapidly increased. Keeping in mind the connection between belief and knowledge, and given that philosophical experimentation on belief attribution is non-existent, I think it is reasonable to expect that examination of experimental attempts in epistemology will be constructive for the purpose of addressing belief reporting empirically. A review of some of experimental works on epistemology and how this precedent can be useful for designing future experiments on belief reporting is the topic of the following chapter.

Chapter Three

Experimental Philosophy

Experimental philosophy or X-Phi, is a recent branch in philosophy that brings philosophical questions to empirical experimentation on the assumption that it is important to take into account the intuitions of the laymen and not only those of "*isolated philosophers*" who are thinking about what seems natural or intuitive to them. The pioneers of the movement, Stephen Stich, Joshua Knobe and Shaun Nichols (among others) contend that analytic philosophers have been using an incomplete methodology, namely the exclusive use of one's own intuitions. This, they sustain, at the very least needs to be complemented by intuitions of non-philosophers so as to validate theoretical assumptions. As Knobe (2003) says, "it can sometimes be helpful to ask ourselves what people would ordinarily say about the situation under discussion" (Knobe, 2003, p. 2).

The thought that fuels the X-Phi movement is that it might be useful to check whether the intuitions of a single philosopher, which are supposed to generalize what people would do or think in a particular context, are shared by those people, by the laymen. In view of that, experimentation comes as a useful guidance for elucidation of philosophical matters, not with the intention of finding an absolute answer in what ordinary people think, but with the expectation that turning to them might help obtain a clearer picture of the issues and help us revise our theories on the basis of what has been found. Still, people are still reflecting on the data obtained and trying to determine exactly what impact experimentation actually has on philosophical theories.

Generally, research in X-phi has been using methods that correspond to cognitive sciences and cognitive psychology such as surveys, brain scan analyses, psycho-physiologic techniques, observation of behaviour under experimental conditions, and such the application of these methods has given very controversial results that have provoked a lot of discussion and growth around questions related to intentional action,

determinism and responsibility (research on free will, moral and emotional implications on one's judgements on responsibility), and epistemology. Philosophical experimentation has been prolific; but the field of belief attribution is still unexplored. Even though there are several works on belief in areas like belief acquisition in children and adults, philosophical theories on belief and belief reporting have not been tested empirically.

Recently, several authors have taken interest in testing empirically whether the intuitions on context sensitivity of Knowledge Attributions that are accepted by philosophers are supported by non–philosophers, based on the assumption that data about the mental states and attitudes of ordinary people is significant in assessing epistemological concepts and theories. The literature in epistemological X-phi is progressively larger, and even though we are not near to arriving to a consensus about which methodology is adequate for addressing philosophical questions empirically and collecting relevant data, some progress has been made.

I here depart from the assumption that experimental works on areas related to belief reporting such as knowledge attribution, can probably guide the way to my objectives. With that in mind, this chapter will be dedicated to methodology in experimental philosophy. I will introduce in the first section different experiments previously performed by different authors in epistemological X-phi, in order to illustrate and analyse with detail the methodology followed and the probes employed, exploring what features in them might be responsible for the type of data collected. I will then present some methodological remarks, exploring criticisms to the experiments presented and bringing up other experiments that follow unconventional approaches. I will also examine other ways of experimenting in areas besides epistemology, with the aim of constructing a path for improvement that hopefully allows me to make documented and reliable experimental choices for achieving tests that report consistent data on belief attributions.

3.1. *Putting thought experiments to test*

The aim of this dissertation is to explore the possible context-sensitivity of belief attributions via empirical testing, following the lines of previous works on experimental philosophy, mainly, on epistemology. The methodology of these experiments has conventionally been to reproduce the thought experiments found in the literature, and

testing if variables such as the amount of evidence one has or the importance of being wrong have an influence in our attributions.

I will present in detail, showing thoroughly the methodology applied, two prototypical studies where the thesis that the value of knowledge attributions vary according to conversational context has been tested. Wesley Buckwalter (2010) and Mark Phelan (2010), each offer a particular interpretation of contextualism and several series of experiments that allow them to reach the conclusion that knowledge attributions are not affected by contextual variables. In what follows I present an overview of both studies focussing on the variables tested and the methodology followed, aiming to claim that there are important experimental mistakes that invalidate the conclusion reached in the tests performed.[43]

3.1.1. Testing thought experiments directly

Wesley Buckwalter (2010) explores the pattern of folk knowledge attributions testing two fundamental standards of conversational context that affect truth value: *Stakes* and *Salience of Error Possibilities* (such as counterevidence), proposing alternative explanations of sceptical challenges. Based on contextualist theories, one could expect that according to particular contextual standards truth value of knowledge attribution would vary so that to higher stakes or greater salience of error possibilities, lower knowledge is attributed. Thus, the experimental hypothesis that the author puts to test is whether "[...] the truth conditions of knowledge claims in the following cases will fluctuate when the stakes or error possibilities of those cases are raised" (Buckwalter, 2010, p. 7).

To examine this, the author designs three probes based on DeRose's bank cases and test them in 544 undergraduate students, between 18 and 20 years of age, and 55% of them males. The design consists in three vignettes, one where it is not particularly important that the checks are deposited, and a second one where stakes are high and a third one where there is salience of error possibilities. The vignettes are the following:

> BANK. Sylvie and Bruno are driving home from work on a Friday
> afternoon. They plan to stop at the bank to deposit their paychecks, but as

---

[43] Apart from my own observations to arrive to this conclusion, I also mention others who have made criticisms o these tests, such as Pinillos (2012; 2010) and Hansen (forthcoming (a); forthcoming (b)).

they drive past the bank they notice that the lines inside are very long. Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. When they return to the bank on Saturday, it is open for business.

BANK-HS. Sylvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks. Bruno has written a very large check, and if the money from his pay is not deposited by Monday, it will bounce, leaving Bruno in a very bad situation with his creditors. As they drive past the bank, they notice that the lines inside are very long. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. When they return to the bank on Saturday, it is open for business.

BANK-ES. Sylvie and Bruno are driving home from work on a Friday afternoon. They plan to stop at the bank to deposit their paychecks, but as they drive past the bank they notice that the lines inside are very long. Although they generally like to deposit their paychecks as soon as possible, it is not especially important in this case that they be deposited right away. Bruno tells Sylvie, "I was just here last week and I know that the bank will be open on Saturday." Instead, Bruno suggests that they drive straight home and return to deposit their paychecks on Saturday. Sylvie says, "Banks are typically closed on Saturday. Maybe this bank won't be open tomorrow either. Banks can always change their hours, I remember that this bank used to have different hours." When they return to the bank on Saturday morning, it is open for business. (Buckwalter, 2010, p 8)

Participants were asked to specify if they considered true the knowledge statement that the character makes in each vignette, using a Likert scale as follows: "On a scale of 1 to 5, circle how much you agree or disagree that Bruno's assertion, "I know the bank will be open on Saturday" is true" (Buckwalter, 2010, p. 9). Two tests were performed, one

comparing BANK and BANK HS to see folk model of agreement to certain truth conditions between groups in high and low stakes probes; a second test for comparison between BANK and BANK-ES to see the pattern of accord toward cases of low and high error possibility. If the hypothesis was correct, one should expect the "mean level of agreement with the target question will decrease below 3 between groups in both tests" (Buckwalter, 2010, p. 9), that is, there should be a statistically relevant difference in levels of agreement towards the knowledge claim as the stakes or the error possibilities are raised.

If stakes mattered to knowledge attributions, the results of Test 1 should have shown that people are more likely to agree that the knowledge claim in the vignette is true when the low stakes situation is presented. However, there was no significant statistical difference in agreement scores between the groups of BANK (mean score 3.83) and BANK-HS (mean score 3.71). Also, both groups' means are greater than 3. Given that a score of 3 constitutes agreement, the author concludes that the hypothesis should be rejected and therefore, that the contextualist's prediction on the pattern of knowledge attribution didn't show

Likewise for test two, if the salience of error (low and high epistemic standards) was influential for knowledge attributions then speakers would accept easier that the character in the story possesses knowledge in low error conditions than in high error ones. Instead, what we see with the comparison between BANK and BANK-ES is that subjects generally show a stable level of agreement in both cases. Therefore there is not significant statistical difference between BANK (mean score of 3.64) and BANK- ES (mean core 3.83). Again, the mean scores are significantly greater than 3, results that contradict the hypothesis showing that the answer patterns of non-philosophers don't go along contextualist estimates.

To summarize the scores, sample and standard deviation for each vignette Buckwalter offers us the following table:

Survey Type	Sample	Mean	Std. Variation
BANK	183	3.83	1.065
BANK-HS	181	3.71	1.108
BANK-ES	180	3.64	1.102

TOTAL	544	3.73	1.093

Table 1. Sample size, mean, and standard deviation for all probe types. (Buckwalter, 2008, p. 11)

As a conclusion, the author sustains that even though testing a few isolated cases from the literature in epistemology cannot and should not be taken as a proof against the thesis that knowledge attributions vary across contexts, the little divergence shown by the examinees at least warns us about a chance that the conclusions reached by contextualists about intuitions in cases like those presented may not be correct. The author remarks the importance of future testing to get more accurate and trustable results, and limits himself as regards further elaboration on consequences of the experimentation to this field of philosophy. As we will see in the next subsection, this conclusion is congruent with what others found out running a slightly different kind of experiments.

3.1.2. Altering thought experiments based on conceptual assumptions

Mark Phelan (manuscript) focuses on Anti–Intellectualist accounts about knowledge such as Jason Stanley's, which sustain that the cost that being wrong might have for the reportee matter to knowledge attributions. Apparently the theme is really similar to the study previously summarized, as Buckwalter also takes into account the views that endorse the relevance of stakes in knowledge attribution. However, Phelan's study includes new elements that were not taken into account in Buckwalter's illustration of the problem.

The author here includes *evidence* as a central piece for the characterization of knowledge attributions. He takes into account the important connection that exists in epistemology between knowledge and evidence, and justification and evidence, and understands that the role of stakes or practical costs in the analysis of knowledge attributions is closely tied to evidence. According to his description, Anti-intellectualism about evidence (AIE) can be expressed in two ways, one closely related with evidence and the second to do with the quality of that evidence: "Whether p is an element of X's evidence set at a particular time t constitutively depends upon the costs of X's being wrong about p at t. The higher those costs, the more stringent are the conditions required for p to be an element of X's evidence set at t." (Phelan, manuscript, p. 2) He also alludes to Stanley on the link between the quality of one's evidence and

the costs of being wrong, saying that "The quality of evidence for a person X at a time t provided by some information, or some method of gaining information, depends upon whether it is used to support X's belief in a proposition that is a serious practical question for X at t." (Stanley cited by Phelan, manuscript, p.2).

On the basis of this characterization, Phelan's idea is to test anti-intellectualist intuitions on cases modelled to address whether stakes are connected to evidence, in order to draw conclusions about the possible difference these elements might make in truth value of knowledge attributions. Thus, the experimental question that guides the research is "Do our intuitive judgments about cases support the hypothesis that there is a relation between one's evidence and the practical costs of being wrong?" (Phelan, manuscript, p. 6). According to the author, it is extremely important for theories that are constructed on the basis of hypothetical cases which aim to be generalized and resume a standard attitude towards a specific situation, that they be supported by the data coming from non philosophers. Consequently, the experimenter suggests that if it is found that intuitions with respect to the cases are not as predicted by Anti-intellectualists, the view should be thoroughly revised.

The way Phelan approaches the experimental question has certain particularities as he believes there is no ground to think that the common use of the word 'evidence' is in any way similar to the way epistemologist use it. For this reason, he resolved that for gathering "intuitive judgments about whether or not a particular proposition, person, time triple falls into the extension of the theoretical concept of evidence" (Phelan, manuscript, p.6), it was necessary further elaboration on this concept as it appears in folk's usage. For this purpose, the author devises a procedure as follows:

> My procedure is to locate a point of contact between ordinary terms of epistemic appraisal, on the one hand, and the epistemologist's concept of a person's evidence, on the other. I assume that one point of contact […] is between the epistemic notion of evidence and the ordinary concept of confidence. Specifically, I assume that the more or better one's evidence for a proposition, the greater one's confidence in that proposition ought to be, and vice versa.
>
> I'll put this point as follows:

Bridge from Rational Confidence to Evidence (BRCE): Peoples' implicit commitments about an agent's evidence set or quality of evidence are reflected in their explicit intuitive judgments about how confident that agent ought to be in various propositions supported by that evidence. (manuscript, p.6-7)

Thus, the idea is that measuring confidence of non philosopher's attitude towards the cases, the author can assess what epistemologists call evidence, and through it, knowledge. Phelan presumes that confidence is a valid indicator of how much a person commits to a determined proposition provided certain evidence. He explicitly states that this is the only point of contact between evidence and confidence he could find and that he would be open to other suggestions about what could represent the epistemic concept of evidence in people's language. [44]

Having accepted these suppositions, Phelan runs a study to measure whether folk's intuitions are compatible with Anti-intellectualist theories of knowledge. To do so, he proposes four studies: the first one aimed to test AIE. If AIE was correct (and also BRCE was), agents should be more confident about p in cases where the practical costs of being wrong about p are low, and be less confident when they are high. In this first study he gave a population of 71 undergraduate students, cases where the stakes were low, cases where they were high, and cases where low and high stakes are juxtaposed. The vignettes used were identical except for the parts underlined, which were modified to make the costs of being wrong vary in each case as follows:

> *Unimportant (Passerby):* <u>Kate is ambling down the street, out on a walk for no particular reason and with no particular place to go</u>. She comes to an intersection and asks a passerby the name of the street. <u>"Main Street," the passerby says</u>. Kate looks at her watch, and it reads 11:45 AM. Kate's eyesight is perfectly normal, and she sees her watch clearly. Kate's hearing is perfectly normal, and she hears the passerby quite well. She has no special reason to believe that the passerby is inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather

---

[44] This reveals how big the assumption they are making is and how controversial might be their definition of a BRCE. Clearly if this definition is not correct then there is not a valid conclusion that can be extracted out of analysing people's confidence, which can affect the relation between the costs of being wrong and the evidence one possesses. It seems quite dubious that the construction made by the authors corresponds to an actual connection either in our brains or in our linguistic practices.

further evidence that she is on Main Street (she could, for instance, find a map), but she doesn't do so, since, on the basis of what the passerby tells her, she already thinks that she is on Main Street." (Phelan, manuscript, p. 8)

Important (Passerby): Kate needs to get to Main Street by noon: her life depends upon it. She comes to an intersection and asks a passerby the name of the street. "Main Street," the passerby says. Kate looks at her watch, and it reads 11:45 AM. Kate's eyesight is perfectly normal, and she sees her watch clearly. Kate's hearing is perfectly normal, and she hears the passerby quite well. She has no special reason to believe that the passerby is inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, find a map), but she doesn't do so, since, on the basis of what the passerby tells her, she already thinks that she is on Main Street. (Phelan, manuscript, p. 20)

Two streets (Passerby): Kate needs to get to Main Street by noon: her life depends upon it. She comes to an intersection and asks a passerby the name of the street. "The street you're on is 'State Street' and this other street is 'Main Street''' the passerby says. Kate looks at her watch, and it reads 11:45 AM. Kate's eyesight is perfectly normal, and she sees her watch clearly. Kate's hearing is perfectly normal, and she hears the passerby quite well. She has no special reason to believe that the passerby is inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, find a map), but she doesn't do so, since, on the basis of what the passerby tells her, she already thinks that she is on Main Street. (Phelan, manuscript, p. 20)

The subjects were asked three different questions depending on the case they received, but for the three cases it was constant to ask "How confident should Kate be that she is on Main Street?" In the *Two Streets* case they were also asked "How confident should Kate be that she is on State Street?" The other questions were not considered relevant and don't appear in the paper published. The answers were classified in a seven point

scale ranging from 'not confident' to 'very confident'. The results obtained show no statistically important difference in the answers to *Important* (average punctuation 5.29) and *Unimportant* (average score 5.00). Given that no difference appeared in the results towards *Important* and *Unimportant*, Phelan looked at Two streets to determine "what explains the motivation for AIE and the significance that epistemologists attach to the topic of anti-intellectualism?" (Phelan, manuscript, p.10), as judging by the results, there is no difference between high and low stakes cases.

Let's have a close look to the third case. Remember that the researcher's intention was to create a situation where there was a contrast between high and low stakes. This vignette is identical to *Important* except that the passerby gives information that includes two variables (instead of one) to Kate. It is difficult to find how exactly low and high stakes are supposed to combine, but as Phelan explains it, with the alternative option present the stakes were high for the information about Main Street but not for State Street. The results in this vignette are supposed to show that the difference in people's judgements about the confidence an agent shows towards important and unimportant information was significant: "There was a highly significant difference between subjects' average confidence ratings for unimportant State Street (5.39) and their average rating for important Main Street (3.13)" (Phelan, manuscript, p.10).

However, seeing the design of the third vignette and the questions asked, one might say that it is possible that rather than the contrast, the difference spotted is due to difficulties in the design. Observe that in this vignette the passerby says clearly "The street you're on is 'State Street' and this other street is 'Main Street'", then people are asked about the confidence Kate should have about being in State Street and in Main Street on the assumption that whatever difference is found in the answers, should be due to the fact that Kate's life depends on being at Main Street, while there is nothing particularly at stake for her about State Street. But in the information given by the passerby it is emphasized that she is on State Street, while about Main Street the passerby says it is "this other street". It seems likely to me that the difference in the information given about both streets can be a source for the difference in judgements about confidence observed, not the difference in what is at stake. However, this is only an observation that would need to be tested and revised by making information about both streets more homogeneous.

The author concludes that the results of the first test show that whereas no significant difference was found between *Important* and *Unimportant* in non-juxtaposed cases, the difference was 7.50 times larger when the judgements were made on juxtaposed vignettes. This shows that "70% responded to the juxtaposed case in a way consistent with the hypothesis that a piece of information's status or quality as evidence constitutively depends upon the costs of being wrong." (Phelan, manuscript, p. 11).

Looking for an explanation for the asymmetry found, Phelan hypothesizes that even though the importance of being wrong doesn't usually play a role in people's assessments of an agent's evidence, when confronted with juxtaposed cases people tend to judge "as though a proposition's status or quality as evidence constitutively depended upon the practical costs to an agent of being wrong about it" (Phelan, manuscript, p. 12), that is, as though Anti-intellectualism about evidence was correct. He conducts a second study to test this hypothesis. This study, therefore, aims to neutralize possible effects of the language or the situation portrayed in the first experiment, in order to evaluate why people faced with juxtaposed cases tend to show intuitions that seem to support Anti-intellectualism. In this study 72 undergraduates were presented with two identical vignettes, *Important* and *Unimportant*. The juxtaposition in the second study, as opposed to the first one, was created by handing both *Important* and *Unimportant* cases to the subjects to judge. The vignettes are the following, altered in the underlined parts:

> *Unimportant (Drunks)*: <u>Kate is ambling down the street, out on a walk for no particular reason and with no particular place to go. She comes to an intersection.</u> Two drunks are standing on the corner in an intense discussion. Kate asks the drunks the name of the street. "Main Street," the drunks say. Kate looks at her watch, and it reads 11:45 AM. Kate's eyesight is perfectly normal, and she sees her watch clearly. Kate's hearing is perfectly normal, and she hears the drunks quite well. She has no special reason to believe that the drunks are inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, find a map), but she doesn't do so, since, on the basis of what the drunks tell her, she already thinks that she is on Main Street. (Phelan, forthcoming, p. 13)

Important (Drunks):   Kate needs to get to Main Street by noon: her life depends upon it. She comes to an intersection. Two drunks are standing on the corner in an intense discussion. Kate asks the drunks the name of the street. "Main Street," the drunks say. Kate looks at her watch, and it reads 11:45 AM. Kate's eyesight is perfectly normal, and she sees her watch clearly. Kate's hearing is perfectly normal, and she hears the drunks quite well. She has no special reason to believe that the drunks are inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, find a map), but she doesn't do so, since, on the basis of what the drunks tell her, she already thinks that she is on Main Street. (Phelan, forthcoming, p. 21)

In this vignette, instead of a passerby, the main character asks some drunks -an unreliable source of information- about her current location. The scoring mechanism and the questions were the same as in the first study. The results confirmed that the subjects' judgements were very close for *Unimportant* (4.72) and *Important* (4.39) cases when run independently. Contrary to this, when the cases were juxtaposed, the difference in judgements was salient: for the *Unimportant* case the average score was 4.53, while for the *Important* probe it reached a 3.69: 42% of the subjects emitted judgements consonant with Anti-intellectualism.

Phelan reports that ordinary factors that intervene in epistemic appraisal, such as the reliability of the source of information -the passerby and the drunks- in the cases presented, can be of help to explain the differences obtained between the first and the second study. Along with Phelan's explanation, it is important to remark that the difference between the results in the isolated cases and the juxtaposition is importantly smaller than in the first study. This might be due to the particularities signalled about the methodology of contrast in the first study, and also to the fact that the juxtaposition as it is done here is very different to what was proposed in the first study. However, it would be necessary to confirm these hypotheses in further research.

Adding the results obtained in the first and second studies, the author concludes that "Importance does not usually factor into peoples' assessments of evidence" (Phelan, manuscript, p. 15), that is, that anti-intellectualism as it is described by the authors

(related to evidence which is assessed through confidence) is not supported by the data found. However, having as a continuous result that the participant's judgements don't go along Anti-intellectualist predictions when cases are presented separated but they do when juxtaposed, Phelan decides to conduct yet a third study to test whether non-Anti-intellectualist intuitions are produced by some kind of difficulty that prevents assessment of subject's authentic epistemological intuitions in cases where the information cannot be contrasted. The question for this round of experiments was whether epistemic factors were capable of affecting confidence judgements in non-juxtaposed cases. For this study 141 undergraduate students were tested with a model similar to the second study with the following vignettes modified where underlined and juxtaposition created by handing both cases to subjects. Street signs were the source of information this time:

> *Unimportant (Street Signs):* <u>Kate is ambling down the street, out on a walk for no particular reason and with no particular place to go.</u> At 11:45 AM, she comes to an intersection and looks up at a street sign that says "Main Street". She looks at her watch, and it reads 11:45 AM. The ambient lighting is perfectly normal. Kate's eyesight is perfectly normal, and she sees the street sign clearly and sees her watch clearly. She has no special reason to believe that the street sign is inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, ask a passerby or find a map), but she doesn't do so, since, on the basis of the street sign, she already thinks that she is on Main Street. (Phelan, manuscript, p.28)

> *Important (Street Signs):* <u>Kate needs to get to Main Street by noon: her life depends upon it.</u> At 11:45 AM, she comes to an intersection and looks up at a street sign that says "Main Street". She looks at her watch, and it reads 11:45 AM. The ambient lighting is perfectly normal. Kate's eyesight is perfectly normal, and she sees the street sign clearly and sees her watch clearly. She has no special reason to believe that the street sign is inaccurate. She also has no special reason to believe that her watch is inaccurate. Kate could gather further evidence that she is on Main Street (she could, for instance, ask a passerby or find a map), but

she doesn't do so, since, on the basis of the street sign, she already thinks
that she is on Main Street. (Phelan, manuscript, p.28)

The results showed, again, no significant difference between *Important* (5.79) and
Unimportant (5.61) cases when presented separated, but when contrasted, the
judgements changed and there was a (small) difference between *Unimportant* (5.76) and
Important (5.26) vignettes. With the information obtained in the third study Phelan
compares responses in non-juxtaposed cases where the source of information was
highly, moderate and not reliable, concluding that indeed the confidence lowered
notably in line with the reliability of the source of information, while the importance
remained stable in the three studies. Upon comparative analysis of the three studies
conducted the author concludes that reliability and not importance have a marked effect
on subject's epistemic judgements. He also concludes that "anti-intellectualism appears
truly predictively impotent in non-juxtaposed cases." (Phelan, manuscript, p. 18)

To complete his testing on evidence and its relation with the importance of being right
in juxtaposed cases, where the results suggest conformity with anti-intellectualism, the
author puts forward a fourth study. In this, 69 undergraduates were presented with the
following vignette:

Kate needs to get to Main Street by noon. She comes to an intersection and
asks a passerby the name of the street. The passerby says, "Main Street."
Kate looks at her watch, and it reads 11:45 AM. Kate wonders how
confident she should really be that she is on Main Street. (Phelan,
manuscript, p.20)

In this test the researcher asked the subjects to answer "What factors should affect
Kate's confidence that she is on Main Street?". Participants were also invited to check
all factors "that apply for nine different factors ranging from the country in which Kate
lives, to the passerby's emotional state, to the reliability of past information from
random people" (Phelan, manuscript, p. 20). Besides, subjects were asked "how
important it is that Kate be on Main Street at noon" in the vignette presented. The
results of this application show that 43% of the subjects signalled that it was important
that Kate be at noon in the on Main street, percentage that was extremely similar to that
obtained by subjects' Anti-intellectualist judgements in juxtaposed cases (mean 44%
across all previous cases). The author concludes that:

> The striking similarity of the percent of subjects who profess a general commitment to importance as one factor that should affect confidence and the percent who judge as anti-intellectualists when presented with juxtaposed cases suggests that our hypothesis is correct. Although importance does not usually factor into people's assessments of an agent's evidence in particular cases, juxtaposing important and unimportant cases or information seems to trigger, on the part of subjects who have the relevant commitment, a tendency to adjust for importance when assessing evidence. (Phelan, manuscript, p.20)

Based on the four experiments tested, the general conclusion Phelan arrives to is the following:

> I conclude that the contrast cases to which anti-intellectualists appeal do not actually support the general principle that what counts as evidence depends upon the costs of being wrong. But I have also found that many people do profess a commitment to such a general principle.

