COOPERATION AND PRAGMATIC INFERENCES

GIULIO DULCINATI

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

UCL, Department of Linguistics

2018

Declaration

I, Giulio Dulcinati, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

The study presented in Chapter 6 was conducted in collaboration with Michael Franke.

The experimental paradigm presented in Chapter 7 was developed in collaboration with Richard Breheny.

Parts of this thesis are currently under review as follows:

Chapter 6: Dulcinati, G., Franke, M., Pouscoulous, N. (under review). *Strategies of deception: under-informativity, uninformativity and lies - misleading with different kinds of implicature*. Topics in Cognitive Science

Acknowledgements

My thanks go to my supervisor, Nausicaa Pouscoulous, who along my research journey has given me guidance, freedom, encouragement and chocolate. All were essential in this project.

Thanks to the people who offered ideas, advice and helpful discussion about my research: Richard Breheny, Diana Mazzarella, Nathan Klinedinst, Napoleon Katsos, Elspeth Wilson, Ye Tian, Wing Yee Chow.

Thanks to the friends in the PhD room who helped me in my PhD in one way or another: Bruno Fernandes, Caterina Paolazzi, Chao Sun, Irini Symeonidou, Louie Pollock, Nick Neasom, Patrick Elliot, Zoë Belk.

Thanks to Saverio, who has been my anchor and my sail.

Abstract

This thesis investigates the role of cooperation for pragmatic inferences. The notion of cooperation that is proposed as relevant for discussing the relationship between cooperation and communication is that of *joint action*. Different theories of communication are reviewed together with the different roles that they assign to cooperation in the context of communication. The study of communication in non-cooperative contexts is used as a way to inform the role of cooperation in communication. Different predictions are derived from Grice's (1989) account and Sperber and Wilson's accounts (Sperber & Wilson, 1986/1995; Sperber et al. 2010) regarding what happens to implicatures in non-cooperative contexts. A series of experimental studies investigates communication in non-cooperative contexts and tests the prediction derived from Grice's account that hearers will not derive implicatures from the utterances of uncooperative speakers. Overall, the results of these studies are not in support of Grice's prediction. They instead support the view that because of a dissociation between comprehension and epistemic acceptance of communicated content (Sperber et al., 2010; Mazzarella, 2015a) uncooperative contexts do not affect the inference of implicatures but only the acceptance of their content. Lastly, this thesis touches on the topic of the *source of relevance* for an utterance, which is treated as a theory neutral notion corresponding to what different theories formalise as the Question Under Discussion (Roberts, 1996/2012) or the *shared purpose* of interlocutors (Grice, 1989). The results of an experimental study on this topic suggest that the exhaustivity of an utterance as an answer to the possible QUDs in a context affects the choice of which QUDs the utterance will be taken to be addressing. Ultimately, this thesis provides initial experimental evidence on how cooperation (or lack thereof) affects pragmatic inferences and puts forward a novel experimental approach to this line of research.

0 Table of Contents

List of figures

List of Tables

1   Introduction and outline

Both communication and cooperation are vague notions. Both occur in simple organisms incapable of sophisticated intentional states such as bacteria as well as in humans, who cooperate and communicate in a complex intentional way. Multiple accounts (e.g., Grice, 1989; Sperber & Wilson, 1986/1995; Clark, 1996) aim to capture the complex intentional communication that happens between humans. Some of these accounts (Grice, 1989; Clark, 1996) cast communication as a cooperative effort which involves interlocutors working together towards *shared goals* by the means of conversation. Some aspects of communication such as *implicatures* have been proposed to depend on an assumption of cooperation (Grice, 1989). Considering the importance that the relationship between communication and cooperation has in these theories, it receives little attention by researchers of human communication. In this thesis I aim to explore the relationship between communication and cooperation both from a theoretical and from an experimental perspective. In doing so I will focus on what happens to communication in non-cooperative contexts, with particular attention to pragmatic inferences. I believe that a better understanding of the relationship between communication and cooperation will further the understanding of the nature and mechanisms of intentional communication itself.