> Proponents of anti-intellectualism in epistemology generally regard the view as a surprising truth, which can be supported only by appeal to our intuitive judgments about whether certain epistemic properties are instantiated in hypothetical cases. In fact, anti-intellectualism about evidence is not a surprising truth, for it is not surprising, and it is not true. (Phelan, manuscript, p. 21).

I deem highly possible that only the difference in the design of the studies together with the dissimilitude in the questions and situations would be enough to cause the divergence in judgements the data shows and in which the author relies on to draw these conclusions. Nevertheless, this is an observation that would need to be empirically contrasted and revised before it can be seriously considered. Regardless of this, it is remarkable that the conclusion reached by Phelan is parallel to Buckwalter's: that contextualism about knowledge attributions is not supported by the empirical evidence. Moreover, in a recent study by Joshua May, Walter Sinnott-Armstrong, Jay G. Hull and Aaron Zimmerman (2009) the authors test parameter-sensitivity claims (Stanley, 2005; Schaffer, 2006) relying on a similar methodology as the studies earlier summarized. They hand the participants the Bank Cases (low and high stakes with and without

alternative, for a total of 4 cases) and evaluate their agreement to the statement that Hanna knows the bank will be open on Saturday. Their results basically go along the studies presented previously, showing that it seems that competent speakers do not differently attribute knowledge to an agent depending on what is at stake.

Essentially, then, the state of the art in epistemological X-phi shows a series of studies that develop sets of experiments following a common methodology and reaching similar conclusions. The data gathered suggest we should dismiss theories that hold EPS and uphold a position that leaves contemporary philosophers rather baffled, as EPS for the cases presented is widely accepted among the philosophical community. Philosophers' intuitions seem to converge for the variation of knowledge attributions across contexts (independently of the explanation they give to it), while X-phi endorses this is probably not correct, so there is a marked and worrying divergence.

I think it is possible that the phenomenon we are witnessing is that philosopher's *intuitions* (not their theories) are not represented in the results obtained by the experimental methods explained above. Experimental data are taken to disprove theoretical approaches; however, it is quite intriguing that the uniform pattern of intuitions found in experts differs so greatly from that found by means of surveys with folks. In fact, this has lead to major discussions about assessment of folks' intuitions and whether expertise matters for forming correct judgements on philosophical matters. In any case, when such a big accordance in intuitions happens among specialists it is worth asking if the experiments are getting it right, if they are precise enough or designed in a way that actually manages to obtain the relevant information needed. So it seems to me, on the basis of these experiments, that the question is open as to whether we are actually collecting the right data. In the following section I consider some bits of this controversy in order to argue in favour of a methodology that allows testing of the right intuitions.

3.2. Methodological considerations to improve the methodology

The main aim of X-phi is to confirm whether philosopher's intuitions toward the hypothetic situations that have been used as incentive for theories are shared by folks. The general conclusion of the experiments presented is that this is not the case. I want to argue in this section that it is highly possible that the methodologies that work for eliciting intuitions towards a case in trained people may differ from those needed for

non specialized participants. I have doubts as to whether it is a good idea to simply take the cases from the literature and test them with folks without any further contextualization. I also want to show there are probably various methodological choices in the experiments presented above that can be improved to guarantee more accurate results. In this section I present diverse methodological considerations as well as other studies that seem to show a path to advance in the direction of a reliable method in X-phi of propositional attitude attributions.

3.2.1. Granting adequate epistemic conditions

Claiming that experiments done in the past follow incorrect approaches, Angel Pinillos (2012; 2011) offers a methodological alternative for empirical assessment of knowledge attributions[45]. His work provides a critical evaluation of the methodology of previous experiments on epistemology and at the same time develops a methodology for testing, which gives results compatible with intuitions of philosophers, and therefore strikingly different from the previously presented.

According to Pinillos, previous experiments on knowledge attributions have failed to collect the data we aim for since they do not provide the participants with the *adequate epistemic conditions to* be in a position to make a correct knowledge attribution. The author believes that, subtle features of the design such as incorporating ascriptions in the stories can be confusing and produce difficulties guaranteeing that the subjects perceive that the agents in the different vignettes are in the exact same epistemic position. This, allegedly, alters judgement of the subject and can invalidate the data gathered. His proposal, then, follows a "seeking evidence" strategy (honouring the connection between knowledge and evidence) with vignettes where there are no internal ascriptions and the epistemic state of the main character is clearly kept fixed.

This is the type of vignettes used in Pinillos' (2012) study:

> *Typo-Low*: Peter, a good college student has just finished writing a two-page paper for an English class. The paper is due tomorrow. Even though Peter is a pretty good speller, he has a dictionary with him that he can use to check and make sure there are no typos. But very little is at stake. The teacher is

---

[45] It is important to mention that Pinillos' intention with the experiments here presented, is not only methodological, but mainly that of defending Anti-intellectualism and, ultimately, IRI, against the apparently challenging results obtained by previous experiments.

just asking for a rough draft and it won't matter if there are a few typos. Nonetheless Peter would like to have no typos at all.

Typo-High: John, a good college student has just finished writing a two-page paper for an English class. The paper is due tomorrow. Even though John is a pretty good speller, he has a dictionary with him that he can use to check and make sure there are no typos. There is a lot at stake. The teacher is a stickler and guarantees that no one will get an A for the paper if it has a typo. He demands perfection. John, however, finds himself in an unusual circumstance. He needs an A for this paper to get an A in the class. And he needs an A in the class to keep his scholarship. Without the scholarship, he can't stay in school. Leaving college would be devastating for John and his family who have sacrificed a lot to help John through school. So it turns out that it is extremely important for John that there are no typos in this paper. And he is well aware of this. (Pinillos, 2012, p. 13)

Immediately after the subjects were given a vignette, they were given the following prompt: 'How many times do you think Peter [John] has to proofread his paper before he knows that there are no typos? _____ times.' (Pinillos, 2012, p. 13) Subjects were told to insert an appropriate number in the blank space.

Applying this new methodology which corrects the features previously mentioned about other experiments, had as a result that folks displayed intuitions parallel to those of philosophers, that is, the data showed that the knowledge attributions do vary across contexts depending on what is at stake.

Apart from what he himself mentions, I propose there are at least four more features in Pinillos' experiments that seem to intervene in getting the results we observe.

i). Knowledge attributions are assessed through *a quantifiable discrete variable*, which makes it very easy to compare intuitions.

ii). *The attribution is made directly* by the experimental subject over the knowledge of the character in the vignette, instead of being an indirect appreciation of an attribution someone else makes in the story. All the cases in the studies presented before test the intuitions of the participant towards external cases where it is not they themselves but someone else who is the "main character" in the story, and agent's judgement is on that

character's dispositions.[46] It certainly would be an interesting twist to test cases where the agent's judgement is not made about someone else's intentions, but where he is involved and is asked to evaluate if he himself would make a certain attribution, or how correct would he consider to utter it given a specific context. In fact, this is the original form in which DeRose presented the cases: he himself was playing the main role in the bank cases where knowledge attributions were made to himself, in introspection. I reckon that making this small twist might have an important impact on judgements as it should give the examinee the emotional and cognitive background to enable the researchers to elicit the intuitions we target. The reason this change of perspective can be of importance is that actually the intuitions we are interested in are those of the participant, not those he thinks someone else might have, which are the ones that the experiments cited above are reporting.

iii). *The design of the vignette is simple and clear.* It is possible that variables such as comprehension of vignettes are affecting the participants' competence in the tasks, and therefore the results. The attributor (the experimental subject) understands that the attributee is in the same state in every scenario and the situations are familiar to the subjects. Testing the bank cases directly in undergraduates might be misleading both because they are quite difficult to follow for a non-philosophy trained public, and because, as it has been suggested by Jason Stanley (http://experimentalphilosophy.typepad.com/, 2010[47]), testing contexts that are common to the experimental subjects is a better idea than testing situations that may be alien to the subjects (it is dubious that undergraduates are familiar enough with situations such as check deposits and bank hours). Thus, my hypothesis would be that proper contextualization and a simplified structure on the tests might allow an optimum cognitive state for testing and therefore, can give us access to the correct intuitions.

iiii) On the negative side, however, it is possible that the design and the situation portrayed in Pinillos' vignettes make easy that *levels of neuroticism* of the participants interfere with the variable to test –i.e. correction of mistakes in a high stakes situation could trigger obsessive trades of a person's character, making them more prone to

---

[46] For a version of this problem also found in other experimental philosophy works see Marti, G. (2009).
[47] This comment was retrieved in the blog on November 2010. The specific page is not available at the source anymore.

elevate the number of readings one considers necessary for a task to be flawless-. However, this is a hypothesis that would need to be addressed in further research.

3.2.2. Eliminating biases

Through X-phi, validity of methods in philosophy has been questioned, and henceforth, the theories produced and the data they rely on. However, I suggest that there might be at least two reasons why there is a gap between the data obtained and the shared intuitions of philosophers, one to do with the population tested and a second one related to methodological choices for testing philosophical matters.

The first explanation goes along the lines of a debate raised within X-phi and its detractors, which can be expressed as Weinberg et al. (2010) do it in the form of a question: "Are philosophers expert intuiters?" (Winberg et al, 2010, p. 1). The central idea is that *expertise might make the difference when testing intuitions on philosophical affairs*. There are different positions towards this: On the one hand there are some[48] who defend that, just as it is in any other science, assessment of correct intuitions on higher philosophical issues requires especial training and thus experiments performed on non-philosophers may not be reliable. On the other hand, experimentalists hold that if philosophical theories aim to apply universally to human processes, behaviour or intentions, they should be testable on non-philosophers and the adequate intuitions be appraised by the laymen.

Solving this controversy might be a tough task. Still, it appears to me that perhaps there is a middle ground between these two positions: In epistemological theories and theories about belief ascription there seems to be indeed an aim to explain general phenomena. It looks as if these are not the kind of theories to which only one positive case counts as enough evidence, rather they seem to seek a global explanation and thus we should be able to test intuitions that base them in non-philosophers. Moreover, having experimental data available in these areas can be nourishing and supportive of the theories. Now, the way these intuitions are assessed in trained people ought to be different from the way they are measured in common people because training makes an important cognitive difference between the subjects.

---

[48] See Williamson, T. (2010); Sosa (forthcoming), among others, for a defence of this view. Or classic articles in psychology, like Pfungst, O. (1911); and Rosenthal, R. (1976).

It is widely known that expertise affects performance towards experimental tasks, and it is possible that such is also the case here. This doesn't mean that the intuitions can't be tested in non philosophers, rather, that we need to be able to design tests that actually get the data we aim for (in the case of epistemology, if the truth values of knowledge attributions of the participants vary across contexts, for instance). That is, we should be able to test folk intuitions but we are responsible for designing tests for non-specialized audiences and this could mean changing the format, the questions, and making several other methodological decisions in the process of testing the intuitions that ensure we are presenting the subjects with vignettes that provide the data we want.[49]

This takes us to the second possible explanation I attempt to give to the difference between the results achieved experimentally and philosopher's intuitions: *the design of the tests matters*. There are several studies that show the multiple ways in which the design of the tests can affect the data collected. This has been a topic of study in experimental and social sciences traditionally, and it has recently interested philosophers concerned about methodological issues. Recently, Nat Hansen, (forthcoming (a); forthcoming (b)) develops an argument specific to epistemology, defending that thought experiments used to support theses about parameter-sensitivity, are influenced negatively by the way they are designed. According to Hansen, it is necessary to re-evaluate the way intuitions are generated by contextualists' thought experiments in order to obtain data that is correct and free of biases[50].

Above I referred to the great agreement there is among philosophers on contextualist intuitions regarding paradigmatic cases. Well, Hansen's criticism intends to show that a possible source of such convergence is an issue with the design of the thought experiments that creates two types of bias well described in social sciences: *experimenter bias* and *exclusive reliance on absolute truth judgements.*

Experimenter bias is a methodological mistake we make when we design our experiments towards a desired outcome. It has been found that the expectations of the experimenter can affect the performance of subjects. This is the *modelling effect* and concerns the relation between the experimental subject and the experimenter, where the

---

[49] In actuality, there is a major debate over the expertise defence. I only raise the topic here to point out at methodological issues that need to be taken into account.

[50] Observe that here Hansen goes beyond a critique to experimental methodology and applies his reasoning to the production of original thought experiments found in the literature. I discuss some bits of his argument here because I think them relevant for the methodology on experimental research.

former modifies his performance to adjust the experimenter's expectations. According to Hansen, this is present in contextualist literature. In his words: "Looking back at the contextualist thought experiments with those kinds of experimenter effects in view, it is remarkable that the expectations and biases of experimenters are explicitly stated before readers have a chance to mull over the thought experiment on their own. It is therefore a reasonable worry that the appearance of agreement in intuitions about contextualist's thought experiments is the result of experimenter bias" (Hansen, forthcoming (a), p. 9)

The idea, then, is that context-shifting thought experiments would be more reliable if they allowed the reader to form her own idea of the case and answer freely before she is offered the experimenter's answer. A clear way of controlling this bias, then, is to refrain from offering assertions about one's own intuitions and be very careful that the design of the experiment does not reveal them. Also, it is possible to *counterbalance* opposed biases, that is, instead of avoiding judgement, one may accompany it with opposed judgements that give the experimental subject other ideas amongst which to form her own criteria.

On the other hand, Hansen mentions *exclusive reliance on absolute truth value judgements*, which is the tendency in contextualist literature to force absolute responses about the truth value of assertions in the thought experiments, not allowing subjects' reluctance to judge if a sentence is true or false in certain contexts. The author holds that "Through the range of possible intuitions about what is said in contextualist thought experiments is not explicitly limited to intuitions about truth and falsity, contextualists implicitly limit the possible range of response to those two options by not making the possibility of different kinds of responses salient" (Hansen, forthcoming (a), p. 12).

The idea behind this problem, is that it is possible for subjects to be baffled, or give "a shrug of the shoulders" when faced with the cases, simply because they do not know some of the concepts in the assertion. For example, one can ignore what it is to *know* something, or for the case I'm interested in, it is possible for one to not know what exactly is to *believe* something. Therefore, it is an interesting suggestion to allow a wider range of responses to the cases so that one can measure more accurately the answers to the tests. Also, as I mentioned in previous sections, the fact that belief might not be representational and discrete suggest that the tentative of measuring intuitions in

absolute terms might not be adequate. Instead, a methodology that allows for the variation in responses allows both the possibility of belief being gradable or discrete.

As a way of controlling this bias, Hansen proposes to allow for *dominant* and *dominated* responses, that is, allow for the subject to choose two responses which seem correct to some degree but, one is more inclined, or has a stronger intuition, towards one than towards the other. On his view, "once options are expanded beyond the stark choice between *true* and *false*, it is far from obvious that intuitions line up in the neat way that contextualists claim" (Hansen, forthcoming (a), p.18), encouraging a revision of the data that bases contextualist approaches.

To be methodologically coherent with the decision of giving subjects the option to say their dominant and dominated intuitions, you must not ask for Yes/No answers, but rather grant them to rank their response in a Likert scale. Other useful tool is to ask open questions; however, this introduces methodological difficulties that would require a combination of qualitative and quantitative approaches to analyse the data.

As we see, designing good experiments that address the intuitions we target is a difficult and demanding task, and following the argumentation here presented it is possible that further research with improved methods would be necessary. In the next section I will present an example that lacks some of the methodological complications presented here, and portrays other forms of making experiments as well as the importance of following an experimental method of repeated testing in order to corroborate the results and refine each time more the tests.

3.2.3. Refining test administration

Other examples of experiments outside experimental epistemology seem to also contribute to produce a methodology that can elicit solid data in this particular field. This is the case of experiments in cognitive and developmental psychology that aim to understand formation of belief and understanding of other minds in children. One of them is the so called *False Belief Task*.

For years psychologists and philosophers were convinced that children younger than 4 were incapable of understanding that belief can be false, until a series of experiments with a different methodology introduced new results that contradicted those findings. I deem possible that methodological conclusions based on their experiments can also be a

referent for designing vignettes in this project. I'll explain this in depth in the following lines so that my point is clear enough.

Psychologists, from 1983 performed various experiments to determine the cognitive development in children, between them a very famous task to test *false belief* understanding. The classic version by Wimmer and Perner (1983) with its results is the following:

> The child is told a story in which another child, 'Maxi' say, puts some chocolate in a green cupboard and then he goes out to play. While Maxi is gone, his mother moves the chocolate from the green cupboard to a blue cupboard. Maxi then returns, seeking his chocolate. Question: Where will Maxi look for the chocolate? Three-year-old children tend to answer that Maxi will look in the blue cupboard, where the chocolate really is, while 4 or 5-year-old children tend to answer as an adult would, that Maxi will look in the green cupboard. (Schwitzgebel, 1999 p. 284)

This experiment was replicated and many variations with a similar structure were performed getting the same results: that acquisition of false belief understanding occurs around the four years of age. For years the scientific community was convinced that the experiments were getting the correct data, until new experiments appeared showing belief understanding from an earlier age than four. Variations on the experiments where the design emphasises on how fun is to trick someone or searches for explanations of past behaviour instead of prediction of future behaviour, yield results that suggest that 3-year-olds recognize people can have mistaken beliefs. This is an example:

> Bartsch and Wellman [1989] presented children with a Band-Aid box and a plain white box and showed them that the plain white box but not the Band-Aid box contained Band-Aids. A puppet with a cut, "Bill", was then introduced, who went to the Band-Aid box and started to open it. The children were then asked to explain why Bill was doing that. If the children did not mention anything about the puppet's beliefs, the where prompted with the question, "What does Bill think?" Three year olds tended to say that Bill thought there were Band-Aids in the box. In this experiment -which relies in the explanation of past behaviour rather than the prediction of

future behaviour- 3 year olds' responses seem o suggest that they do realize that people can be mistaken in their beliefs. (Schwitzgebel, 1999, p. 284)

With this example it is evident how important small details can be in the design of an experiment to get the correct data for matters such as belief. It is likely that also in the particular case of belief attribution, devising the right questions and experimental situations can be crucial to allow us get the data we aim for. Even though the general features of the experiments are comparable, the conditions of administration seem to give finer results. This reiterates the importance of controlled conditions and stimulus – i.e. cases that are familiar, short, easy to approach, with a structure that is looked for in order to assure fair and valid comparison if needed-.

Similarly, the evolution the task has gone through and the results of its refinement make salient the need for following a method of re-applications through time. It can take years to get to a point where methodology has improved in a way that allows us to assess more accurately the data. If we apply this to the experimental philosophy field, it becomes obvious that all the discussion around methodology and the different attempts to test a single phenomenon are part of a path that eventually will get us to a more precise, better methodology for experiments on philosophical matters. I hope this dissertation makes a contribution to this end.

Chapter Four

Methodology and Results

4.1. Methodology

In this section I plan to explain the methodology followed for the experimental part of this study. I will present the details of the design, data sources and methodological decisions made in the process.

4.1. Research design or tradition

The present is a quantitative theoretical study that aims to answer the research questions and test the hypotheses made previously about the sensitivity to contextual parameters of belief reports. It is an exploratory observational study that makes use of statistical surveys in the form of vignettes for data collection and analysis.

Vignettes are short descriptions of hypothetical situations designed to assess judgments on other people's or one's own ideas, actions, beliefs, etc. Applying vignettes as an experimentation method, is part of the empirical research tradition in social sciences, and has been increasingly popular to measure people's attitudes towards certain objects. Also, vignettes are especially helpful to allow exploration of possible actions in context, to clarify people's judgments, and to grant a less personal and therefore less biased way of looking at delicate topics. Likewise, well designed vignettes are made to emulate real world situations but narrowing down the number of factors in play, which makes them very useful to isolate the judgments one wants to elicit.

The way vignettes are employed in this research integrates survey methodology and experimental designs. Relevant features of the object of study -3rd person belief attributions- are controlled thanks to a modeling that relies on a specific experimental design. Thus, the surveys were designed to observe the behavior of two independent variables: Knowledge of the Audience and Stakes. They were handed out randomly, on groups of participants in a controlled application setting.

As an exploratory study that approaches BR's from an experimental point of view for the first time, it centers on testing the most basic issues as proposed in philosophers' examples (RKotA) and as suggested from the literature in epistemology (Stakes). The results here obtained are intended to inspire new lines of work that lead to a more comprehensive and deep study in the future. Multiple suggestions for further work can be found in chapter 5, where, on the basis of the analysis of the data gathered possibilities of refinement and improvement of the work are revealed.

4.1.2. Participants

228 undergraduates at the Universitat of Barcelona and the Universitat Autònoma de Barcelona majoring in philosophy and philology participated. All participants were fluent speakers of Spanish, 149 tests were answered by females and 157 by males, with a mean age of 23.94[51]. Most participants were majoring in philosophy and philology, in their first years. Students with advanced knowledge of philosophical fields related to the study -such as those which had taken courses on epistemology and were familiar with contextualist theories- were excluded. Seven master students doing undergraduate courses as a prerequisite for their MA were included since they were not philosophers and did not have any Philosophy background.

Choosing a sample of students that were not familiar with matters related to the study was a priority, to guarantee that the data collected was homogeneous and not biased. As my intention was to address folk's intuitions, it was necessary to exclude "expert intuiters" from taking part of the experiment. Once the population was defined, the sample was randomly framed.

Participants were rightfully informed that the data collected was going to be used for research purposes and that it was part of a study to be registered in a PhD dissertation.

4.1.3. Instruments and Procedure

4.1. 3.1. Probes

The instruments used were two pairs of written vignettes focused on the two variables that the study aims to examine: Referential Knowledge of the Audience -RKotA- and Stakes. More precisely, they were designed to test whether the referential knowledge possessed by the addressees of a belief report, or what is at stake for the agent at the

---

[51] Mean age for Knowledge of the Audience applications was 24.59, and for the Stakes one was 23.29. 77 females and 81 males answered the Knowledge of the Audience vignettes, and 82 females and 76 males, the Stakes ones.

time of attribution, affects the disposition of the subjects to attribute a certain belief to an agent. The vignettes were modeled on the basis of similar ones in the field of knowledge attributions –as previously explained-, taking on account the discussion arisen form them and the conditions for experimentation of philosophical thought experiments explained previously.

All vignettes were tested on a pilot group and revised by different experts before they were used to collect the data here presented. Here I offer and English version of the probes. For the Spanish version (the originally applied), see Appendixes 1 and 2. The vignettes are:

A.) *Stakes.*

Scenario 1

Suppose you and two of your friends, Luisa and Andrés, are in a Japanese restaurant. Luisa is an expert on Japanese food and she usually visits that restaurant, reason why you and Andrés are completely disposed to let her order the dishes for everybody and follow her recommendations. Especially Andrés has been very enthusiastic about tasting the food at this restaurant, as he has heard that the food is delicious there.

Before you order, Andrés goes to the toilet for a moment, leaving you alone with Luisa at the table. She starts to look at the menu and says "Oh, here they make a terrific Tuna, we all should order one and taste it". Next, she leaves for a moment. The waitress, who knows Luisa, as she is a regular costumer, wants to take the order knowing that you and Andrés will follow Luisa's advice. Only you are at the table at that moment, so the waitress starts to ask the order of everyone and asks you: Did Luisa believe Andrés should order the Tuna?

Please choose in the following scale from 1 to 6 how correct do you think would be this answer:

"Luisa Believes Andrés should order the Tuna"

	1	2	3	4	5	6	
Totally correct							**Totally incorrect**

86

Scenario 2

Suppose you and two of your friends, Luisa and Andrés, are in a Japanese restaurant. Luisa is an expert on Japanese food and she usually visits that restaurant, reason why you and Andrés are completely disposed to let her order the dishes for everybody and follow her recommendations. Especially Andrés has been very enthusiastic about tasting the food at this restaurant, as he has heard that the food is delicious. Before you order, Andrés goes to the toilet for a moment, leaving you alone with Luisa at the table. She starts to look at the menu and says "Oh, here they make a terrific Tuna, we all should order one and taste it".

You know –but not Luisa- that Andrés is severely allergic to Tuna and that one single bit of it could cause him to die. Luisa leaves the table for a moment. The waitress, who knows Luisa, as she is a regular costumer, wants to take the order knowing that you and Andrés will follow Luisa's advice. There's only you at the table at that moment, so the waitress starts to ask the order of everyone and asks you: Did Luisa believe Andrés should order the Tuna?

Please choose in the following scale from 1 to 6 how correct do you think would be this answer:

"Luisa believes Andrés should order the Tuna"

Totally correct

1	2	3	4	5	6

Totally incorrect

B.) *Knowledge of the audience*

Scenario 1

Suppose you are a great fan of the pop singer publicly known as 'Madonna'. You have always loved her music, have gone to her concerts, seen her movies and read her books. Recently, you have even started to frequent the Kabbalah cult of which she is part. There, no one calls her 'Madonna' because the cult is strict and they have their own baptism

ceremonies and only call people by the name they have been given in the community. The singer, for instance, is only known as 'Esther'.

One day you and your friends are celebrating a party. You have the opportunity to play the CD of your idol, and play a song that you like a lot.

A friend of yours, Jorge, who knows very little about the singer (he only knows she is a pop singer, but he cannot really identify her songs and has never heard that she is also known as 'Esther'), comes to you saying "wow, this singer is incredible!".

A few days after, you are with a group of friends. Everyone knows the pop star and recognizes some of her songs, but they don't know details of her life nor about her identity at the Kabbalah as 'Esther'. You are telling them the story about what happened at the party.

Please choose in the following scale from 1 to 6 how correct you think it would be to say:

"Jorge believes Esther is an incredible singer"

1	2	3	4	5	6

Totally correct   **Totally incorrect**

Scenario 2

Suppose you are a great fan of the pop singer publicly known as 'Madonna'. You have always loved her music, have gone to her concerts, seen her movies and read her books. Recently, you have even started to frequent the Kabbalah cult of which she is part. There, no one calls her 'Madonna' because the cult is strict and they have their own baptism ceremonies and only call people by the name they have been given in the community. The singer, for instance, is only known as 'Esther'.

One day you and your friends are celebrating a party. You have the opportunity to play the CD of your idol, and play a song that you like a lot.

A friend of yours, Jorge, who knows very little about the singer (he only knows she is a pop singer, but he cannot really identify her songs and has never heard that she is also known as 'Esther'), comes to you saying "wow, this singer is incredible!".

A few days after, you are with a group of friends from the Kabbalah cult. They hardly go out of the temple and they know very little about the outside world as they are cloistered devotees. For this reason they only know the pop singer as 'Esther' and most probably they have never heard that she has other names. You are telling them the story about what happened at the party.

Please choose in the following scale from 1 to 6 how correct you think it would be to say:

"Jorge believes Esther is an incredible singer"

Totally   correct | 1 | 2 | 3 | 4 | 5 | 6 | Totally incorrect

4.1.3.2. Procedure

In a classroom setting, the probes were applied to 4 groups of people where the four vignettes were presented either Single or Juxtaposed. Groups one and two were presented the Juxtaposed probes, meaning, they were given both vignettes corresponding to each variable together (either both Scenarios for RKotA or the two Scenarios on Stakes), so they could contrast both Scenarios and establish or correct judgment considering that comparison. Groups 3 and 4 had Single applications where the subjects were given two vignettes that did not correspond to the same variable. Group 3 got randomly any of the following combinations: Scenario 1 of RKotA with Scenario 1 of Stakes, Scenario 1 of RKotA with Scenario 2 of Stakes. And Group 4 got either of these two following pairings: Scenario 2 RKotA with Scenario 1 Stakes, and Scenario 2 RKotA with Scenario 2 Stakes. Order of appearance of the vignettes in both groups was random.

The purpose of setting the administration of the probes this way was, on one hand, to be able to observe whether the pattern of attributions was affected also by the possibility of comparing scenarios in the Juxtaposed probes. I deem it possible that having both vignettes of the pair available for comparison might affect the patterns of response. On the other hand, as regards the Single applications, mixing stakes and RKotA Scenarios arguably helps two important conditions of the experimental design: 1. It allows

comparison between the results obtained with Single and Juxtaposed applications within each variable, with no interference of variables such as tiredness of participants, as for all applications participants got tests of similar length, and general features. 2. As it can be noted, Stakes and RKotA vignettes are very different in topic but similar in structure, presenting them together does not allow for comparison of any type, on the contrary, it prevents biases that would affect Single applications, such as *collaboration with experimenter* or *sabotage* (which happen when subjects find out what the experiment is about and want to collaborate or sabotage the experiment). Each vignette serves as a decoy vignette for the other in pairings of Single applications.

A total of 316 valid probes were applied[52]. Application took an average of 10 minutes, though subjects were instructed to take the time they needed. It was made in Spanish, which is one of the official languages in Catalonia, and the vignettes were designed to fit the regional culture so that they were easy to understand and portrayed familiar situations featuring public figures or imaginary characters with local names. Participants were given the following instructions here translated into English (for the original Spanish instructions, please go to Appendix 3):

- Instructions:

Please read carefully the story told and answer the question at the bottom using the scale just as you are instructed. Please, do it quickly and without reflecting too much. Thank you.

4.1.4. Data collection and analysis

The data were collected with paper and pencil questionnaires in which the vignettes were presented with a Likert scale to measure correctness/incorrectness of attribution in every case, as shown previously. Scoring, therefore, was straightforward on a scale with an even number of choices (and no central mark) in order to push decision. Data were collected during the months of November and December of 2011, and the month of February 2012. The test was applied by the researcher herself under the conditions previously explained.

The data collected are analyzed by means of statistic methods. Specifically, it was calculated the mean, mode and standard deviation of each group. These measurements

---

[52] Tests with unclear or incomplete answers were invalidated.

help us see the patterns that the data follow and make easy to appreciate important points and preferences in the scoring. They give us information about what is the central tendency of a collection of data (mean), which scores are more popular within the sample (mode) and how much dispersion from the average mark exists within the answers. Also, a standard chi squared (χ2) test was applied for all inter-test comparisons, with the aim to evaluate the significance of the difference between the groups. The value given by this test represents the probability of obtaining the observed difference 'by chance', if the null hypothesis h_0 (i.e. that there is no difference between the compared probes) were true. Typically, values of p under 0.05 are considered to show statistically significant difference between the probes compared[53]. Finally, I present pie charts, bar graphs and radial graphs that enable a quick visualization and organization of the data.

4.4.1.Expected results

The general hypothesis that guided the construction of the experiments was that if correctness of belief attributions was affected by the referential knowledge the audience has or by changes on what is at stake, one would expect variations in the distribution of answers across the Likert scale from Scenario 1 to Scenario 2 in each pair of vignettes. More precisely, for the RKotA probes one would expect a tendency towards answers 4, 5, and 6 (scores in the "totally incorrect" zone) for Scenario 1 (the one where the audience does not have knowledge of the proper name employed in the attribution), showing that subjects were more prone to judge the attribution as incorrect in cases where the audience did not have the required referential knowledge for articulating the attribution. To contrast, we should expect a comparative increase in answers 1 (totally correct) for Scenario 2, where the audience has the relevant referential knowledge, supporting the hypothesis that context of utterance of a belief attribution matters.

As regards the Stakes tests, one would expect exactly the opposite behavior with respect to the polarization of answers in each scenario. In Scenario 1 responses should concentrate in scores 1, 2 and 3, while in Scenario 2 they should converge in marks 4, 5, 6. In order to support the claim that in contexts where what is at stake increases, it is

---

[53] Throughout the results here offered I attach to the significance value, in a footnote, the raw value of the test applied for each comparison, in order to provide more precise and complete statistical information. I also present the degrees of freedom The number of degrees of freedom represents the amount of values in the final calculation of the test that are free to vary, and therefore is helpful information to appreciate the accuracy of the test.

less likely to attribute a belief that in low stakes scenarios would be considered accepted.
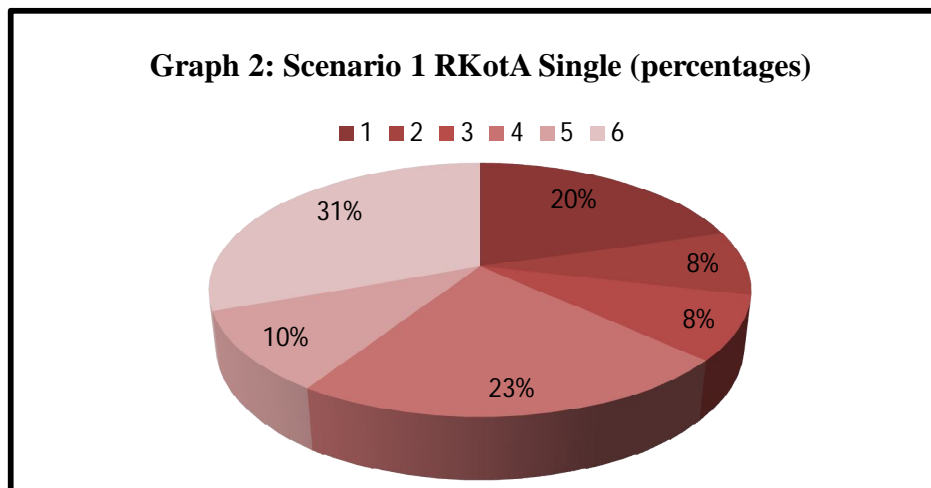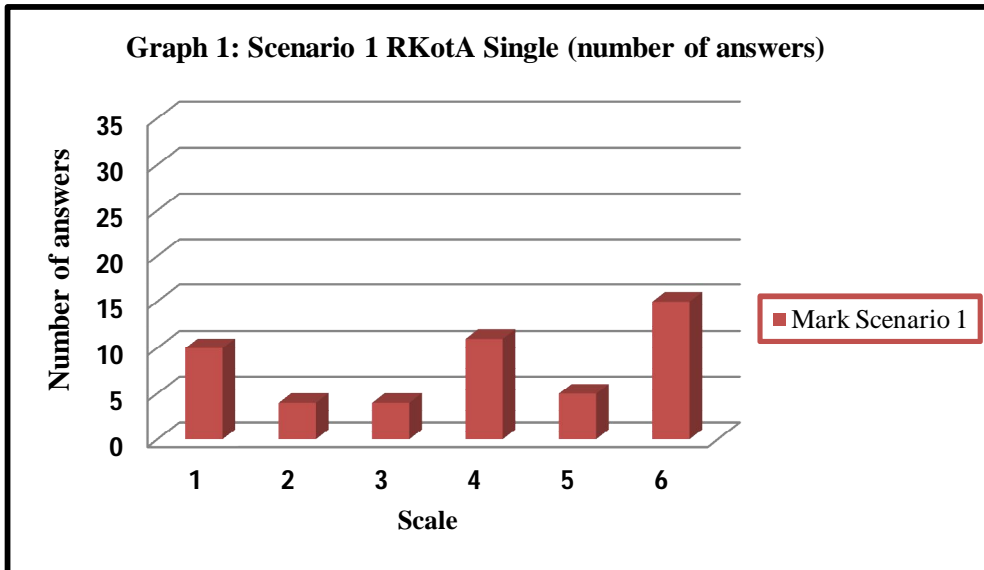
4.2. Results

In this section I present the results obtained after the application of the probes. I will first present those obtained on the Referential Knowledge of the Audience (RKotA) experiment, and after that, I will offer the results of the Stakes administration. Pie charts illustrate the allocation of answers in percentages, and bar graphs show the number of answers in each mark and help us make inter-test comparisons between the groups of Single and Juxtaposed applications; radial graphs allow comparison with a simple view of the distribution of answers with respect to a central point.[54]

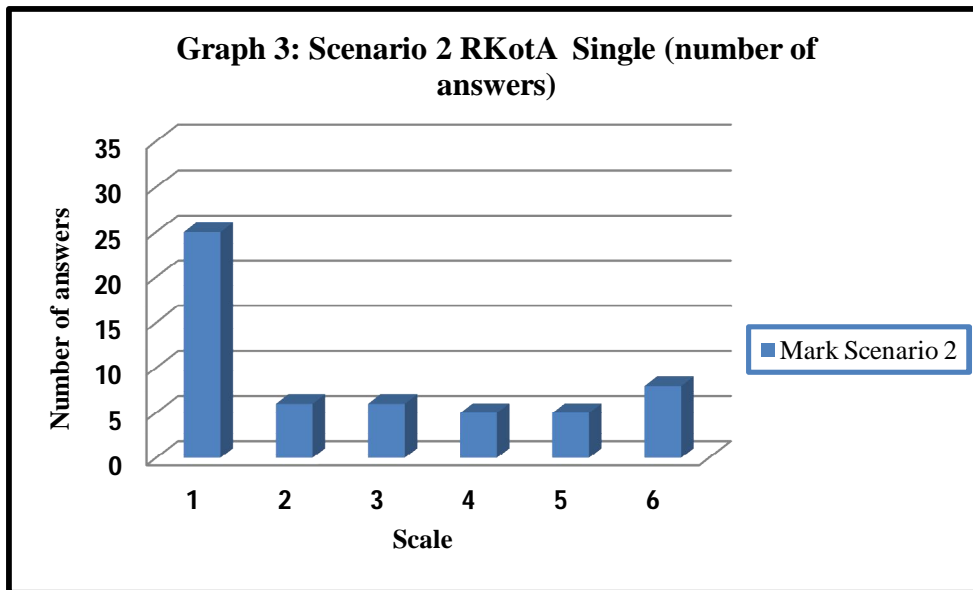### 4.2.1. On the Referential Knowledge of the Audience vignettes

If correctness of belief attributions was affected by Referential Knowledge of the Audience, we would expect a significant change in the distribution of answers from Scenario 1 to Scenario 2. More precisely, one would expect a tendency towards answers 4, 5, and 6 in the scale (scores in the incorrectness area) for Scenario 1 (the one where the audience does not have knowledge of the proper name employed in the attribution), showing that subjects were more prone to judge the attribution as incorrect in cases where the audience did not have the required referential knowledge for articulating the attribution. In contrast, we should expect a comparative increase in answers in the correctness area for Scenario 2, where the audience has the relevant referential knowledge, supporting the hypothesis that context of utterance of a belief attribution matters.

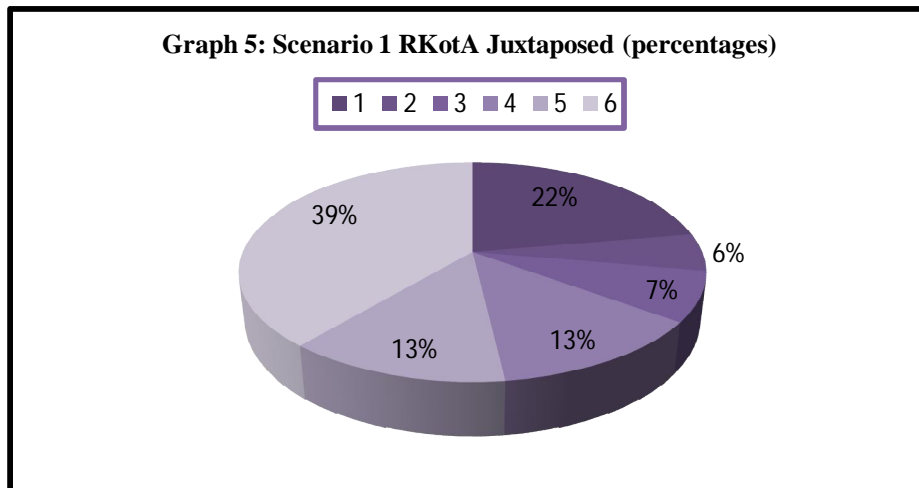In Graphs 1 to 14 I present the results obtained after the administration of the probes.

---

[54] The individual marks of both the RKotA and the Stakes tests can be consulted in Appendix 4

Graph 1: Scenario 1 RKotA Single (number of answers)


Graph 2: Scenario 1 RKotA Single (percentages)

Graphs 1 and 2, show the results of the Single application of Scenario 1. As we can observe, the tallest bar corresponds to mark 6 -'totally incorrect'- in the scale, with 15 (31%) answers. This is followed by marks 4 (11 answers, 23%) and 1 (10 subjects, 20%). The remaining 26% is distributed between scores 2, 3 and 5. Notice that 64% of the answers correspond to judgments of incorrectness, while the 36% remaining are allocated in the correctness area. The mean score for this administration was 3.857, the mode was 6, and the standard deviation 1.90.

Graph 3: Scenario 2 RKotA Single (number of answers)

Graph 4: Scenario 2 RKotA single (percentages)
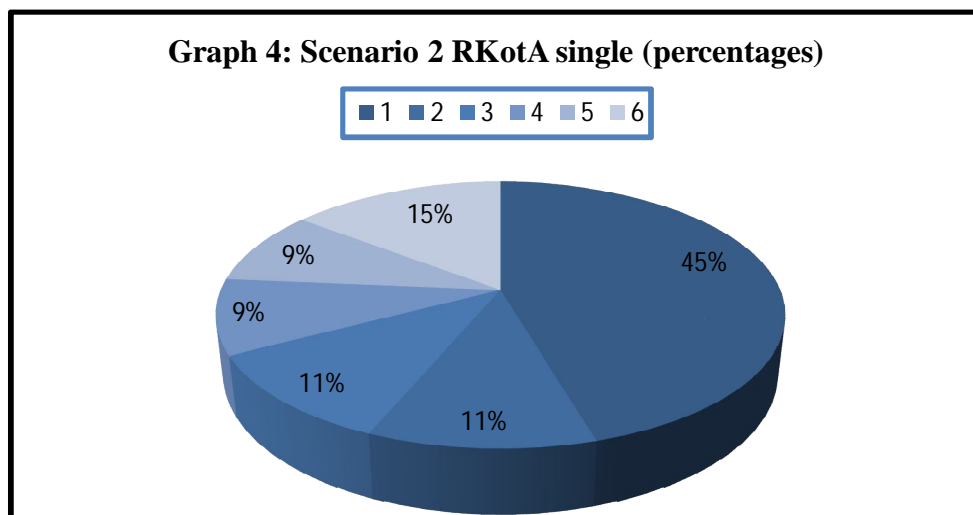
In Graphs 3 and 4 we observe the results collected from the Single application of Scenario 2. We see a clear dominance of score 1 ('totally correct') with 25 of the answers which corresponds to a 45%. Score 6 follows, but only with a 15 % (8 answers). Other marks exhibit percentages between 9% and 11%, no more than 6 participants chose them. Altogether, we can observe a 77% preference for scores in the correctness zone, against a 23% shared by scores 4, 5, and 6. For this application the mean was 2.69, the mode was 1 and the standard deviation 1.91.
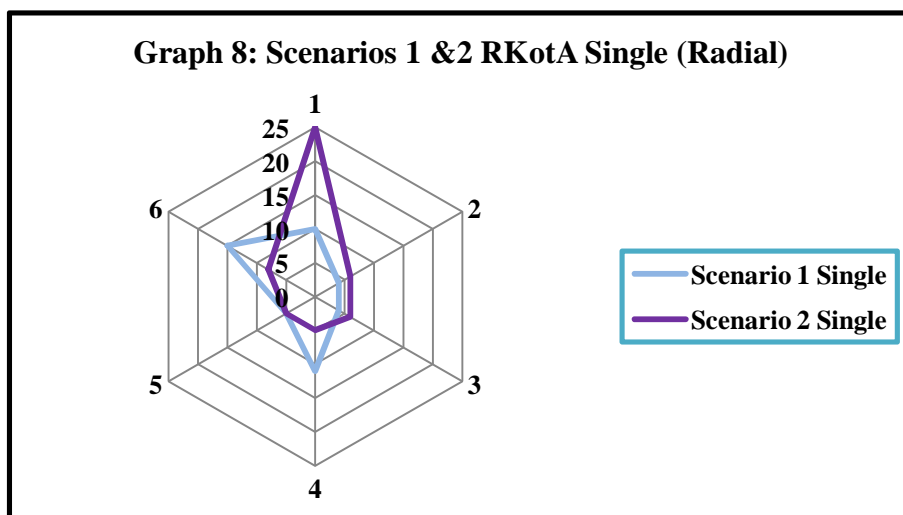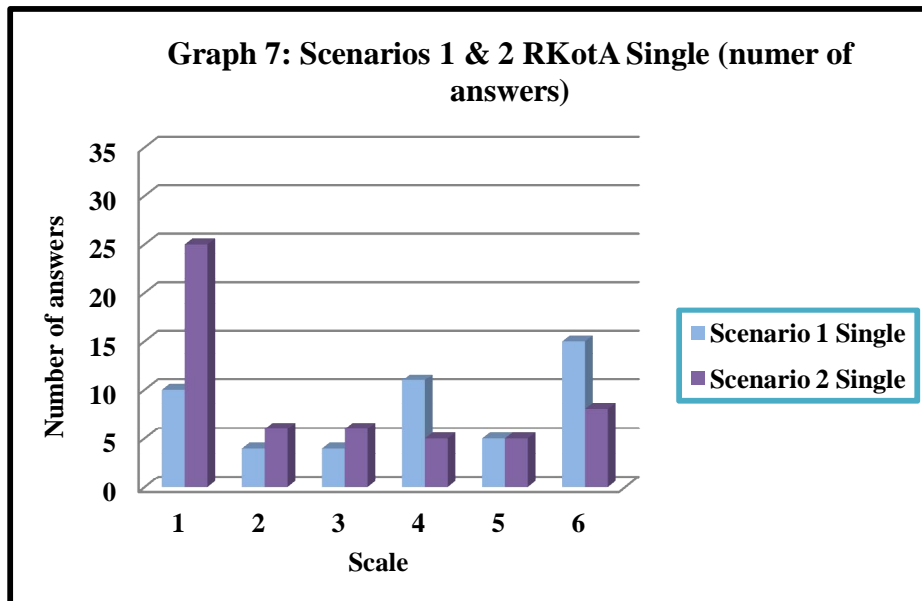
Graph 5: Scenario 1 RKotA Juxtaposed (percentages)

Legend: 1 2 3 4 5 6

22%
6%
7%
13%
13%
39%

In Graph 5 I present the results obtained with Scenario 1 in the Juxtaposed application[55]. In this case the participants had the opportunity to evaluate both vignettes, and therefore could contrast their answers. In the graph for Scenario 1 we see a majority in score 6 (39%), followed by 1 (22%). 35% of the answers correspond to the correctness area, while 65% correspond to the incorrectness judgments (almost doubling the 35% on the incorrectness pole). The mean for this application was 4.05, the mode was 6 and the standard deviation 2.01.

Graph 6: Scenario 2 RKotA Juxtaposed (percentages)

Legend: 1 2 3 4 5 6

0%
5%
7%
6%
17%
65%
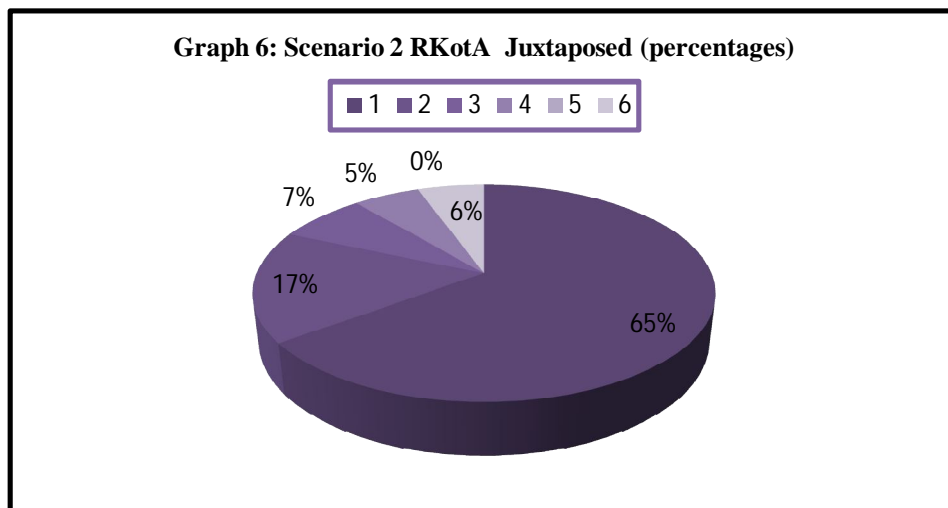
Graph 6 shows the results for Scenario 2 in the Juxtaposed application. Here we see 65% agreement on answer 1. Marks 2 and 3 correspond to 24% of the answers. Therefore, a considerable 89% of the answers concentrate on the correctness area, leaving only an 11% in the opposite area. The mean of this administration was 1.75, the mode was 1 and the standard deviation 1.34.

---

[55] Bar graphs of the juxtaposed application, which correspond to pie charts 5 and 6, can be found below in graph 9.

Graph 7: Scenarios 1 & 2 RKotA Single (numer of answers)

Graph 8: Scenarios 1 &2 RKotA Single (Radial)

Graphs 7 and 8 show a comparison between the scores obtained in Single applications of both Scenario1 and 2 (i.e., comparing the results displayed in Graphs 1 and 2). Light blue corresponds to Scenario 1 and purple corresponds to Scenario 2. In the radial graph it is possible to appreciate the disparity in answers, the substantial dominance of answer one for Scenario 2, which contrasts with the prevalence of marks 4 and 6 for Scenario 1.

In the bar graphs, we see that the largest disparity in scores is found in score 1, where Scenario 2 obtains more than two times the number of answers of Scenario 1. We can see that scores allocated in 4 for Scenario 2 are less than half of those in Scenario 1. Mark 6 –'totally incorrect'-, on the other hand, reaches 15 answers for Scenario 1, while only 8 people chose it when evaluating scenario 2.

Consequently, p= 0.046[56] for Scenarios 1 and 2 in the Single application, showing a significant difference between the groups.

Graph 9: Scenarios 1 &2 RKotA Juxtaposed (number of answers)

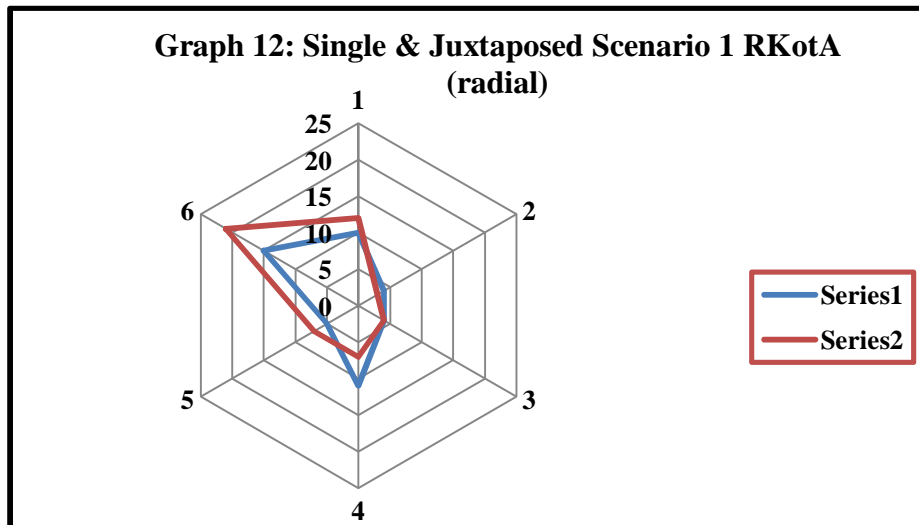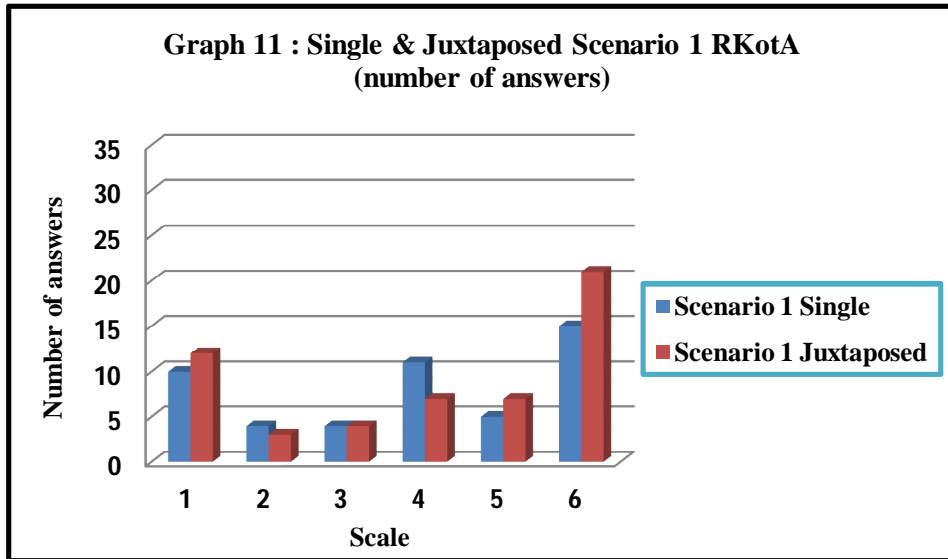Graph 10: Scenarios 1 &2 RKotA Juxtaposed (radial)
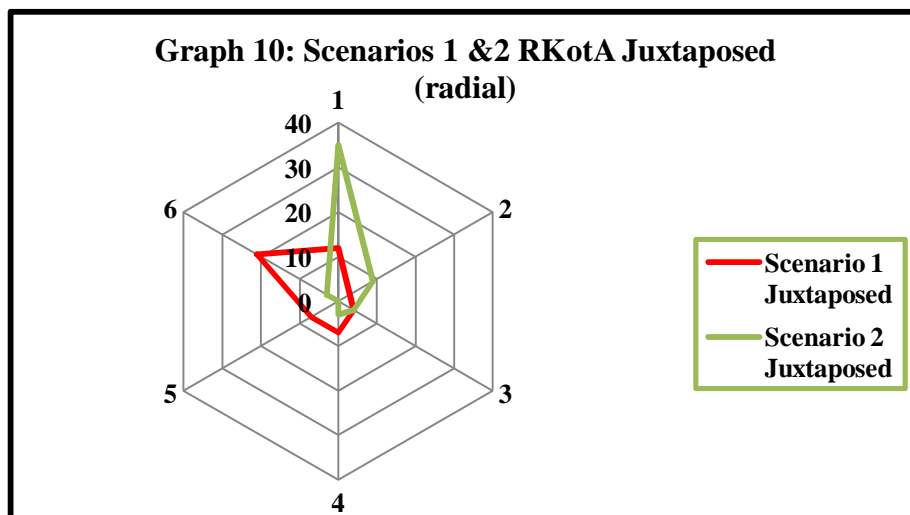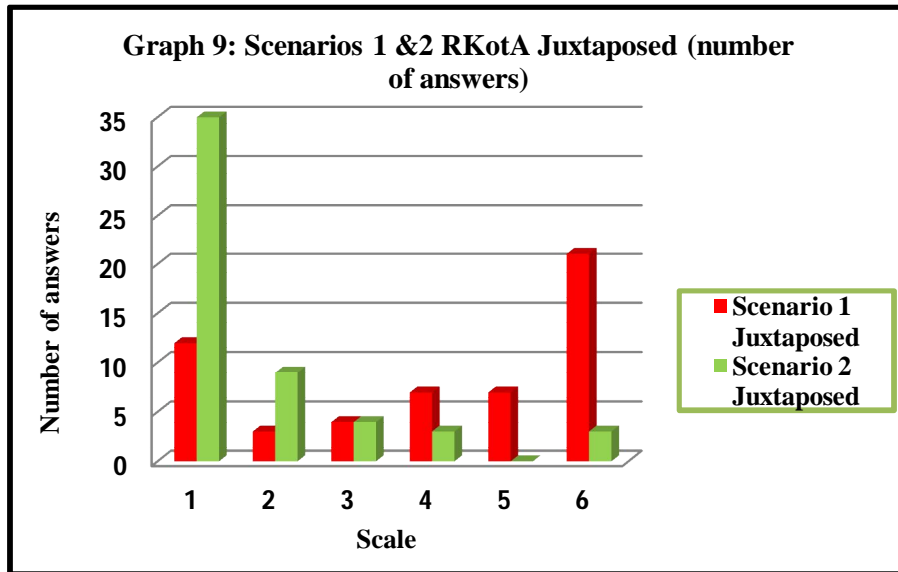
In Graphs 9 and 10 we see the contrast between Scenarios 1 and 2 for Juxtaposed administration (i.e., the comparison between the results shown in Graphs 5 and 6). Scenario 1 is identified in red and Scenario 2 in light green. The disparity seen in polar answers 1 and 6 is considerable, which can by verified in the radial graph. Almost three times the people who thought it was totally correct to make the belief attribution in question in Scenario 2 judged it so in Scenario 1. Conversely, 21 participants considered it totally incorrect to attribute the belief in Scenario 1, but only 3 deemed it so in Scenario 2. In the middle answers (2 to 5), there is a higher number of answers in the

---

[56] χ^2 value= 11.300. Degrees of freedom= 5.

incorrect zone for Scenario 1 and a contrary increase of answers in the correct zone for Scenario 2. Zero participants scored 5 and only one marked 4 in Scenario 2.

For scenarios 1 and 2 in the Juxtaposed application p<.001[57], showing a very significant difference between them.

Graph 11 : Single & Juxtaposed Scenario 1 RKotA (number of answers)

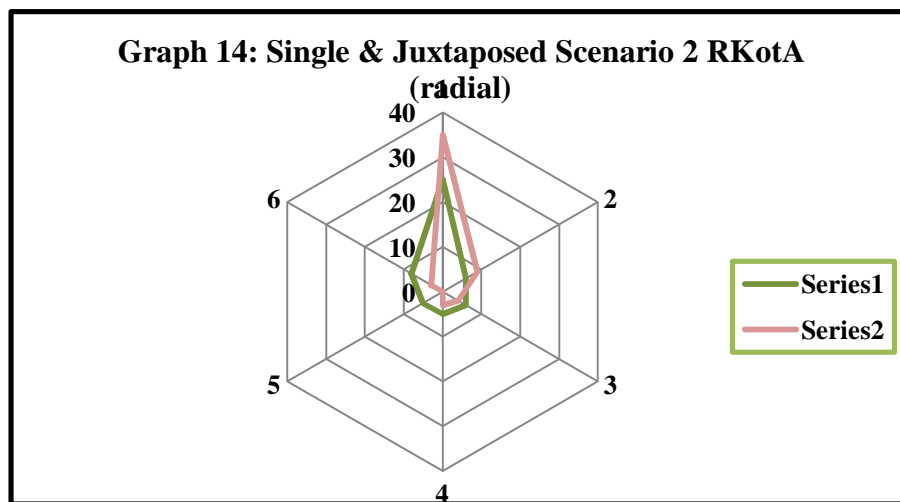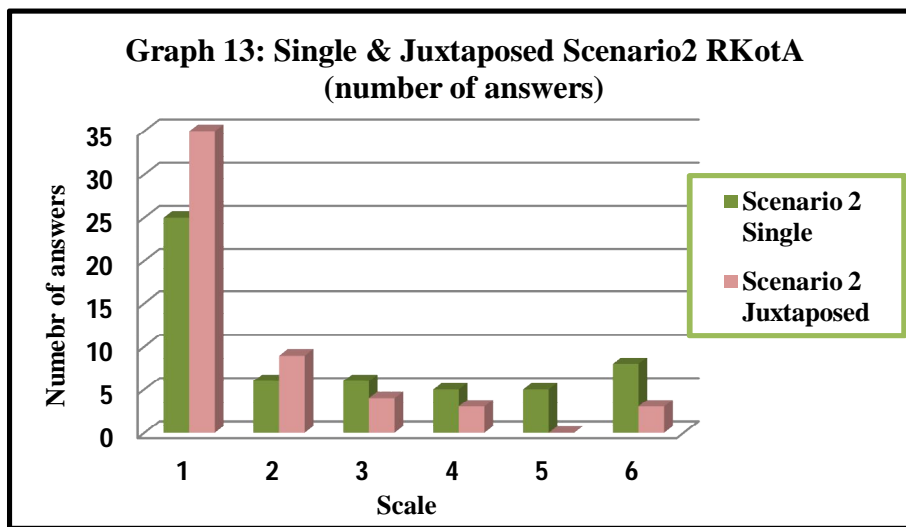**Graph 12: Single & Juxtaposed Scenario 1 RKotA (radial)**

Graphs 11 and 12 offer a comparison between the results obtained by Single and Juxtaposed applications of Scenario 1 (which can be traced back to Graphs 1, 2, 5 and partially 9). The blue colour represents the Single application while red bars and lines stand for the Juxtaposed test. In the radial graph we see a fairly matching dynamic in answers for both applications, with only mark 4 standing out for the Single test. It is possible to observe some parity in the answers to polar score 1, with a difference of just

---

[57] χ^2 value= 36.355. Degrees of freedom= 5.

a few points between Juxtaposed and Single administrations. Conversely, we can see an increase of responses for the Juxtaposed case in the other polar answer, 6. Additionally, in the central scores of the scale we can appreciate uniformity, except for mark 4, where there is variation to the Juxtaposed case, in which 7 people chose 4, and the Single vignette, where 11 were inclined towards this answer.

For the comparison between single and juxtaposed applications of scenario 1 p= 0.805[58], showing no significant difference between them.

Graph 13: Single & Juxtaposed Scenario2 RKotA (number of answers)

Graph 14: Single & Juxtaposed Scenario 2 RKotA (radial)

Finally, Graphs 13 and 14 compare Scenario 2 in Single and Juxtaposed applications (results displayed in Graphs 3, 4, 6 and 9). Single application can be identified in green and Juxtaposed application in pink. A parity in most of the scores can be noticed, with a dominance of score 1 for both applications. We can also see how they have matching

---

[58] χ^2 value= 2.310. Degrees of freedom= 5.

shapes in the diagrams of answers in the radial graph. However, the bar graph reveals some differences in marks 1, 5 and 6. The number of answers 1 increases, while the number of responses 6 decreases in Juxtaposed applications of Scenario 2, compared to Single ones, and 0 people chose mark 5 when presented the Juxtaposed vignettes.
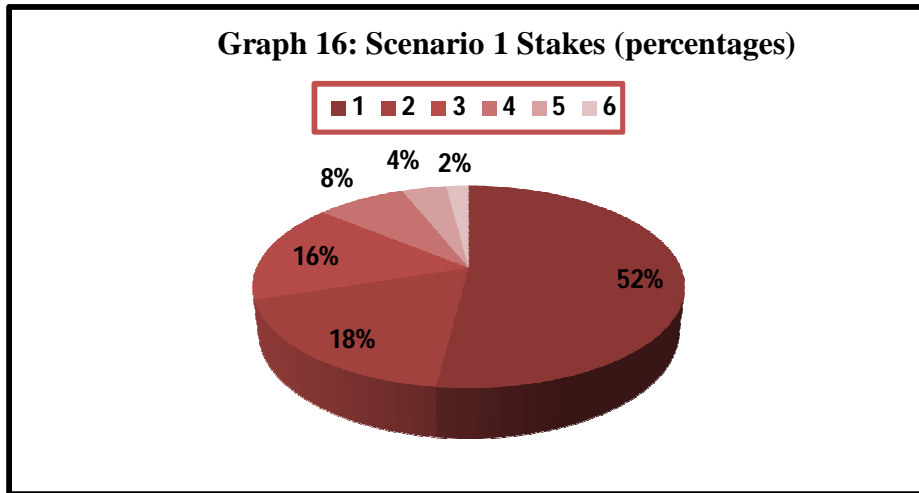
The difference between Single and Juxtaposed applications of scenario 2 is p= 0.064[59]; it barely has significance at the conventional p=0.05 level. Nevertheless, given that the difference between the two groups of results is important, one might expect that with more statistic potency (a bigger sample) it would reach greater significance.

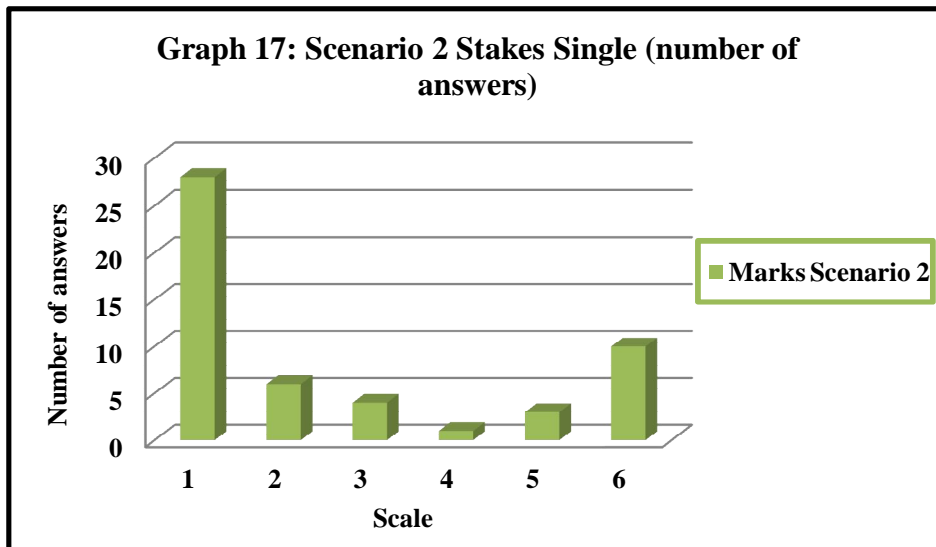*4.2.2. On the Stakes vignettes*

If correctness of belief attributions was influenced by stakes in the agent's context, just as with the previous variable, we would expect a considerable change in the distribution of answers from Scenario 1 to Scenario 2. More precisely, one would expect a tendency towards answers 1, 2, 3 in the scale (scores in the correctness area) for Scenario 1(in which stakes of the agent are low), showing that subjects were more likely to judge the attribution as correct in cases where the attribution did not present any risk for the agent. In contrast, we should expect a comparative increase in answers in the incorrectness area for Scenario 2, where there is an important price for being wrong with the attribution. If these patterns show in the distribution of answers, this would corroborate the hypothesis that the stakes of the context of the agent matter for belief attribution. In Graphs 15 to 28 I present the results obtained after the administration of the probes.
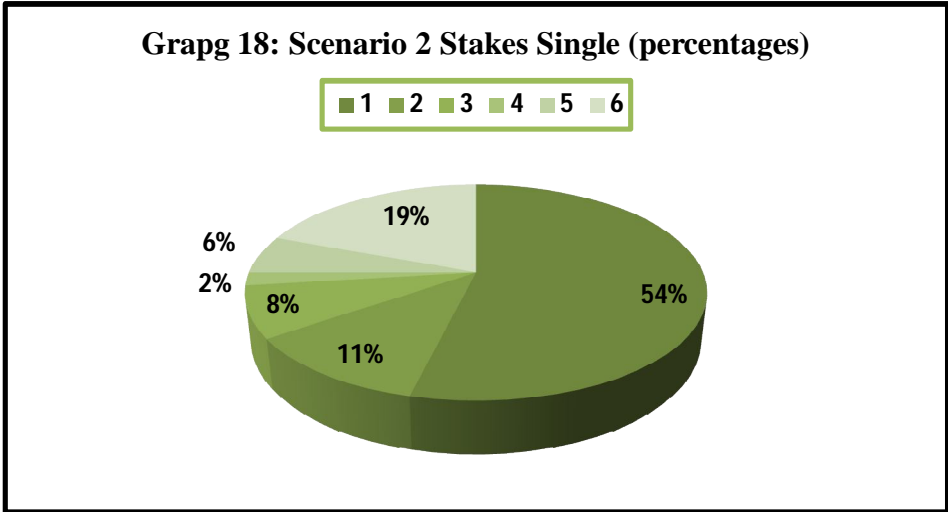

Graph 15: Scenario 1 Stakes Single (number of answers)

---
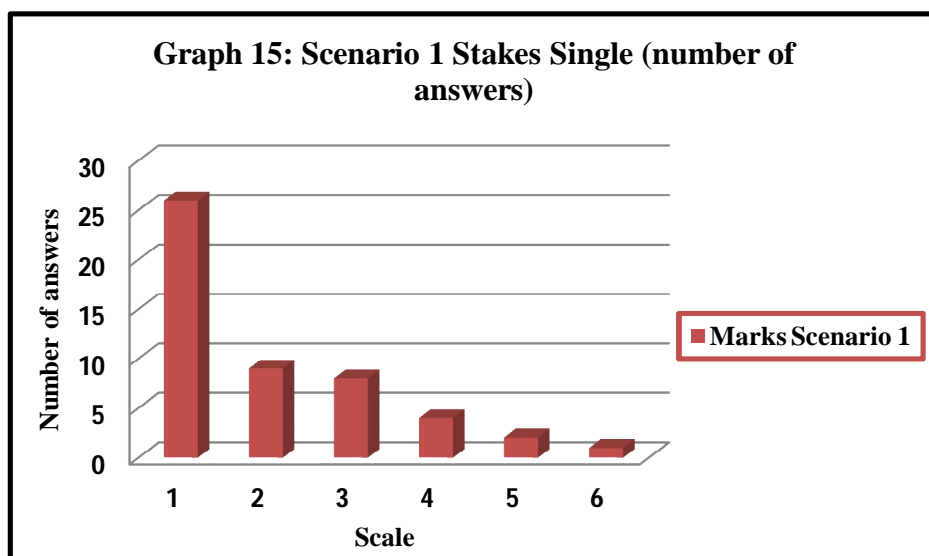
[59] χ^2 value= 10.431. Degrees of freedom= 5.

Graph 16: Scenario 1 Stakes (percentages)

■ 1 ■ 2 ■ 3 ■ 4 ■ 5 ■ 6

8% 4% 2%

16%

52%

18%

Graphs 15 and 16, show the results of the Single application of Scenario 1of the Stakes probe, where we can appreciate dramatic preference of answer 1 -'totally correct'-, which 26 subjects chose, the equivalent to 52% answers. Marks 2 (9 answers, 18%) and 3 (8 subjects, 16%) are the most selected after 1. We observe that the number of answers decline gradually to mark 6, only selected by one person (2%). The difference between percentages in the correctness and incorrectness area is striking: 86% of the answers correspond to judgments of incorrectness, while only 14% pertain to the correctness area. The mean score for this administration was 2, the mode was 1 and the standard deviation was 1.30.

Graph 17: Scenario 2 Stakes Single (number of answers)

■ Marks Scenario 2

Number of answers

Scale

**Grapg 18: Scenario 2 Stakes Single (percentages)**

In Graphs 17 and 18 we observe the outcome of the Single application of Scenario 2 of the Stakes vignette. Here we also observe a substantial dominance of score 1 -'totally correct'- with 28 answers, which corresponds to 54%. Score 6 follows, but the difference is quite dramatic, it only obtained 19% (8 answers). Mark 2 rose to 11%, selected by only 6 people. On the whole, 73% of the participants preferred scores in the correctness zone, against a 27% equivalent to marks 4, 5, and 6. For this application the mean was 2.51, the mode was 1 and the standard deviation was 2.02.

Graph 19: Scenario 1 Stakes Juxtaposed (percentages)
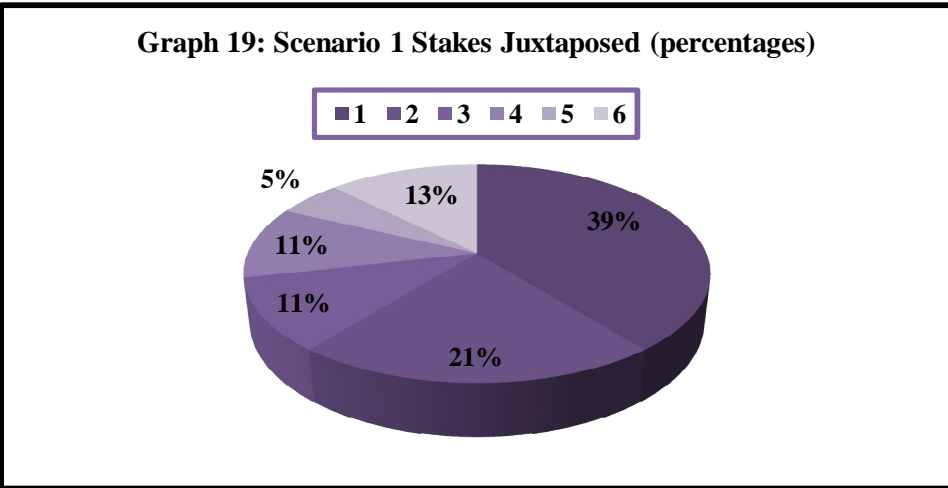
In Graph 19[60] I present the results obtained with Scenario 1 in the Juxtaposed application. In this case, as previously explained, the participants were given the chance to evaluate both Scenario 1 and Scenario 2 vignettes; they could compare and rectify their answers if they wanted. In the chart, we see a majority in score 1 (39%), followed by 2 (21%), and 1 (135). It seems remarkable that 71% of the answers correspond to the

---

[60] Bar graphs of the juxtaposed application, which correspond to pie charts 19 and 20, can be found below in graph 23.

correctness judgments area, number than surpasses by more than double the remaining 29% that correspond to the incorrectness area. The mean for this application was 2.58, the mode was 1 and the standard deviation was 1.76.
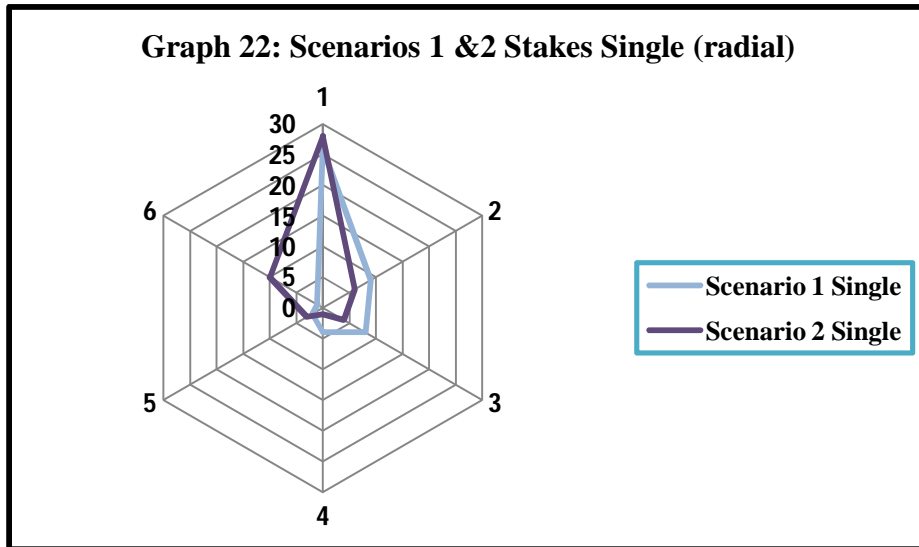
Graph 20: Scenario 2 Stakes Juxtaposed (percentages)

Legend: 1 2 3 4 5 6

- 34%
- 5%
- 6%
- 7%
- 16%
- 32%
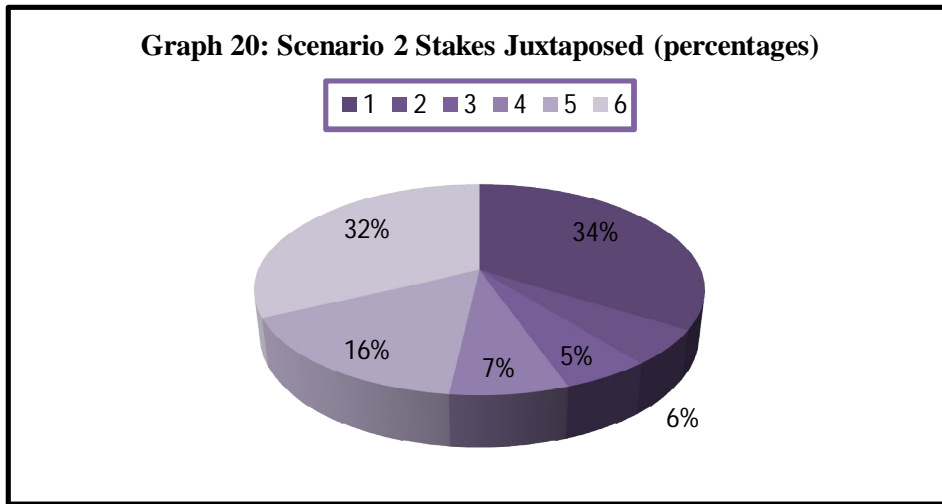
Graph 20 shows the results for Scenario 2 in the Juxtaposed application. Again, we see majority of choices on score 1, represented by 34% (19 answers), followed very closely by mark 6, with 32%. Surprisingly, the answers are fairly equally distributed between the incorrectness and correctness areas: 55% in the incorrectness share, against 45% of participants that judged the attribution correct to a certain degree. The mean for this application is 3.625, the mode was 1 and the standard deviation was 2.17.
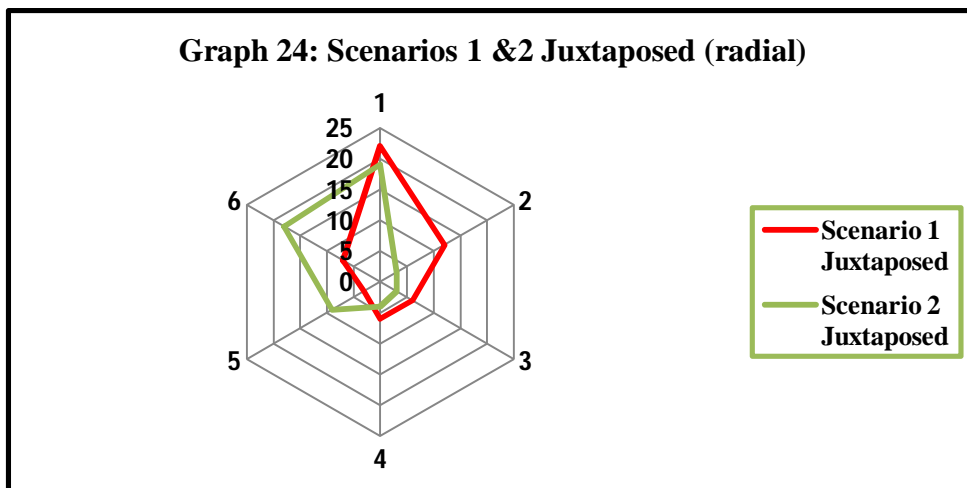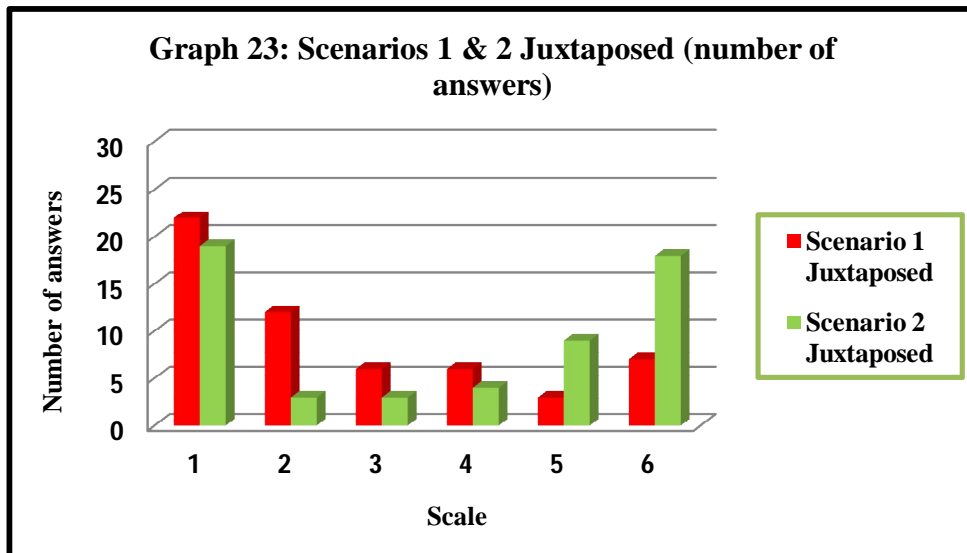
Graph 21: Scenarios 1 &2 Stakes Single (number of answers)

Y-axis: Number of answers (0, 5, 10, 15, 20, 25, 30)
X-axis: Scale (1, 2, 3, 4, 5, 6)

Legend:
- Scenario 1 Single
- Scenario 2 Single

Graph 22: Scenarios 1 &2 Stakes Single (radial)
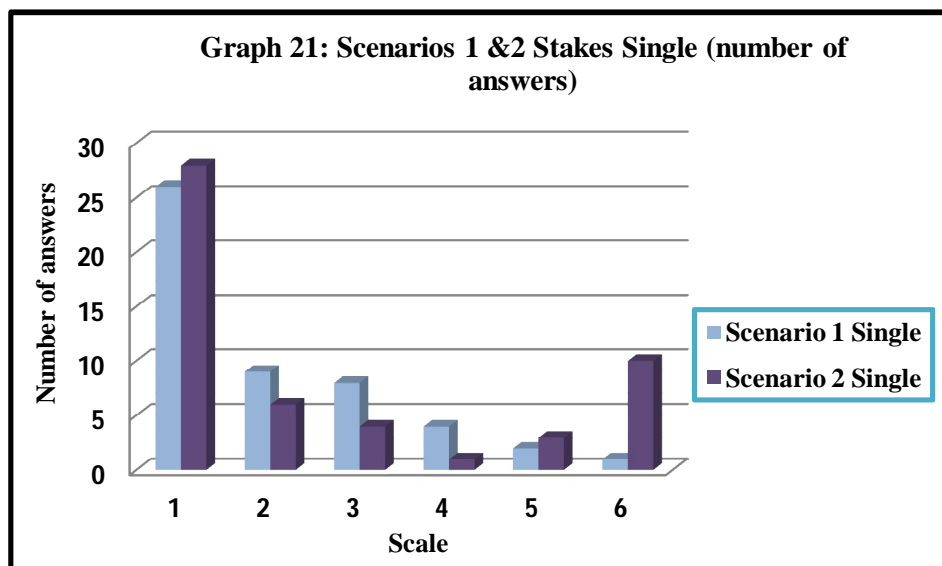
Graphs 21 and 22 contrast the scores obtained in Single applications of both Scenario1 and 2 (consequently, it compares the results displayed in Graphs 15 and 16). Light blue corresponds to Scenario 1 and purple corresponds to Scenario 2. In the radial graph we can see how the pattern of answers to both scenarios forms sharp shapes pointing towards mark one. The marked preponderance of answer one in both scenarios is noteworthy.

In the bar graph, it is possible to corroborate the parity and dominance of answer one in both contexts, but it is also evident that for the single Scenario 2 the number of answers allocated in mark 6 –'totally incorrect'-is 10 times the number we find in Scenario 1. As for the rest of the scores, the only important difference we find between scenarios is in mark 3, where Scenario 1 doubles the number of answers of Scenario 2.

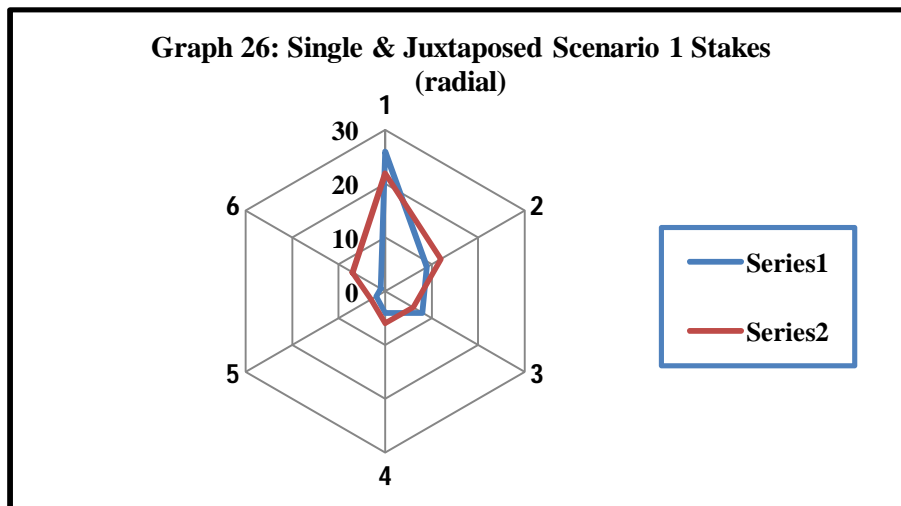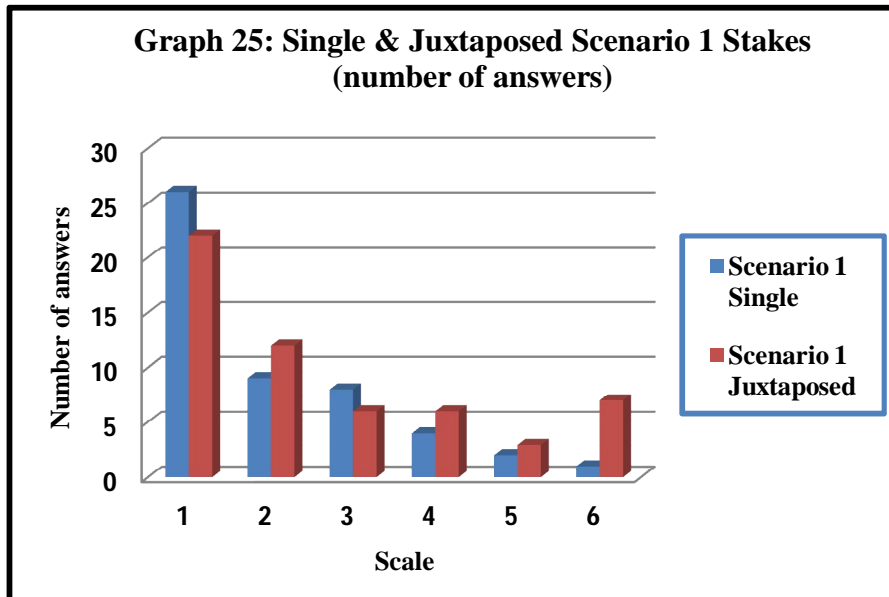For Scenarios 1 and 2 in the Single application p= 0.50[61], showing no significant difference between the groups.

---

[61] χ^2 value= 10.776. Degrees of freedom= 5.

Graph 23: Scenarios 1 & 2 Juxtaposed (number of answers)

Graph 24: Scenarios 1 &2 Juxtaposed (radial)

In Graphs 23 and 24 we find a comparison between Scenarios 1 and 2 for the Juxtaposed administration (therefore, this integrates the results in Graphs 19 and 20). Scenario 1 is in red and Scenario 2 is represented in light green. Even though for both vignettes the dominant answer is 1-totally correct-, it is noteworthy that the general distribution of answers is quite dissimilar in the groups. As we can appreciate in the radial graph, the shapes that represent the answer pattern for both probes are different in that they expand towards different sides of the hexagon, showing the prevalence of correctness judgments for Scenario 1, and incorrectness for Scenario 2. This can be clearly observed if we look at marks 6, 5 and 2 in the bar graph. For mark 2, in Scenario 1 there is four times the number of answers than in Scenario 2. On the contrary, mark 5 was selected three times more for people presented with Scenario 2 than with Scenario 1, and answers on mark 6 for the latter probe are more than double compared to Scenario 2. It is also quite remarkable that while we see a gradable decrease in number

of answers from 1 to 6 in Scenario 1, in Scenario 2 we observe that both poles have barely the same number of answers, and that in all central marks (2, 3, 4, 5), we notice a progressive increase in number of answers towards the incorrectness area.

Generally, the pattern of answers for both vignettes is quite dissimilar, despite their overlap in score 1. Consequently, for scenarios 1 and 2 in the Juxtaposed application p=0.010[62], shows a marked difference between them.

Graph 25: Single & Juxtaposed Scenario 1 Stakes (number of answers)

Graph 26: Single & Juxtaposed Scenario 1 Stakes (radial)
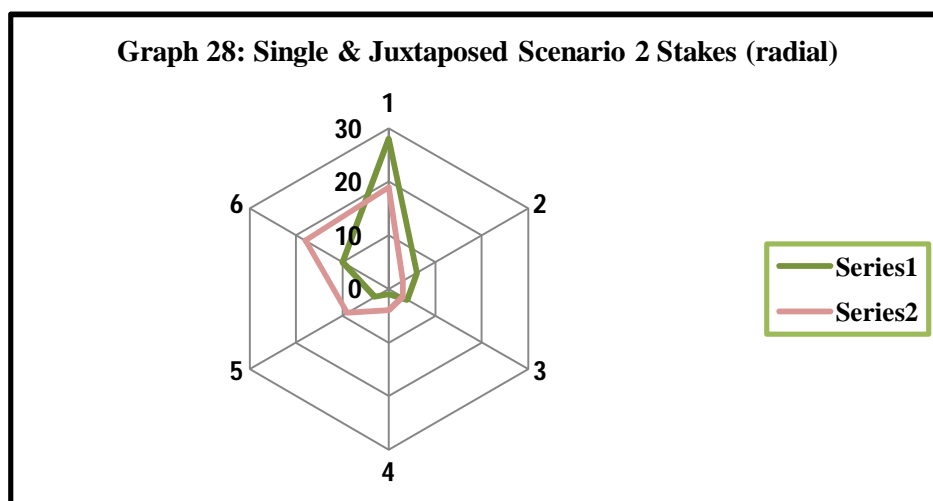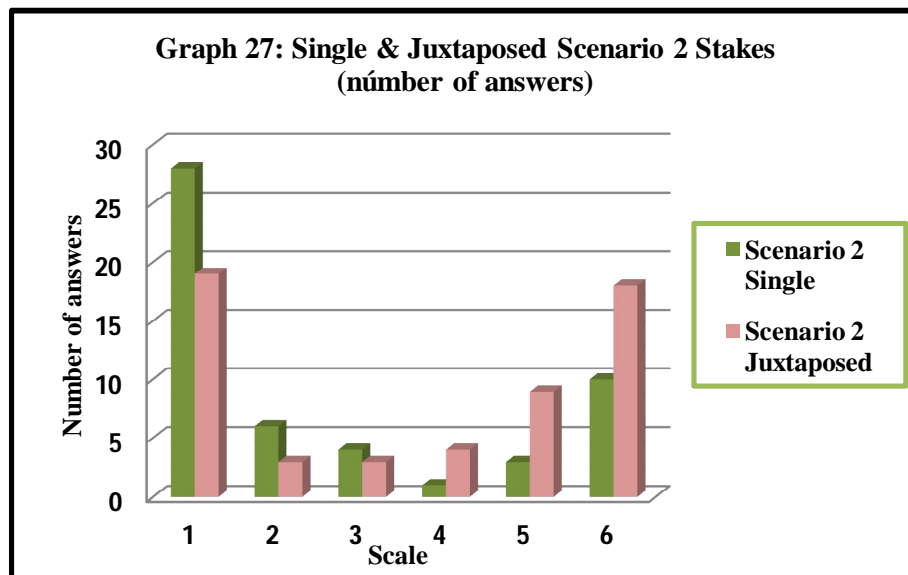
Graphs 25 and 26 offer a comparison between the results obtained by Single and Juxtaposed applications of Scenario 1 (found before in Graphs 15, 16, 19 and partially 23). Single application is in blue while red stands for the Juxtaposed test. In the radial graph we see how both series of answers portray 1 as the dominant choice, and the

---

[62] χ^2 value= 14.860. Degrees of freedom= 5.

patterns in other scores are also similar, except that Scenario two shows a larger amount of answers in marks 2 and 6. If we consult the bar graph we corroborate the similarity in the answers to polar score 1, with a difference of just a few points between Juxtaposed and Single administrations. Furthermore, it can be noted that for the Single application there is a gradual decline in the number of answers from 1 to 6, while the pattern in the Juxtaposed application is less organized and surprises us with a steep rise in the number of answers for score 6. For this reason, while in the Single application for Scenario 1 only one person chose this score, 7 people did so when presented with the Juxtaposed case.

The comparison between Single and Juxtaposed applications of scenario 1 p= 0.372[63], shows no significant difference between the groups.

Graph 27: Single & Juxtaposed Scenario 2 Stakes (númber of answers)

Graph 28: Single & Juxtaposed Scenario 2 Stakes (radial)

---

[63] χ^2 value= 5.497. Degrees of freedom= 5.

Lastly, Graphs 27 and 28 contrast the results obtained for Scenario 2 in Single –green bars and lines- and Juxtaposed –in pink color- applications (it compares the results in Graphs 16, 17, 19 and 22). The radial graph shows a dominance of answer one for the Single application, which stands out against the parity of scores 1 and 6 for the Juxtaposed vignette. In the bar graph we can see how mark 1 in the Single application obtained 9 points over the Juxtaposed one. This is the largest difference present between both applications in all scores, followed by marks 5 and 6. As regards mark 5, the Juxtaposed probe obtained three times more selections than the Single one. Similarly, we observe that in juxtaposition almost the double of participants judged the attribution to be totally incorrect (mark 6).

The difference between Single and Juxtaposed applications of scenario 2 is p= 0.077[64], and therefore, is not significant.

Obviously, it is to be generally expected that larger samples, of at least 100 participants per application, would provide more solid evidence. Nevertheless, I consider the results obtained allow for some interesting preliminary conclusions.

---

[64] χ^2 value= 9.817. Degrees of freedom= 5

Chapter Five

Discussion, Conclusions and Future Work

In this section I will analyse the data collected under the light of the theories presented along this dissertation. The reader will first find the discussion on the tests on Referential Knowledge of the Audience –RKotA-, and secondly, the corresponding analysis of the results of the Stakes application. Finally, I present a general conclusion

5.1. About the Referential Knowledge of the Audience variable

5.1.1. General observations on the RKotA test

In general terms, the results of the RKotA test show that the subjects have a tendency to judge the attribution of belief as incorrect in Scenario 1 and as correct in Scenario 2. This arguably supports the hypothesis that the attribution of beliefs is sensitive to the Referential Knowledge of the Audience parameter in the context of utterance of a belief report.

This is especially clear in Graph 7 and 8, which compare the results obtained in Single applications of the two Scenarios, and in Graphs 9 and 10, that compare the results obtained for Scenario 1 vs. Scenario 2 in the Juxtaposed application. We can observe there is a remarkable difference in both Scenarios as regards the amount of people that think that 'Jorge believes Esther is an incredible singer' is totally correct (mark 1). Both in Single and Juxtaposed applications, the proportion of people that think that the attribution is totally correct is more than double in Scenario 2 than in Scenario 1. As regards mark 6, i.e., the subjects who think that the attribution is incorrect, we have the inverse result: the proportion of those who regard the report as incorrect is much higher in Scenario 1 than in Scenario 2, confirming that people are definitely more inclined to regard the attribution of the belief as correct or incorrect depending on the Referential Knowledge of the Audience.

It is interesting to note that the proportion of those who judge the belief report incorrect is reduced quite dramatically from Scenario1 to Scenario 2 in the Juxtaposed

application in comparison with the Single application (Graphs 5 and 6 confirm a reduction from 65% to 11% in marks 4, 5 and 6 for the Juxtaposed administration, contrasted with a decrease from 64% to 33% -hardly more than a half- in marks 4, 5, 6 for the single administration, shown in Graphs 2 and 4).

It appears that participants that can contrast both Scenarios (that is when it is salient that the difference between the two Scenarios has to do with the audience's familiarity with the name 'Esther' as a name for Madonna) are even less disposed to regard the attribution of belief in Scenario 2 as incorrect. Bear in mind that the relevant difference between Scenarios 1 and 2 is just that the addressees of the report in Scenario 1 are not familiar with the name whereas those in Scenario 2 are. Let me stress again that Jorge's state of mind is held fixed. Nothing in the vignettes would suggest that there is a difference in Jorge's mental representations, ideas, notions or dispositions between Scenario 1 and Scenario 2.

The Referential Knowledge of the Audience, I conclude, is a relevant parameter to BRs. This is supported by the significant difference between Scenarios 1 and 2 in both Single and Juxtaposed applications. But also, it seems that when it is more obvious that there are differences in the referential knowledge of the addressees, namely when participants can compare the two Scenarios in the Juxtaposed application, we observe important variations in their judgments. Given that the only difference between the two Scenarios is the Referential Knowledge of the Audience -their familiarity or lack of familiarity with 'Esther'- it is possible that the difference in results between the Single and Juxtaposed applications that we observe in Graphs 11, 12, 13 and 14 is due to the fact that the comparison makes this parameter, Referential Knowledge of the Audience, more salient and it is this that affects the judgments. Observe that for Scenario 2 (Graphs 13 an 15), even though the dominant answer in both Single and Juxtaposed applications is 1 ('totally correct'), in the Juxtaposed application we see a substantial increase in the number of subjects that choose 1 -comparing Graphs 4 and 6 we see that the increase is 20%. And in general the percentage of answers in the incorrectness area (marks 4 to 6) goes down from 33% in the Single application to just 11% in the Juxtaposed

Prima facie, one might have expected a similar change, but towards incorrectness, in Scenario 1, when comparing Single to Juxtaposed applications. However, we do not

observe that. Although there is an increase in the 'totally incorrect' (mark 6) answers in the Juxtaposed application, it is not as dramatic as the increase in 'totally correct' answers that we observe in Graph 13 and 14 (corresponding to Scenario 2), for the increase here is just an 8% (compared to an increase of 20% for mark 1 as we saw in the case of Scenario 2). Thus, it appears that comparing the two vignettes does not result in more people being convinced that the belief attribution is incorrect. In fact, the proportion of answers in the incorrectness zone (marks 4 to 6) is practically invariant between the Single and the Juxtaposed application.

Nevertheless, it is noticeable that there is more confidence in their judgments of incorrectness as regards Scenario 1, among the participants that were simultaneously exposed to the two Scenarios. Mark 4, even if it is in the incorrectness area, is a relatively neutral response, the kind of response that one would be likely to choose when not completely confident in one's judgment. When we compare the mark 4 response in the Single and the Juxtaposed application of Scenario 1, we can notice a tendency to abandon the more neutral response in favour of the more definitely incorrect ones in the Juxtaposed application.

We see that people are slightly more confident in their judgment that the report was incorrect in the Juxtaposed than in the Single application of Scenario 1. But the exposure to the two vignettes, namely the exposure to a story where the addressees of the report are familiar with the name 'Esther' did not contribute to make the participants more prone to make choices within the incorrectness area: across Single and Juxtaposed applications of Scenario 1 the distribution of percentages in the two correctness and incorrectness zones remains rather stable.

Contrary to my initial assumption, it might not be so surprising that the exposure to the two vignettes would not result in more people judging the report in Scenario 1 as incorrect. The participants who did judge the report as incorrect in the Single application of Scenario 1 did so, I hypothesize, because the reporter used a name that was unfamiliar to the addressees. Exposing the participants in the Juxtaposed application also to a story where the audience is familiar with the name in question does not alter the reason to regard the report as incorrect in Scenario 1, for it does not change the fact that the addressees of the report in that Scenario are not familiar with 'Esther.'

It appears that the comparison with Scenario 2 does not make the report look any more incorrect than it looks when evaluated independently.

Just to summarize, the proportion of Incorrect/Correct answers in Scenario 1 remains invariant across Single and Juxtaposed applications, but the responses of the participants in the Juxtaposed application are more confident. On the other hand there is more of a tendency to judge the report as correct in Scenario 2 among participants that can compare the two vignettes.

5.1.2. Specific observations on the RKotA test

Focusing now on the non-mainstream responses to each of the Scenarios, first we can observe that as regards Scenario 1 there is a percentage of participants that consider the attribution 'Jorge believes Esther is an incredible singer' as correct in spite of the fact that the addressees of the report are clearly not familiar with the name 'Esther' and nothing in the vignette suggests that Jorge would express his belief using that name, and that remains constant across Single and Juxtaposed applications.

And as regards Scenario 2, we still see a percentage of people that regard the attribution as incorrect, even when it is directed to an audience that is familiar with the name 'Esther.'

A possible explanation for the non-mainstream responses is, quite simply, that part of the population does not have contextualist leanings, at least as regards the parameter tested. In Scenario 1, the 36% of participants that give a response in the correctness zone in the Single application (20% of them ready to regard the report as totally correct), and the 35% that give a response in the correctness zone in the Juxtaposed application (of which 22% regard the report as totally correct) perhaps are simply focusing on the state of affairs that is being attributed to Jorge, regardless of the wording the audience is familiar with or the believer would he himself use. And in Scenario 2, there seems to be an opposite tendency by part of the population, i.e., the tendency to consider paramount the fact that the believer would not use the word 'Esther', so one could say that part of the population, perhaps, are sensitive to the believer's mode of presentation, something that they express by classifying the report as incorrect.

However, in Scenario 2 we do observe a moderate reduction in the incorrectness zone from the Single to the Juxtaposed applications. Participants exposed to both vignettes are less prone to regard the attribution as incorrect (33% regard it as incorrect in the Single application vs. 11% who do so in the Juxtaposed) as if the exposure to a vignette that makes the attribution seem more incorrect (as in Scenario 1) makes the difference. But it seems to me that if the participants that regarded the report as incorrect, either in the Single or the Juxtaposed application, had done so because it was important to them that Jorge would not use the word 'Esther,' we should not observe such a decrease, for it is clear that the believer's mode of presentation is invariant across the two Scenarios. It is as if concern for the audience trumped the way in which the believer would express his belief.

5.1.3. Possible worries on the RKotA vignettes

I would like to express some possible worries about the RKotA vignettes, which would suggest that the probes might not be testing just what I wanted to test (i.e. whether correctness of belief attribution is affected by this contextual parameter).

A major worry would be that the variation in responses observed between Scenarios 1 and 2 might not be due to matters about attribution of belief, but is caused by other factors. It could be said that the reason why people would consider incorrect the attribution in Scenario 1 (where the name is unfamiliar to the audience), responds simply to the fact that the attribution judged employs a name that the audience would not understand or recognize. This can be perceived as a violation of the rules of communication: by using the name 'Esther' with a hearer that does not recognize it as a name for Madonna, I am doing something similar to speaking to him in a language alien to him. If this is the case, the data gathered does not show a phenomenon about contextuality of belief attributions, but is simply due to the mere fact that understanding between users requires a proper introduction of new words, a common language. I'll call this possible problem the *foreign language difficulty*.

This is a very important issue for testing whether Referential Knowledge of the Audience might be a variable that affects BRs, precisely because the variable to measure is to do with recognition of the name. It seems to me that it is extremely difficult to design an experiment in which this variable isolated and controlled but does not have the foreign language difficulty. Evaluation of Referential Knowledge of the

Audience requires, as it is obvious, the contrast between identical situations where the only variation is precisely the familiarity with a certain proper name. One could suggest that by making experiments that are parallel to the ones here presented but test the variable (i.e. familiarity against unfamiliarity with a name) also in situations that are not about belief attribution, one could appreciate by comparison whether the results gathered here do indeed reflect something about belief reporting or are due to the foreign language difficulty (or other factors). For instance, one could modify the RKotA vignettes to portray a situation in which Jorge *says* that Esther is an incredible singer and where the report to evaluate is about what he said. The idea would be that if the vignettes are parallel but in one we do not intend to measure BRs but other attitudes -or as in the example, indirect speech-, and we get the same or very similar data, then it can be said that the results do not reveal facts about our belief attributing practices, but other issues.

However, this can't be a solution to the matter either, because even if the variable behaves differently in other situations, unless one can specify the correlation between belief attribution and whatever other variable is tested, that won't give us the information wanted. In fact, reports of other propositional attitudes, assertion, etc. are very likely to be affected differently by RKotA than BRs, or not be affected by it at all; and even if they vary according to it, this might not have implications on belief reporting. Drawing conclusions about whether the data reflects or not a matter about BR by testing similar vignettes on other situations apart from belief reporting would be methodologically incorrect, so that can't be an alternative. For the moment, it seems this is an issue that the RKotA vignettes have to tolerate. Also, other similar experimental designs that I have considered present the foreign language difficulty as well; however, in next section I offer an experimental attempt that can be illuminating on this respect.

Now, independently of the particular matter about experimental design, it is important to remark that the vignettes here presented are modelled on the basis of the cases suggested by the likes of Wettstein, Perry and Crimmins, and Richard (and also the 'chartreuse'/ 'the colour of that dress' case by Capellen and LePore, or the 'Donostia'/ 'San Sebastian' case by Perry). The basic structure of the 'Esther'/ 'Madonna' case here presented follows that of these cases. So, one could argue that the very examples that some philosophers have used to illustrate the need for an appeal to contextuality also suffer from the foreign language difficulty. If this is the case, the intuitions elicited by

these cases that might make one agree that there is indeed a context shifting phenomenon, are due to a conversational issue, not to a semantic or even a pragmatic one. Hence, the original examples by the philosophers might not be as suggestive as they look at first sight.

In any case, the results gathered here show that in cases similar to those offered by some philosophers to encourage contextualism of BRs, we find variability in intuitions. Whether this is to be explained by means of a pure conversational, a pragmatic, or a semantic phenomenon is left open. I think this is a matter that most likely won't be determined by the use of surveys.

A second and slightly less complicated worry is to do with the uniformity between the names used. As 'Madonna' is generally speaking a more widely recognized name than 'Esther', one could think that this might affect the evaluation of familiarity/unfamiliarity by the participants. In this line, it probably would be best to have cases in which the names used are equally known and introduced in the story.

I performed a pilot test[65] of new vignettes with the aim of overcoming difficulties in the design of the vignettes that could have been derived by asymmetry of the names. In this pilot I address the second worry, using more symmetrical names. As regards the major worry, I think that with this pilot I offer a valuable input of information about both the familiar and the unfamiliar name which I hope provides a hint to a deeper understanding of the patterns shown. This won't solve the foreign language difficulty, but can certainly be a helpful guide to a better methodology for testing correctness of BRs in cases where RKotA varies.

The specific features of the application and the results are presented below.

5.1.4. RKotA Remake (RKotA-R):

Following a methodology similar to that of the previous tests a new experiment was designed for the evaluation of the RKotA variable. The vignettes in this test were also designed to resemble even more the kinds of cases some philosophers have used as illustration of the need to appeal contextuality.

---

[65] More information as regards pilot tests can be found in section 6.1.1.

This application is performed as a *pilot test*. Pilot tests are usually performed to check whether the researcher is correct in all his assumptions about the sample, design and type of questionnaire. As mentioned by Richardson, Ampt and Meyburgh (1993), "If the survey designer has been correct in all the assumptions they have made in the design of the sample and the questionnaire, then the pilot survey will not pick up any problems and, in many cases, the data obtained can be combined with the rest of the data obtained from the main survey" (p. 213). Pilots are generally –and are here- applied with small samples (usually around 10% of the actual a number of participants wanted for the study). In this subsection I present details about the application and design of this test.

5.1.4.1. *Participants*

26 undergraduates at the University of Barcelona majoring in psychology (3rd -4th year) participated. 22 female and 4 male, with a mean age of 29, 53[66]. Participants did not have previous higher education in philosophy and were fluent speakers of Spanish, the language of application.

5.1.4.2. *Instrument and procedure*

5.1.4.2.1 *Probes*

For this pilot, the probe consisted in two contrasting Scenarios where the knowledge of the audience varies. In these vignettes the story told is intended to allow more parity between both the name explicitly said is familiar to the audience and the one that is unfamiliar to the hearers in each story. Also, to provide better conditions for testing and gather more insightful data from the reduced sample, I here change the type of question made and scale used. In the previous experiments only one attribution was asked to be evaluated in both Scenarios, namely, that "Jorge believes Esther is an incredible singer". In the RKotA-R vignette, instead, I ask two questions in both scenarios, intending to evaluate the pattern of attribution and to contrast familiarity/unfamiliarity not only across scenarios, but also within each vignette.

The probes are the following (for the original version in Spanish of these probes, go to Appendix 5):

---

[66] A considerable number of the participants were in a program called *The University of Experience* addressed to people over 55.

Scenario 1

Suppose you are friends with John, a first year Classics student in New York. One day, you and John are having lunch together and engage in a conversation about oratory where he says: "Cicero is the best orator ever". You know that 'Cicero' and 'Tully' are both names for the same orator, but in the US people, including John, commonly do not know this, and generally they only use the name 'Cicero' to refer to the roman orator.

Now suppose you go as a visitor to another American university and you are explaining the conversation you had with John to a few students you met there.

Please mark how correct would it be for you to say to your audience either of the following:

a.   John believes Cicero is the best orator ever.

	Totally Correct
	Correct but not the best
	Incorrect but not the worst
	Totally incorrect

b.   John believes Tully is the best orator ever.

	Totally Correct
	Correct but not the best
	Incorrect but not the worst
	Totally incorrect

Scenario 2

Suppose you are friends with John, a first year Classics student in New York. One day, you and John are having lunch together and engage in a conversation about oratory where he says: "Cicero is the best orator ever".

You know that 'Cicero' and 'Tully' are both names for the same orator, but in the UK people commonly do not know this is so, and generally they only use the name 'Tully' to refer to the roman orator, while in the US happens just the opposite and people, including John, normally use the name 'Cicero' when referring to the orator.

Now suppose you go as a visitor to a British university and you are explaining the conversation you had with John to a few students you met there.

Please mark how correct would it be for you to say to your audience either of the following:

a.   John believes Cicero is the best orator ever.

Totally Correct
Correct but not the best
Incorrect but not the worst
Totally incorrect

b.   John believes Tully is the best orator ever.

Totally correct
Correct but not the best
Incorrect but not the worst
Totally incorrect

5.1.4.2. Procedure

All aspects of the procedure including instructions, time, setting and language of the probes were the same as for the RKotA vignettes, with the particularity that for this pilot test the participants were only presented with a Juxtaposed probe. The decision to apply exclusively the Juxtaposed vignette instead of the single Scenarios was taken considering both that the sample was small and the fact that, according to conclusions achieved in this study and previous suggestions by others as already explained, this type

of application seems to aid a finer contrast between Scenarios, allowing me, thus, to reveal with more clarity the pattern of answers.[67]

6.1.4.2.1. Data collection and analysis

The data were collected with paper and pencil questionnaires in which the vignettes were presented with a scale of 4 points[68] to measure correctness/incorrectness of attribution in every case. The scores were spelled out as 'totally correct', 'correct but not the best', 'incorrect but not the worst' and 'totally incorrect'. Scoring was straightforward. The data were gathered during the month of May 2012.

The data collected are analyzed by means of statistic methods. Specifically, for this pilot, taking into account the sample size, all p-values[69] of the contrast between groups are obtained by the application of a standard Kendall's $\chi 2$ test.

5.1.4.2.2. Expected Results

If the test is well designed, no complications for answering or ambiguous questions should be found by the participants and no odd answer patterns should appear. The expected results for the RKotA-R probes are the following: in Scenario 1 for attribution *a* -the one in which a the name with which the audience is familiar-, there should be a tendency towards answers 1 and 2 (scores in the correctness zone), contrasted with answers to attribution *b* –the unfamiliarity one-, expected to be located primarily in marks 3 and 4 of the incorrectness zone. On the contrary, for Scenario 2 we should expect an increase in answers 3 and 4 (incorrectness zone) for attribution *a* –the name in this attribution is unfamiliar to the audience for this situation-, and vice-versa for attribution *b,* where one would expect scores in the correctness zone.

5.1.4.3. Results[70]

---

[67] For this reason also, it is to be expected that answers to Single applications in a test with a larger sample, be more moderate than the one showed by these results.

[68] The choice of a 4 mark point instead of a 6 point one was for statistical reasons. As the sample for this test was small, to have less possible answers allows for a better assessment of judgements.

[69] In this section I offer in a footnote attached to each p-value, the raw values of both Kendall's and $\chi 2$ tests, as well as the degrees of freedom for each contrast.

[70] The individual marks of this pilot test can be seen in Appendix 6

Graph 29: Contrast *a* & *b* Scenario 1

Graph 30: Contrast *a* & *b* Scenario 1 Radial
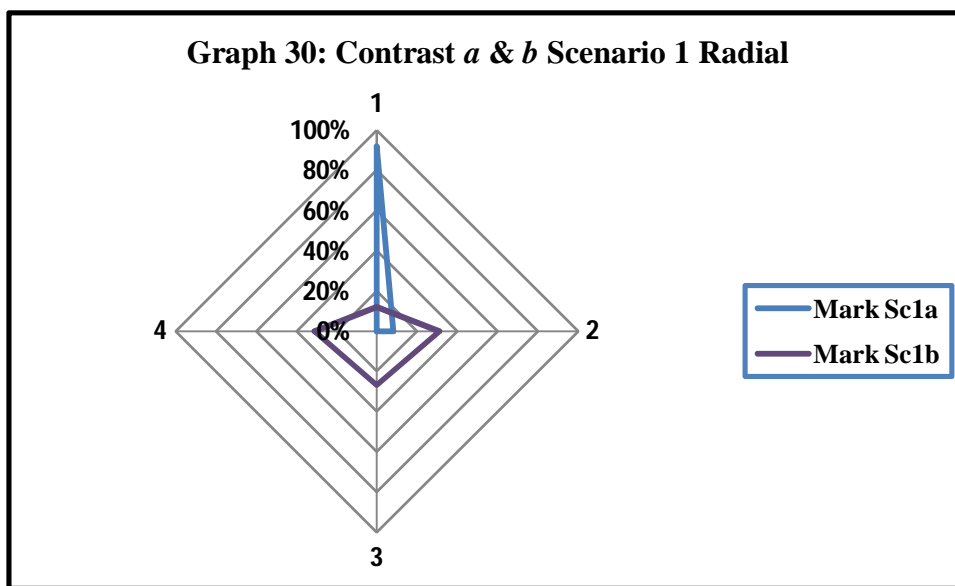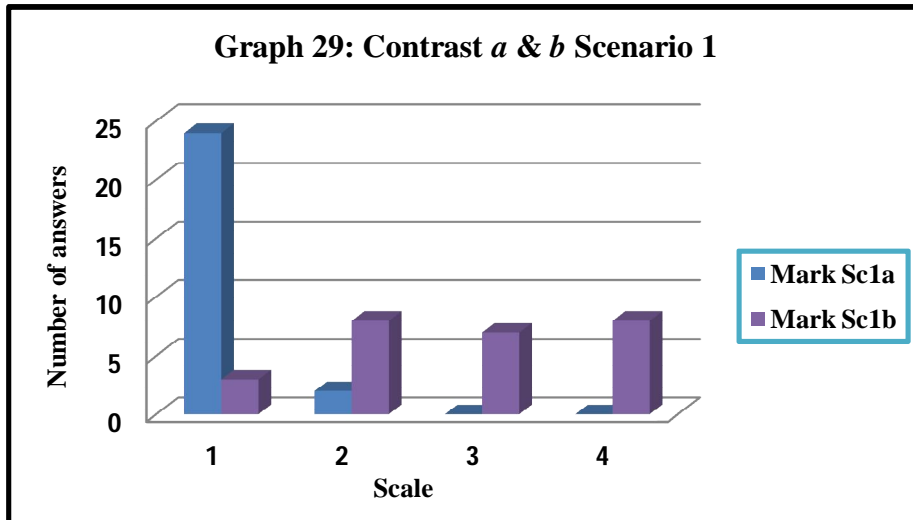
In Graphs 29 and 30 we observe the contrast between attributions *a* and *b* in Scenario 1. Light blue stands for attribution *a*, and purple represents attribution *b*. We see a decisive dominance of mark 1- totally correct- (24 out of 26 participants chose it) for the attribution that 'John believes Cicero is the best orator ever' (*a*), where 'Cicero' is the name that the audience is familiar with. For this attribution, marks 2, 3 and 4 obtain a bare minimum percentage of answers, confirming a clear tendency towards the polar answer in the correctness zone. On the contrary, for attribution *b* -that employs a name unfamiliar to the audience- we see that answers are evenly distributed between marks 2, 3, and 4, with no more than 8 answers in each score. When we look at the radial graph it is possible to see clearly a sharp figure for attribution *a*, showing a striking tendency towards answer one, while for the answers to how correct would it be to make attribution *b* we see a flattened shape that reveals no clear prevalence of any particular

score. P-value for attributions *a* and *b* in Scenario 1 is p< 0,0001[71], showing a very significant the difference.[72]

Graph 31: Contrast a & b Scenario 2

Graph 32: Contrast *a & b* Scenario 2 Radial
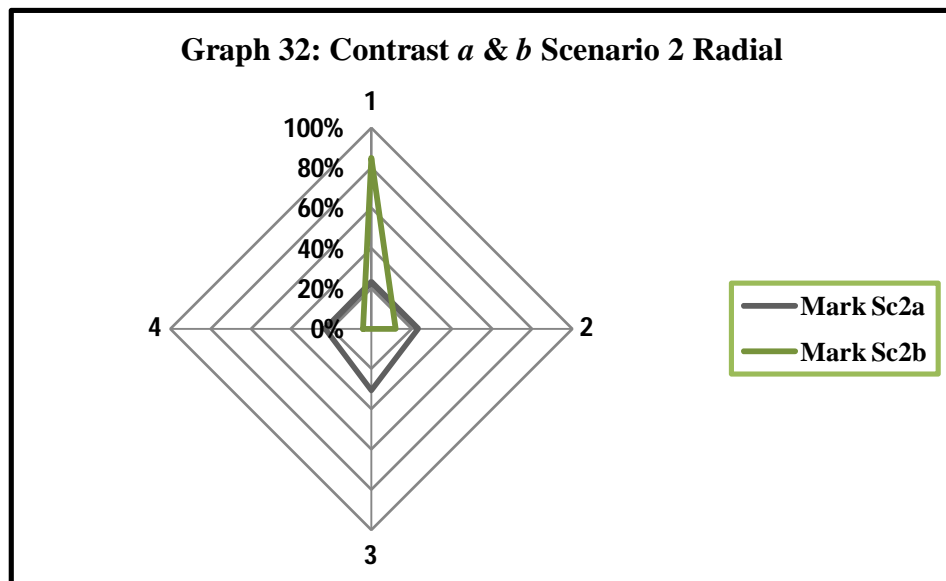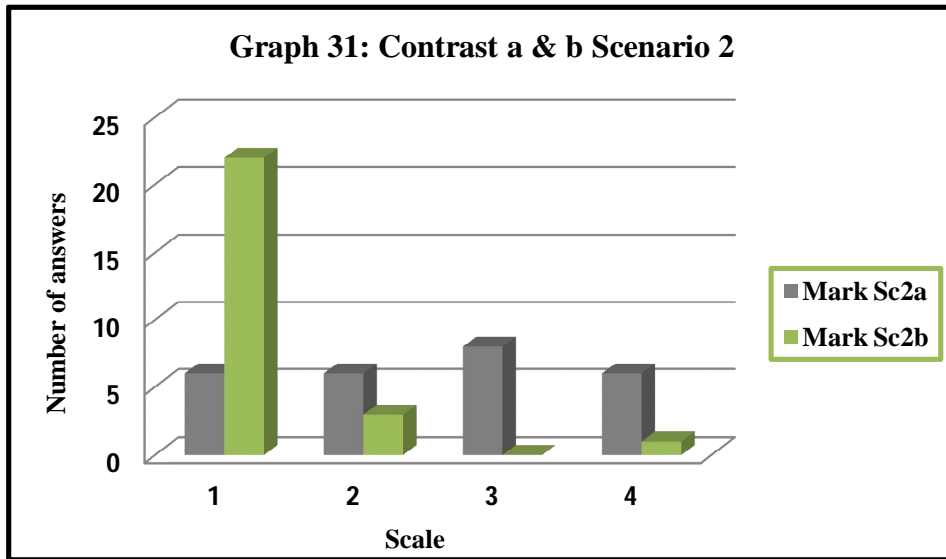
Graphs 31 and 32 illustrate the contrast between attributions *a* and *b* in Scenario 2. Green colour represents attribution *a* and grey corresponds to attribution *b*. In this Scenario the reporter is speaking to an audience that is only familiar with the name 'Tully' for the orator, employed in attribution *b*. Here we observe a uniform number of answers in all four marks as regards attribution *a*, in which a name unfamiliar to the audience is used. In the radial graph we can confirm that the grey shape has no sharp angles and shows only a slight tendency towards mark 3. In contrast, score 1 -totally

---

[71] χ^2 value= 19.000. Kendall's W=0.731 Degrees of freedom= 1

correct- is noticeably overriding for attribution *b* (22 people, a 85% of the sample, scored it), and exceeds in more than three times the number of answers allocated in this score for attribution *a*. Marks 2, 3 and 4 are almost deserted for attribution *b*. We can see in the radial graph that the two figures have very different shapes and are pointing to opposite directions, showing that patterns of response to *a* and *b* attributions are quite dissimilar. P-value for the difference between attributions *a* and *b* in Scenario 2 is p< 0,0001[73], very significant.

Graph 33: Contrast Scenarios 1 and 2 attribution *a*

Graph 34: Contrast Scenarios 1 and 2 attribution *a* Radial
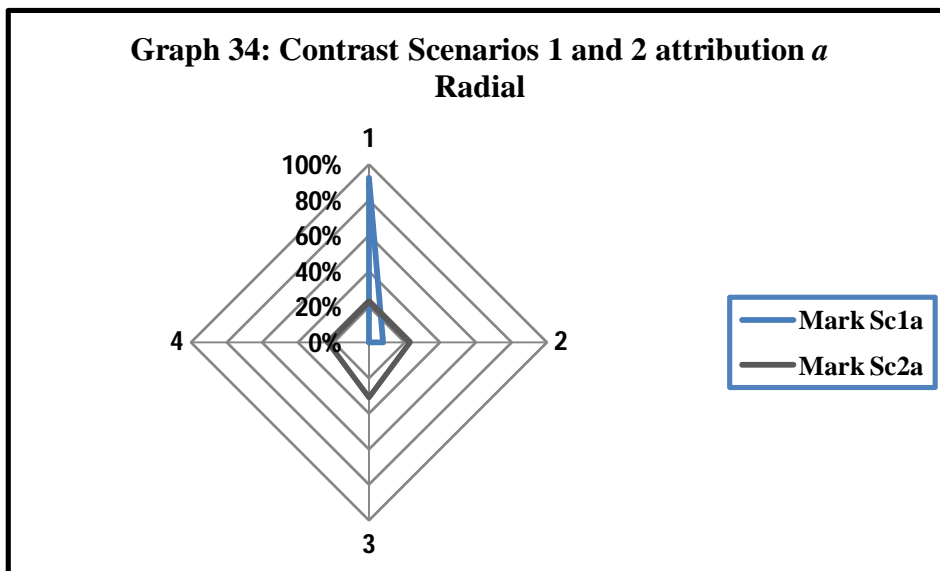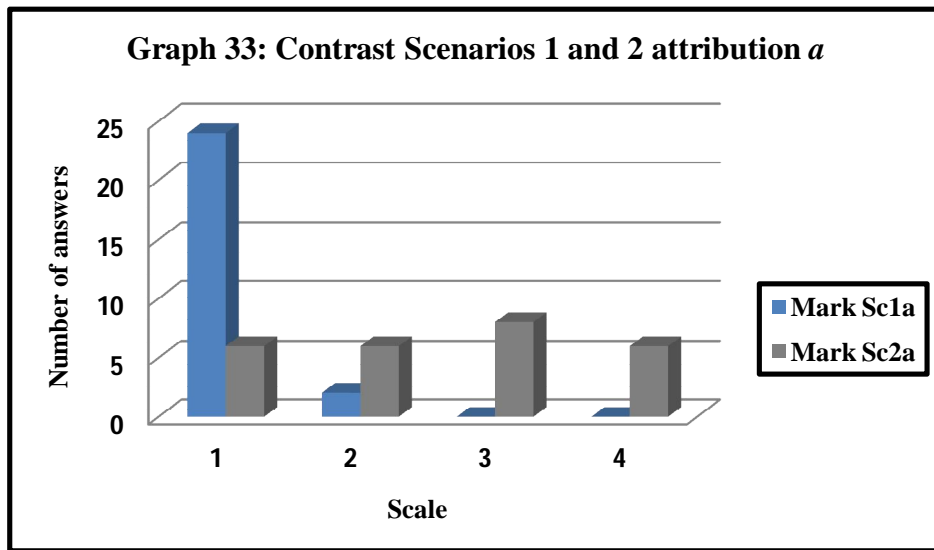
In graphs 33 and 34, I compare answers for attribution *a* in both Scenarios 1 and 2. Scores for Scenario 1 can be appreciated in grey and blue represents Scenario 2. In these graphs we observe a substantial dominance of mark 1 for Scenario 1, contrasted with

---

[73] χ^2 value= 19.000. Kendall's W=0.731 Degrees of freedom= 1

evenly distributed answers for Scenario 2. We can verify the disparity in choice patterns in the radial graph, where we see two very different shapes which barely match on the central percentages. Subjects evaluated attribution *a* as correct more than three times in Scenario 1 than in Scenario 2 (in the radial graph we observe a 92% for Scenario 1, against a 23% for Scenario 2). For scores 3 and 4 -the incorrectness zone- no answers are registered in Scenario 1, while we can see for Scenario 2 slightly more than a half (14) of the subjects chose scores in this area. A $p < 0,0001$[74] shows quite a significant difference between the groups.

Graph 35: Contrast Scenarios 1 and 2 attribution *b*

Graph 36: Contrast Scenarios 1 and 2 attribution *b* Radial
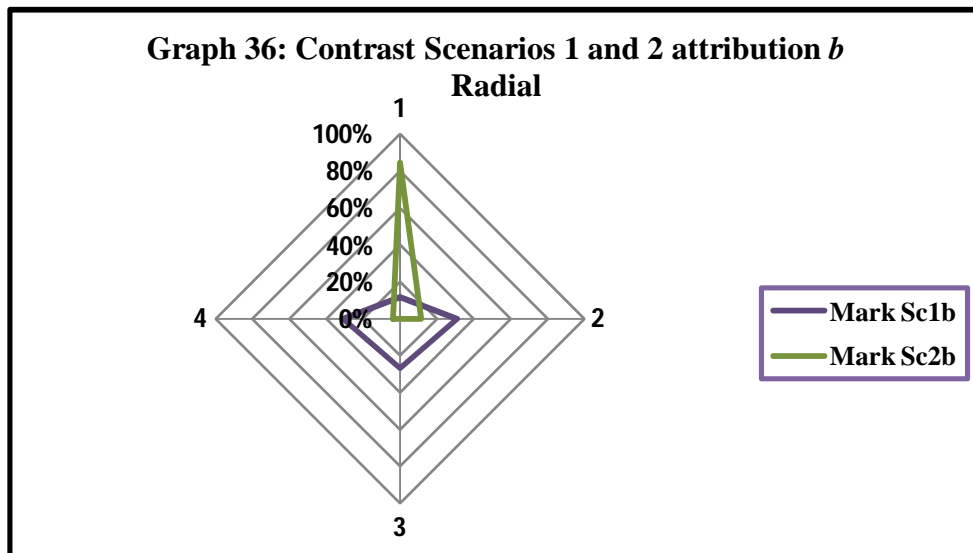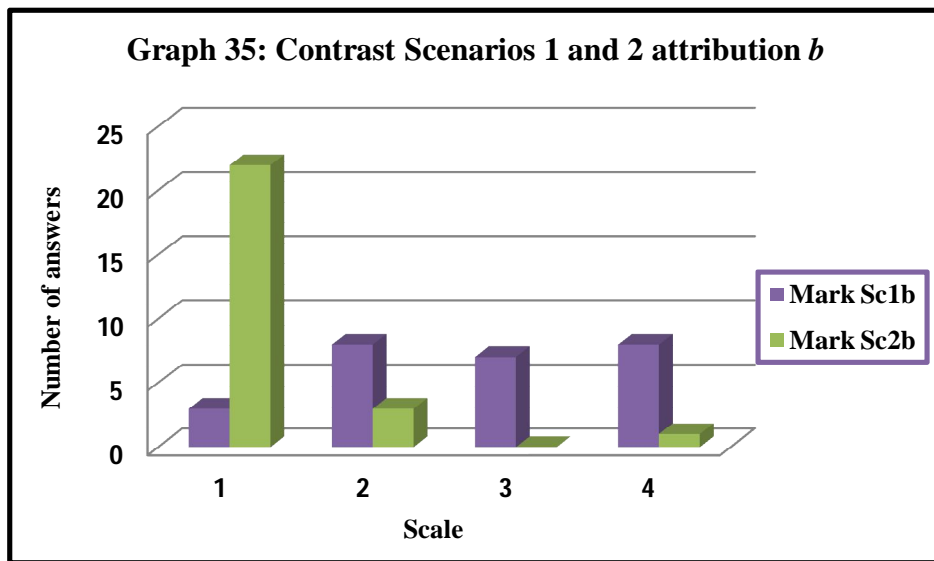
In graphs 35 and 36 we see a comparison between answers to attribution *b* in Scenarios 1 and 2. Scenario 1 is represented by purple and Scenario 1 by green. In the radial graph

---

[74] χ^2 value= 19.000. Kendall's W=0.731 Degrees of freedom= 1

we observe a flattened diamond for Scenario 1 contrasted with a sharp triangle that points to mark 1 for Scenario 2. This, again, shows important dissimilarities between the behaviour of responses to both scenarios, where the most salient disparity is found in score 1 –totally correct- chosen by very few subjects for Scenario 1, but preferred by a great majority in Scenario 2. In fact, for mark 1, Scenario 2 reaches 8 times the number of answers in Scenario 1. Marks 2, 3 and 4 are homogeneously selected in Scenario 1, and correspond to 89% of the answers to this scenario, while for Scenario 2 only a 16% of the answers is allocated in these scores. There is a very significant difference between the groups: $p < 0,0001$[75].

5.1.4.4. Conclusions on the RKotA-R test

As this is a pilot done with 26 people, the conclusions drawn about it should be confirmed by an application with a larger sample and can only be regarded as a tentative approach to possible results of such administration. However, even if they are only provisional, the data collected seem to show clear tendencies that support contextualist claims about belief attribution.

If we compare answers to both Scenarios (Graphs 29-36) we can see that they show opposite patterns in the following way: for the attribution that 'John believes that Cicero is a great orator' in Scenario 1, where the audience is familiar with the name 'Cicero' people show a very marked tendency to evaluate the attribution as totally correct. On the contrary, the attribution that 'John believes that Tully is a great orator' is evaluated as totally correct by the great majority of participants in Scenario 2, where the audience is only familiar with 'Tully' as a name for the orator. This goes in accord with the initial prediction if the h_1 was correct, namely, if the referential knowledge possessed by the audience affected correctness of belief attribution. For this pilot test, people seem to be considerably inclined to adjust attribution in consideration of the familiarity/unfamiliarity that the hearers of a report have with a name.

In principle, one would also expect that for the attribution where the non familiar name is employed, answers would be located in the totally incorrect pole, or at least mainly in the incorrectness zone. However, we observe for both scenarios that these attributions do not seem to be considered particularly incorrect. They are certainly taken by very

---

[75] χ^2 value= 19.000. Kendall's W=0.731 Degrees of freedom= 1

few as totally correct (mark 1), but we see that for marks 2, 3 and 4 there is an even distribution of responses in both Scenarios 1 and 2 for the attribution with the unfamiliar name. This can be associated with lack of decisiveness. I hypothesize that, simply and plainly, participants don't seem to be sure about how correct it would be to make an attribution with a name that the audience might not understand.

In contrast with the RKotA vignettes, the RKotA-R ones seem to show clearer results, with less variability within the scores. Subjects were surer of their answers, as they were more prone to chose scores in the poles of each zone. It seems to me that the design of the RKotA-R test gets us closer to more accurate data, and gives interesting information for analysis that the RKotA did not provide (namely, the two question methodology shows evaluation of attribution with familiar and unfamiliar names for both vignettes). Based on the results here gathered, it would certainly be interesting to expand the study and test it with a bigger sample. It seems highly possible that by doing this one could find empirical backup and input for theories that endorse contextualism of BRs.

5.2. About the stakes tests.

The results of the application of the Stakes vignettes reveal that subjects show a tendency to evaluate the belief report that 'Luisa believes Andrés should order the tuna' as correct for both the Low-stakes (Scenario 1) and the High-stakes (Scenario 2) situations, in the Single application. This can be observed clearly in Graphs 21 and 22, which compare the answers obtained Scenarios 1 and 2 in the Single application. On the contrary, for the Juxtaposed application there seems to be different patterns of response so that for the Low-stakes Scenario there continues to be an inclination to consider attribution of belief as correct, while the High-stakes scenario we see an ambiguous tendency to consider the attribution both as totally correct and as totally incorrect. This can be verified in graphs 23 and 24, that contrast the results for Scenarios 1 and 2 in the Juxtaposed application. These results seem to partially disconfirm the hypothesis that BRs are sensitive to Stakes of the agent. The results of this study allow the EPS theorists who had not considered the possibility that belief attribution would be contextual, to conditionally[76] reject this hypothesis.

---

[76] As I explain in what follows, new tests and re-tests should be performed before we reach definite evidence

As the results are considerably dissimilar for the Single and the Juxtaposed applications, I will next present the analysis and conclusions to each of them separately. I offer a general conclusion for both applications at the end of this section.

5.2.1. About the Single application

For the Single application we see (Graphs 21, and 22) that people showed an analogous tendency to consider the attribution totally correct for both High and Low stakes situations. If we see the data of this application closely, we observe that for the Low-takes Scenario there is a gradual and steady decrease in number of answers from scores 1 to 6 (only 1 person chose the latter), with a dominance of responses in the correctness zone (86%). This is a very solid pattern of answers, as expected for the Low-stakes vignette. However, if my working hypothesis (i.e. that stakes for the agent matter to belief attribution) was correct, one should have also expected that that most subjects' responses to the High-stakes vignette would be allocated in the incorrectness pole of the scale. This does not happen. Instead, not only there is no significant difference between the groups, but the number of answers for scores 1-5 is similar for the two Scenarios. Only mark 6 –totally incorrect- sets apart from the rest in that there is a considerable difference in number of answers for both Scenarios, with 10 answers for the High-stakes Scenario against only 1 for the Low-stakes one. However, it seems yet more important to notice that for mark 1 -'totally correct'- the amount of answers for the High-stakes Scenario (28 subjects chose it) is even higher than for the Low-stakes one (26 participants), quite contrary to the expected result.

Remember that High and Low stakes Scenarios are different in a very crucial fact: Attributing to Luisa the belief that Andrés should order tuna in the High-stakes Scenario raises the stakes in the agent's context: Andrés could die as a consequence of her advice. In contrast, the same recommendation has no particular costs in the Low-stakes situation. All other aspects of the vignettes are stable and the agent's mental state remains fixed. It is indeed intriguing that the results show so little variation between judgements of correctness/incorrectness through both situations in the Single application.

The most obvious explanation for similarity in answer patterns to both Scenarios of the Single application is that people simply do not have contextualist inclinations as regards stakes as a contextual parameter. If this was the case, probably the subjects were more

126

interested in preserving the report faithful to the agent's words than in the possible consequences that the attribution could have. However, I would like to offer yet another possible explanation; my hypothesis is that the participants may have overlooked the connection between the recommending/ordering of the food and the actual food poisoning, and for that reason did not quite get that by making the attribution the agent's stakes rose.

Actually, there is an important difference between recommending/ordering a dish and actually eating it, and this might have been relevant to the subjects at the time of evaluation. Luisa's recommendation might be perceived as an "innocent" act that once the dishes are served will not cause any harm because, for instance, Andrés himself would refuse to eat the tuna once he sees it, or the reporter, aware of his allergy, would prevent Andrés from eating the fish, or alert Luisa so that she takes measures in order to avoid possible consequences, etc. Many things can happen between the moment of ordering a dish and that of eating it, and this might have had the effect that the high-stakes in the experiment turn barely mild to the subjects' perception.

I reckon that constructing vignettes where it is clearer that the relation between recommending the food and the possibility that Andrés dies is causal for the High-stakes scenario, could give more accurate results. This connection can be highlighted, for example, stressing that Andrés would not recognize the raw Japanese tuna if it is served to him and he could perfectly eat it and die, etc. I think it would be worth it to re-design and apply improved stakes vignettes in further work, in order to compare whether the results here obtained are due to matters about the design or they reflect actual intuitions about belief reporting in cases where stakes vary.

For the moment, though, and based on the data collected for the Single application, one should be ready to dismiss the working hypothesis, that is, one should be prepared to say that what is at stake in the agent's context does not affect BRs. However, an interesting twist happens in the data gathered with the Juxtaposed application.

5.2.2. About the Juxtaposed application

In the Juxtaposed application -illustrated in Graphs 23 and 24-, we observe a very different distribution between the answers to High and Low stakes vignettes than we saw for the Single administration. In this administration, the amount of people who

judged the report as totally incorrect –mark 6- for the High-stakes scenario outnumbers for almost three times the total of subjects that chose this answer in the Low-stakes situation, and is significantly above the number obtained for this same score and Scenario for the Single application. This means that when given the possibility to contrast between Scenarios, the participants were more prone to judge the attribution that 'Luisa believes Andrés should order the tuna' as incorrect when stakes rose. This results show a pattern more coherent with the initial assumption that stakes matter to belief reporting. Also, the difference between Scenarios 1 and 2 in the Juxtaposed application is significant (p=0.010). I confirm again that giving the subjects the opportunity to contrast between the stories and correct their judgements does have an effect in the results and is a methodology that seems to help make participant's judgement sharper.

A quick conclusion one might draw just by regarding the significant difference that corresponds to the comparison between Scenarios and the raise of answers 6 for the High-stakes vignette, is that for the Juxtaposed application, the h_1 is confirmed. However, a more detailed examination to the data is required as we see other considerably crucial information that might have an effect on this quick conclusion.

In the Graph 24 (radial) we see that the two figures that represent the answers to both Scenarios overlap in very significant points, for instance in mark 1, showing a considerable number of people (18) who still considered totally correct the attribution in the High-stakes Scenario. Also, in the Graph 23 (bars) it is possible to observe that polar answers 1 and 6 both obtain an almost an equally elevated number of answers for this Scenario (18 answers for mark 1, and 19 for mark 6), presenting a very ambiguous pattern of response to the vignette.

Moreover, if we look at the non-mainstream data, we verify that for the Juxtaposed probe the number of totally incorrect –score 6- answers elevates with respect to the Single application for Scenario 1, where stakes are low: 7 people thought that the attribution was totally incorrect in the Juxtaposed application, while only 1 thought so in the Single administration. This means that comparing scenarios not only did not make participants less likely to consider incorrect the attribution when nothing is at risk, but, surprisingly, made them more prone to judge it as totally incorrect.

Again, this could be simply a part of the subjects that do not have contextualist inclinations. However, the fact that comparison would elevate the likeliness that one would consider the attribution as totally incorrect in the High-stakes Scenario is rather baffling. One possible reason for this is that, considering that the main character in both vignettes is *Andrés* and the situations are similar, participants could have thought that with the new information in Scenario 2 they should not make the attribution in Scenario 1 either. In other words, given the possibility to contrast, subjects took into consideration the information in Scenario 2 to evaluate also Scenario 1. It would be needed to confirm whether this hypothesis is correct by making experiments where High and Low stakes vignettes had different names for the characters. In any case, it would be recommended to do it in further work to avoid possible biases.

For the Juxtaposed application, in general, I think that rather than a decisive contrast between scenarios with the expected dominance of answers in the correctness zone for the Low Stake Scenario and the corresponding polarization towards incorrectness for the High-stakes, the data show an unclear arrangement of the answers. It is highly possible that the difference between groups is significant for this reason only, and therefore cannot be interpreted as evidence in favour of the claim that stakes matter to belief reporting. Certainly, the results are dissimilar between the scenarios, but not quite conclusive in the way one would have expected to confirm the working hypothesis.

Conversely, although the evidence is not decisive, we should not be ready to accept the null hypothesis (h_0) either. There is quite an important rise in the number of people who evaluated as totally incorrect the attribution in the High-stakes Scenario, and that is information that should not be discarded. Instead, the suggestion I make is to refine the probes and perform improved applications including the modifications previously mentioned.

5.2.3. General conclusions on the Stakes variable

Contrary to my assumption here, the results seem to show little support to the claim that stakes is a relevant parameter that affects belief reports. The data collected seem ambiguous and it is suggested to perform improved tests before reaching a definite conclusion on the matter.

Above, I offered a few possible explanations for the data collected. Those are explanations centred on the design of the vignettes and on possible interpretations the subjects could have made about them, however, it might not be so surprising that in the tests performed this parameter does not behave as initially presumed. As described in Chapter 1, proponents of sensitivity of belief attributions actually do not contemplate stakes of the agent as a relevant parameter for belief reporting. It is true that these authors do not declare openly which features of the context could contribute to variability of truth value of BRs, but they surely give several hints between which stakes does not take place. The selection of this parameter as one that was worth testing, was not made on the basis of the suggestions about contextuality of BRs. So at least as regards the specific proposals by some philosophers as regards which contextual parameter might affect BR's, one should not expect variation in judgements towards the stakes vignettes.

Instead, the choice of stakes as a parameter to examine was made on the basis of the literature in epistemology. As explained in Chapter 2, traditionally Contextualist and Interest-relative invariantist theories endorse that truth value of knowledge attributions depends on what is at stake (either in the context of the agent or for the speaker). Proponents of these theories have focussed mainly on the justification for knowing, I explained. According to them, a single attribution may vary in truth-value in low and high stakes situations because with the costs of being wrong, the justification required for knowledge also rises: in one context the agent is justified enough for knowing but not in the other, while the evidence he counts with and his mental state are held fixed. However, these authors have not suggested that stakes should have an import of any kind on belief or belief attributions; precisely that is the reason why it seemed an interesting empirical objective to pursue and a legitimate hypothesis to test. Taking into account the basic definition of knowledge as justified true belief, and given that epistemologists have not considered the possibility that the belief component of knowledge was affected by stakes, I proposed to test if stakes was a relevant contextual parameter in the case of belief attributions.

It is important to emphasize that it is not clear that EPS theorists had considered the possibility that belief attribution was contextual and dismissed it. Instead, on the basis of the literature it looks like they assumed the contextuality of the justification without allowing for the option I here put forward. Now, having considered and tested this

hypothesis, epistemologist can, on the basis of the results here presented, at least provisionally reject it.

In fact, on the basis of the data gathered I conclude that, if it was the case that knowledge attributions were sensitive to contextual parameters, it might be justification that is affected by stakes, not belief. Given that the evidence here is not sufficiently conclusive, drawing conclusions about the relation between findings on belief attribution and knowledge attribution might be risky and not well founded. For the moment, I should simply say that the conclusions reached here cannot be extrapolated to knowledge attributions. In fact, on the face of the results obtained, it cannot be said that stakes affect judgements of correctness on BRs.

On another issue, it would be interesting to examine Jennifer Nagel's proposal under the light of the data here gathered, as she explains intuitions about high/low stakes cases in knowledge attribution relying on modules of cognition that could be generalized for assessment of other several propositional attitudes apart from knowledge. In fact, I want to suggest here that the psychological mechanisms described by Nagel about knowledge attribution (mentioned in section 2.5 of Chapter 2), can be said to also apply for reasoning about belief, and should therefore affect our belief-attributing practices. Actually, Nagel mentions how attribution of mental states such as knowledge and belief seem to work in a similar way based typically on System 1: "One might expect that assessments of the mental states of others would invariably demand higher level controlled processing. However, there is evidence that routine assessments of the knowledge or belief of others is not inherently taxing but rather modular and automatic in character" (2011, Pg.13).

Nagel's proposal is that certain psychological mechanisms underlie our attributing behaviour in high and low stakes situations, and she goes as far as suggesting that this is so not only for epistemological claims but also for the case of BRs. For forming judgements about high-stakes situations we use System 2, designed for tasks that involve bigger and more controlled cognitive effort, while low-stakes cases are usually processed with System 1, designed for making every day, heuristic, automatic and simple judgements. In this line of thought, in Scenario 1 we should use System 1 to evaluate the attribution that 'Luisa believes Andrés should order the tuna', but in Scenario 2 we engage in System 2 to form a judgement of correctness because the

context makes it more difficult, and it requires a processing that exceeds the heuristic cognition.

It seems, prima facie, that it is possible to follow a similar reasoning based on Dual Process Theory, need-for-closure and cognitive processing employed by Nagel for explaining judgements towards high and low stakes situations in knowledge attributions, to understand situations of this nature in belief reporting. If this was the case, one should expect decisive results showing that subjects have contextualist linings for the stakes tests. Moreover, the expected pattern of answers should be even more evident in the Single application of the vignettes, because in Juxtaposed cases the possibility to contrast gives access to more elements of judgement and may affect the automatic cognitive response. However, as we have seen, the results show exactly the opposite.

Clearly, I cannot reject Nagel's proposal that cognitive processes underlie our judgements in high and low stakes situations on the basis of the data here collected. However, for the suggestion that we may attribute beliefs in high and low stakes contexts differently due to cognitive modules, the results presented seem to show this is not the case. Even if the data here are not conclusive, it seems to me that given that Nagel's argument is set on a psycho-biological framework, the results for the tests should be categorically decisive. In any case, it would be coherent with the data available to say that System 1 and System 2 do not regulate assessment of belief in high/low stakes cases.

5.2.4. Possible worries

The major worry when confronted with the ambiguous results is whether the vignettes are well designed. Previously I mentioned several points in the probes that could be improved in further work. In this section I want to explore other details that could be possible sources of error in the experiments.

EPS theorists' proposal is that either stakes for the speaker (Contextualism) or the agent (IRI) affect truth value of knowledge attributions. I have suggested before that it might not be clear for the participants that there is actually something at stake either for the agent or for the speaker. Here I want to take this observation further. It is evident that the probes here presented do not intend to measure whether correctness of attribution is

affected by the costs of getting it wrong for the speaker, as he is not particularly affected by the raise of stakes. Instead, I consider them to be closer to stakes of the agent, as it can be understood by the circumstances described in the High-stakes vignette that there are she is involved and can be held responsible for a very serious consequence of the attribution (food poisoning and ultimately the loss of a human life). However, it can be argued that In Scenario 2 the costs of being wrong do not necessarily affect the agent *directly*, namely, that in the vignette it s not the agent's life itself that is at risk.

This worry can lead to improvement of the vignettes. A re-design of the tests including a high-stakes story where it is more obvious that there is an elevated cost in making the attribution that falls upon the agent directly, might be in order. Further work that integrates the observations here presented will help us verify if by making the adjustments suggested we find more clear and decisive data against or in favour of sensitivity to stakes of BRs. Also, it is left to future work to test cases of belief attribution where stakes are high and low for the attributor, a proposal that would be closer to epistemological Contextualism.

5.3. General conclusions about context -sensitivity of belief reports

Some semantic theories of belief reporting are built on the basis of the claim that the attribution of belief is sensitive to features of the context and some of the examples typically invoked to justify the plausibility of the view are cases where the relevant contextual parameter is the referential knowledge of the addressees. I think the analysis of the empirical data collected from the RKotA and RKotA-R vignettes tend to support that claim.

On the other side, some epistemological theories endorse that what is at stake for the agent or the attributor affects the justification for one's true belief to count as knowledge, and for that reason, the truth-value of knowledge attributions. Taking into account that knowledge is justified true belief, I proposed that it would be interesting to consider and explore whether belief was affected by what is at risk by evaluating whether BR's varied according to what is at stake for the agent. I conclude on the basis of the analysis of the data gathered from the stakes vignette that this claim is not supported, and therefore, that it is in principle correct to suppose that the justification be the aspect of knowledge claims that is sensitive to stakes.

It is important to remark that the scope reached by this research is only partial. Philosophical theories cannot and should not be established or rejected purely on the basis of surveys. The data gathered here show at most what people tend to do: people tend to judge a BR as correct or incorrect taking into account the referential knowledge of the audience, and they are inclined to disregard what is at stake in the agent's context for evaluating correctness of BRs.

It might be argued that philosophers that endorse contextuality of belief reports were already aware of the presence of contextualist intuitions among ordinary speakers. In fact, the discussion as to whether the phenomenon is semantic or pragmatic presupposes its existence. Nevertheless, the possibility remained that philosophers were focusing on a relatively uncommon phenomenon, or a practice that people engaged in, only in very special circumstances. The evidence gathered here shows that the practice is in fact rather common at least as regards the Referential Knowledge of the Audience.

Now, it remains open whether the tendencies found are to be explained as a purely pragmatic phenomenon that does not affect the semantic theory of belief reporting, or as a full blown semantic fact, i.e., something that affects the truth-value of belief reports, as Mark Richard, for instance, would have it. In any case, that it is a common practice among people to gauge contextual factors such as the Referential Knowledge of the Audience perhaps should be of importance for our semantic theories.

On the other hand, even though they were aware of contextualist intuitions and have made several suggestions as to which features in the context might affect variability of belief attributions, philosophers had not focussed on any specific variable in their arguments. This was recognized by some of the authors as a theoretical gap (see Perry and Crimmins' comment on this in section 1.3.1 of chapter 1). The evidence I present here shows, I think, a relevant parameter.

I can conclude that the hypothesis that belief attributions are sensitive to contextual parameters is supported for the Referential Knowledge of the Audience, but not for the stakes variable. I hope that the data presented here and the analysis of their significance, constitute empirical background and input for theories of belief reporting and epistemology. Also, I expect that the methodological observations made throughout this dissertation be a contribution to experimentation on philosophical matters in the future.

Bibliography

Armendt, B. (2007). 'Stake-invariant belief', *Acta Analytica*, 23: 29-43

Armendt, B. (2008). 'Stakes and beliefs', *Philosophical Studies*, 147: 71-87

Bach, K. (1997). 'Do Belief Reports Report Beliefs?' *Pacific Philosophical Quarterly*, 78: 215–241

Bach, K. (1992) 'Sometimes a Great Notion: A Critical Notice of Mark Crimmins'

Talk about Beliefs'. *Mind & Language*, 8 (3): 431-441

Back, E. and Apperly, I. (2010) ' Two sources of evidence on the non-automaticity of true and false belief ascription'. *Cognition*. N115: 54-70

Braun, D. (1998). 'Understanding Belief Reports'. *Philosophical Review*, 105: 555-595.

Braun, D. (2002). "Cognitive significance, attitude ascriptions, and ways of believing propositions." *Philosophical studies* 108: 65-81

Buckwalter, W. (2010) "Knowledge isn't closed on Saturday: a study in ordinary language". *European review of philosophy, psychology and experimental philosophy*, 1 (3):395-406.

Cappelen, H., and LePore, E. 1997: 'On an Alleged Connection Between Indirect Speech and the Theory of Meaning,' *Mind & Language*, 12: 278-296.

Cappelen, H. and LePore, E. (2005). *Insensitive semantics*. Blackwell Publishing

Churchland, P.M. (1981), "Eliminative Materialism and the Propositional Attitudes", *Journal of Philosophy*, 78: 67-90.

Cohen, S. (1986). "Knowledge and context". *The journal of philosophy*, LXXXIII: 574-582.

Cohen, S. (1999) "Contextualism, Skepticism, and The Structure of Reasons". *Philosophical Perspectives* ,13: 57-89

Dennett, D. (1987), "True Believers: The Intentional Strategy and Why it Works" in *The Intentional Stance*: 13-35. Cambridge, Mass.: The MIT Press

DeRose, K. (1992). "Contextualism and Knowledge attributions". *Philosophy and phenomenological research*, I.II (4): 913-929

DeRose, K. (1999). "Contextualism: an explanation and defense", in Greco and Sosa (eds.), *The Blackwell guide to Epistemology*: 187-205. Oxford: Basil Blackwell.

DeRose, K. 2006. ''Bamboozled by Our Own Words': Semantic Blindness and Some Arguments Against Contextualism'. *Philosophy and Phenomenological Research*, 73(2): 316-338.

DeRose, K. (2009). *The Case for Contextualism: Knowledge, Skepticism, and Context*, Vol. 1, Oxford UP

Frances, B. (2002). 'A test for theories of belief ascription'. *Analysis*, 62: 116-25.

Frege, G. (1872). 'On sense and reference'. Reprinted in P. Geach and M. Black, eds., *Translations from the Philosophical Writings of Gottlob Frege*: 26.78, Oxford: Blackwell.

Hansen, N. (forthcoming (a)). 'A slugfest of intuitions: Contextualism and experimental design'.

Hansen, N. (forthcoming (b)). 'On an alleged truth /falsity asymmetry in context shifting thought experiments'.

Hansen, N. and Chemla, E. (forthcoming): 'Experimenting on Contextualism.' *Mind and Language*.

Hawthorne, John. (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hawthorne, John and Stanley, Jason. 2008. 'Knowledge and Action' *Journal of Philosophy,* 105:571–90.

Hewson, C. (1994) 'Empirical evidence regarding the folk psychological concept of belief' in *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*: 403-406.

Knobe, J. (2003), 'Intentional action and side effects in ordinary language'. *Analysis*, 63:190-193.

Kripke. S. (1979; 1988). 'A puzzle about Belief' in A. Margalit (ed.), *Meaning and use*: 239-283. Reidel, Dordrecht

Ludwig, K. (2007) 'The epistemology of thought experiments: first person versus third person approaches'. *Midwest studies in philosophy*, XXXI: 128-159

MacFarlane, J. (2007) 'Nonindexical Contextualism'. *Synthese*, 166: 231-250

Malcom, N. (1954) 'Knowledge and belief'. Analysis 14 (4), Pgs. 94-98.

Mallon, R; Machery, E.; Nichols, S. and Stich, S (2004). 'Semantics Cross-Cultural Style'. *Cognition*, 92: B1-B12.

Marti, G. (2009). 'Against semantic multiculturalism'. *Analysis*, 69 (1). Pgs. 42-48.

May, J. (2010). 'Practical interests, relevant alternatives and knowledge attributions: an empirical study'. European review of philosophy, psychology and experimental philosophy. Machery, E, Lombrozo, T. and Knobe J. eds., 1 (2): 265–273

Moran, R. (1994). 'Arthur Collins' the nature of mental things'. Philosophy and phenomenological research. Vol LIV (4): 917-920.

Nagel, J. (2011). 'The psychological basis of the Haman-Vogel paradox'. *Philosophers' Imprint*, 10 (15): 2-28.

Nagel, J. (2008). 'Knowledge ascriptions and the psychological consequences of changing stakes'. *Australasias Journal of Philosophy*, 86 (2): 279-294

Phelan, M. (manuscript). 'Evidence that stakes don't matter for evidence'.

Perry, J. (forthcoming): 'The Cognitive Contribution of Names.' A. Bianchi (ed.): On Reference.

Perry, J. and Crimmins, M. (1989) "The Prince and the Phone Booth: Reporting Puzzling Beliefs". The journal of Philosophy No. LXXXVI. Pgs. 685-711.

Pfungst, O. 1911. "Clever Hans". *The Horse of Mr. von Osten*. New York: Henry Holt & Co.

Pinillos, A. (2011). 'Some recent work in Experimental epistemology'. *Philosophy Compass*, 6: 675-688.

Pinillos, A. (1012): 'Knowledge, experiments and practical interests.' In J. Brown and M. Gerken (eds.): *Knowledge Ascriptions*. Oxford. Oxford University Press.

Putnam, H. (1975). 'Meaning of "Meaning"'. *In Language, Mind and Knowledge. Minnesota Studies in the Philosophy of Science*, 7: 131-193, ed. Keith Gunderson.

Reimer, M (2003). 'Reference'. *Stanford encyclopaedia of philosophy.*

Richard, M. (1989). 'How I Say What You Think'. *Midwest Studies in Philosophy* , 14 (1):317-337.

Richard, M. (1990). *Propositional Attitudes*. New York: Cambridge University Press.

Richardson, A. Ampt, E. Meyburg, A. (1995) *Survey Methods for Transport Planning.* Eucalyptus Press.

Rosenthal, R. (1976) 'Experimenter effects in behavioural research'. New York: Irvington.

Russell, B. (1905). 'On denoting'. *Mind*, New Series, 14 (56): 479-493.

Rysiew, P. (2011), Epistemic Contextualism, *Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/entries/contextualism-epistemology/

Saul, J. (1999). 'The road to Hell: Intentions and propositional attitude ascriptions'. *Mind and language* ,14 (3): 356-375

Saxe, R., Carey, S., and Kanwisher, N. (2004). 'Understanding other minds: Linking developmental psychology and functional Neuroimaging', *Annual Revision of Psychology,* 55: 87-124

Schaffer, J. (2004). 'From contextualism to contrastivism'. *Philosophical Studies*, 119: 73-103

Schaffer, J. (2006). 'The irrelevance of the subject: against subject-sensitive invariantism'. *Philosophycal Studies*, 127: 87-107

Schiffer, S. (1996). 'Contextualist solutions to scepticism', *Proceedings of the Aristotelian Society*: 317-333

Schiffer, S. (2006) 'A problem for Direct-Reference theory of belief reports'. *Noûs*, 40 (2):361-368.

Shortsleeve, E. (2005). 'Knowledge and certainty: A Speech- Act Contextualist Account'. Available on-line.

Schwitzgebel, E. (2006) 'Belief'. *Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/entries/belief/

Schwitzgebel, E. (2002). 'A phenomenal dispositional account of belief'. *Nous*, 36: 249.275

Schwitzgebel, E. and MacGeer, V. (2006). 'Disorder in the representational warehouse'. *Child development*, 77 (6):1557-1562

Schwitzgebel, E. (1999). 'Gradual belief in children'. *Human development* , 42:283-285.

Schwitzgebel, E (2001). "In-between believing". The philosophical quarterly 51: 76-82.

Sider, T. (1995). 'Three Problems for Richard's Theory of Belief Ascription'. *Canadian Journal of Philosophy* , 25: 487–514

Soames, S. (1988). 'Reference, attitudes and content' in *Propositions and Attitudes*: 197-239, Nathan Salmon and Scott Soames eds., Oxford University Press.

Soames, S. (2009). "Drawing the Line Between Meaning and Implicature –and Relating Both to Assertion". Noûs 42 (3):440-465

Sosa, E. (2010) "Experimental philosophy and philosophical intuition". Special edition of Philosophical Studies ,132: 99-107

Stanley, J. (2002). 'Making it articulated', *Mind and Language* 17: 149-168.

Stanley, J. (2005). *Knowledge and practical interests*. Oxford University press.

Stanley, J. (2007) Précis of Knowledge and Practical Interests. *Philosophy and Phenomenological Research*, 75 (1): 168-172

Stanley, J. (2010) Comment in *http://experimentalphilosophy.typepad.com/*. Retrieved on November, 2010.

Weatherson, B. (2011) 'Defending Interest Relevant Invariantism' *Logos and Episteme* 2: 591-609.

 Weinberg, J., Gonnerman, Ch., Buckner, C., Alexander, J. (2010) 'Are Philosophers Expert Intuiters?'. *Philosophical Psychology* 23 (3):331-355.

Williamson, T. (2010) *Philosophy of philosophy*. Oxford Publishing

Williamson, T. (2004) 'Armchair philosophy, metaphysical modality and counterfactual thinking'. *Meesting of the Aristotelian society*, held in Senate house, University of London, Monday 11 October, 2004.

Wettstein, H. (2004). *The magic prism: An essay in the philosophy of language*. Oxford Scholarship Online Monographs.

Wright, C. (2007), 'New age relativism and epistemic possibility: the question of evidence'. *Philosophical Issues*, 17: 262–283.

Appendix 1

Referential Knowledge of the Audience Vignettes Spanish

- *Scenario 1*

Supón que eres un gran fan de la cantante de pop públicamente conocida como 'Madonna'. Siempre has amado su música, has ido a todos sus conciertos, visto sus películas y leído sus libros. Recientemente, incluso, has comenzado a frecuentar un culto de la Kabbalah al que ella pertenece. Allí nadie la llama 'Madonna', pues el culto es estricto y hacen sus propias ceremonias de bautismo y sólo llaman a las personas por el nombre que se les ha dado en la comunidad. A la cantante, por ejemplo, sólo se la conoce como 'Esther'.

Un día, tú y tus amigos están celebrando una fiesta. En una oportunidad que tienes sacas el CD de tu ídolo y pones una canción que te gusta mucho. Tu amigo Jorge, que sabe muy poco de ella (sólo sabe que ella es una cantante de pop, pero en realidad no puede identificar sus canciones y jamás ha escuchado que también se la conoce como 'Esther'), se te acerca diciendo "¡Wow, esta cantante es increíble!".

Unos días más tarde estás reunido con algunos amigos. Todos conocen a las estrella de pop y reconocen algunas de sus canciones, pero no saben detalles de su vida ni mucho menos sobre su identidad en la Kabbalah como 'Esther'. Les estás contando la historia de lo sucedido en la fiesta.

Por favor escoge en la siguiente escala de 1 a 6 qué tan correcto crees que sería decirles:

"Jorge cree que Esther es una cantante increíble"

Totalmente correcto	1	2	3	4	5	6	Totalmente incorrecto

- *Scenario 2*

Supón que eres un gran fan de la cantante de pop públicamente conocida como 'Madonna'. Siempre has amado su música, has ido a todos sus conciertos, visto sus películas y leído sus libros. Recientemente, incluso, has comenzado a frecuentar un culto de la Kabbalah al que ella pertenece. Allí nadie la llama 'Madonna', pues el culto

es estricto y hacen sus propias ceremonias de bautismo y sólo llaman a las personas por el nombre que se les ha dado en la comunidad. A la cantante, por ejemplo, sólo se la conoce como 'Esther'.

Un día, tú y tus amigos están celebrando una fiesta. En una oportunidad que tienes sacas el CD de tu ídolo y pones una canción que te gusta mucho. Tu amigo Jorge, que sabe muy poco de ella (sólo sabe que ella es una cantante de pop, pero en realidad no puede identificar sus canciones y jamás ha escuchado que también se la conoce como 'Esther'), se te acerca diciendo "¡Wow, esta cantante es increíble!".

Unos días más tarde estás reunido con algunos amigos del culto de la Kabbalah. Ellos nunca salen del templo y saben poquísimo sobre el mundo exterior pues son devotos en claustro. Por esta razón sólo conocen a la cantante de pop por el nombre 'Esther' y probablemente nunca han escuchado que ella tenga otro nombre. Les estás contando la historia de lo sucedido en la fiesta.

Por favor escoge en la siguiente escala de 1 a 6 qué tan correcto crees que sería decirles:

"Jorge cree que Esther es una cantante increíble"

Totalmente correcto | 1 | 2 | 3 | 4 | 5 | 6 | **Totalmente incorrecto**

Appendix 2

Stakes vignettes Spanish

- *Scenario 1*

Supón tú y dos de tus amigos, Luisa y Andrés, están en un restaurante japonés. Luisa es experta en comida japonesa y suele frecuentar ese restaurante, por lo cual tú y Andrés están completamente dispuestos a dejar que Luisa ordene los platos y degustar sus recomendados. Andrés especialmente, ha tenido muchas ganas de probar ese restaurante, pues ha escuchado que la comida es deliciosa. Antes de ordenar, Andrés se retira un momento para ir al lavabo, dejándote solo con Luisa en la mesa. Ella comienza a mirar el menú y dice "Oh, aquí preparan un Atún delicioso, deberíamos ordenar para cada uno y probarlo". Acto seguido, Luisa se ausenta por un momento. La camarera, quien conoce a Luisa por ser una cliente regular, quiere tomar la orden sabiendo que tú y Andrés seguirán su consejo. Sólo estás tú en la mesa en ese momento, así que la camarera comienza a pedirte la orden de cada uno y te pregunta: ¿Luisa cree que Andrés debería ordenar el Atún?

Por favor escoge en la siguiente escala de 1 a 6 qué tan correcto crees que sería responderle:

"Luisa cree que Andrés debería ordenar el Atún"

	1	2	3	4	5	6	
Totalmente correcta							**Totalmente incorrecta**

- *Scenario 2*

Supón tú y dos de tus amigos, Luisa y Andrés, están en un restaurante japonés. Luisa es experta en comida japonesa y suele frecuentar ese restaurante, por lo cual tú y Andrés están completamente dispuestos a dejar que Luisa ordene los platos y degustar sus recomendados. Andrés especialmente, ha tenido muchas ganas de probar ese restaurante, pues ha escuchado que la comida es deliciosa. Antes de ordenar, Andrés se retira un momento para ir al lavabo, dejándote solo con Luisa en la mesa. Ella comienza a mirar el menú y dice "Oh, aquí preparan un Atún delicioso, deberíamos ordenar para

147

cada uno y probarlo". Tú sabes- pero Luisa no- que Andrés tiene una alergia severa al Atún y que sólo probarlo podría causarle incluso la muerte.

Luisa se ausenta por un momento de la mesa. La camarera, quien conoce a Luisa por ser una cliente regular, quiere tomar la orden sabiendo que tú y Andrés seguirán su consejo. Sólo estás tú en la mesa en ese momento, así que la camarera comienza a pedirte la orden de cada uno y te pregunta: ¿Luisa cree que Andrés debería ordenar el Atún?

Por favor escoge en la siguiente escala de 1 a 6 qué tan correcto crees que sería responderle:

"Luisa cree que Andrés debería ordenar el Atún"

1	2	3	4	5	6

Totalmente correcta   **Totalmente incorrecta**

Appendix 3

- Instrucciones:

Lee con atención la historia que se te cuenta y responde la pregunta que se encuentra al final haciendo uso de la escala tal y como se te indica. Por favor, hazlo rápidamente y sin darle demasiada reflexión.

Gracias.

Appendix 4

Result tables Referential Knowledge of the Audience and Stakes tests

- **Result tables for the RKotA tests:**

		Mark Scenario 1
L	1	10
i	2	4
k	3	4
e	4	11
r	5	5
t	6	15

Table 2: RKotA Single application Scenario 1

		Mark Scenario 2
L	1	25
i	2	6
k	3	6
e	4	5
r	5	5
t	6	8

Table 3: RKotA Single application Scenario 2

		Mark Scenario 1	Mark Senario 2
L	1	12	35
i	2	3	9
k	3	4	4
e	4	7	3
r	5	7	0
t	6	21	3

Table 4: RKotA Single application Scenario 3

- **Results tables for the stakes tests:**

		Marks Scenario 1
L	1	26
i	2	9
k	3	8
e	4	4
r	5	2
t	6	1

Table 5: Stakes Single application Scenario 2

		Marks Scenario 2	
L	1		28
i	2		6
k	3		4
e	4		1
r	5		3
t	6		10

Table 6: Stakes Single application Scenario 2

		Mark Scenario 1		Mark Scenario 2	
L	1		22		19
i	2		12		3
k	3		6		3
e	4		6		4
r	5		3		9
t	6		7		18

Table 7: Stakes Juxtaposed application Scenarios 1 and 2

Appendix 5

Referential Knowledge of the Audience remake vignettes Spanish

- *Scenario 1:*

Supón que eres amigo de John, un estudiante de primer año de Filosofía Clásica en Nueva York. Un día, estáis comiendo juntos y comenzáis a conversar sobre oratoria y John dice: "Cicerón es el mejor orador que ha existido". Tú sabes que 'Cicerón' y 'Tulio' son dos nombres para el mismo orador, pero en Estados Unidos la gente, incluido John, comúnmente no lo sabe y usan únicamente el nombre 'Cicerón' para referirse al orador romano.

Ahora supón que vas como visitante a otra universidad Americana y estás explicando a algunos estudiantes que conociste allí la conversación que tuviste con John.

Por favor marca en la escala qué tan correcto crees que sería decirle (a) y (b) a tu audiencia:

a.   John cree que Cicerón es el mejor orador que ha existido.

Totalmente correcto
Correcto mas no el mejor
Incorrecto mas no el peor
Totalmente incorrecto

b.   John cree que Tulio es el mejor orador que ha existido.

Totalmente correcto
Correcto mas no el mejor
Incorrecto mas no el peor
Totalmente incorrecto

- *Scenario 2*

Supón que eres amigo de John, un estudiante de primer año de Filosofía Clásica en Nueva York. Un día, estáis comiendo juntos y comenzáis a conversar sobre oratoria y John dice: "Cicerón es el mejor orador que ha existido". Tú sabes que 'Cicerón' y 'Tulio' son dos nombres para el mismo orador, pero en el Reino Unido la gente comúnmente no lo sabe y usan únicamente el nombre 'Tulio' para referirse al orador romano. Por su parte en Estados Unidos pasa justo lo opuesto y las personas, incluyendo a John, normalmente usan sólo el nombre 'Cicerón' para referirse al orador.

Ahora supón que vas como visitante a una universidad Británica y estás explicando a algunos estudiantes que conociste allí la conversación que tuviste con John.

Por favor marca en la escala qué tan correcto crees que sería decirle (a) y (b) a tu audiencia:

a.   John cree que Cicerón es el mejor orador que ha existido.

	Totalmente correcto
	Correcto mas no el mejor
	Correcto mas no el peor
	Totalmente incorrecto

b.   John cree que Tulio es el mejor orador que ha existido.

	Totalmente correcto
	Correcto mas no el mejor
	Correcto mas no el peor
	Totalmente incorrecto

Appendix 6

Result table Referential Knowledge of the Audience-Remake

Mark Scenario 1a	Mark Scenario 1b	Mark Scenario 2a	Mark Scenario 2b
24	3	6	22
2	8	6	3
0	7	8	0
0	8	6	1

Table 8: Referential Knowledge of the Audience-Remake