In this thesis I will adopt a notion of cooperation as equivalent to Joint Action, which is a collaborative intentional activity where two agents pursue a *shared goal* together. I will argue (Chapter 2) that this is the appropriate notion of cooperation to address the relationship of cooperation and communication. *Shared goals* are common purposes of agents. Shared goals play a pivotal role in some accounts of communication (Grice, 1989; Clark, 1996). I will propose that different types of goals play different roles in communication and after Attardo (1997) I will argue for the importance of distinguishing *locutionary* goals, which consist in aiming to make communication function, and *perlocutionary goals*, which are goals external to communication such as the goal to find out who ate all the cookies in the pantry or the goal to agree on what movie to see at the cinema. In this thesis I will use the terms *conversation* and *communication* interchangeably and I will not dwell on the relationship between these two notions. I do not take a position on the relationship between the two notions but I will assume that it is valid to compare Grice's account of *conversation* and Sperber and Wilson's account of *communication*.

In the same way that is it not possible to understand the effect of a medicine by observing only patients who are taking that medicine, it is not possible to properly investigate the role of cooperation in communication by considering only communication in cooperative settings. Therefore, I will focus on non-cooperative situations and compare them with cooperative

situations. As communication may involve multiple goals, interlocutors may cooperate on some goals and not on others. If agents are not *locutionarily* cooperative, and so they are not willing to invest the effort to make communication function, no communication can happen between them. In order to study *communication* (i.e., and not its absence) in non-cooperative contexts, I will consider contexts where interlocutors are *locutionarily* cooperative but *perlocutionarily* uncooperative (i.e., unwilling to cooperate with respect to some *perlocutionary goals*). In some of the studies I will present I will use *competitive* settings as a way to operationalise a non-cooperative context.

As to the aspects of communication under investigation, I will focus particularly on implicatures. Implicatures are of particular interest for the relationship between cooperation and communication because according to Grice's (1989) account they are afforded by an assumption of cooperation while according to Relevance Theory they are not. This difference between the two accounts will allow deriving contrasting predictions that can be tested experimentally. Furthermore, I will specifically focus on Quantity implicatures. The reason for this choice is that Quantity implicatures and in particular *scalar* implicatures have been extensively investigated in experimental pragmatics. The existence of previous experimental studies facilitates the construction of experimental items and paradigms and it provides a useful reference point for the availability of these implicatures in an experimental setting.

Non-cooperative contexts often involve a conflict of interest that may lead a speaker to try to *deceive*, that is try to cause the hearer to have a false belief (Mahon, 2007). They can achieve this by *saying* something false, i.e. *lying,* or by communicating a false implicit proposition, i.e. a *false implicature* (Meibauer, 2014). Given the fundamental role that the assumption of cooperation has in Grice's account of how implicatures are derived, I will devote particular attention to what happens to implicatures in non-cooperative contexts. This will be of theoretical interest as *false implicatures* in some particular non-cooperative contexts are incompatible with predictions derived from Grice's account but compatible with Relevance Theory and the account of Epistemic Vigilance (Sperber & Wilson, 1986/1995; Sperber et al. 2010) as I will discuss in Chapter 3. In particular, the separation between the processes of comprehension and epistemic evaluation of the content communicated by the speaker that is proposed by the Epistemic Vigilance account (Sperber et al. 2010; Mazzarella, 2015a) will be important in giving an account of what happens to implicatures in non-cooperative contexts from a Relevance Theoretic perspective.

Since the aim of this chapter is to introduce the topic and scope of this thesis, I will now outline the issues that each chapter addresses.

In Chapter 2 I address the question of finding an appropriate notion of cooperation for discussing its role in human communication. As mentioned, I reach the conclusion that Joint Action offers the appropriate framework. I present the role that cooperation has in Grice's (1989), Clark's (1996) and Sperber and Wilson's (1995; Sperber et al., 2010) theories of communication. In Chapter 3 I address the question of what happens to communication in non-cooperative contexts according to Grice's (1989) account and Relevance Theory (Sperber and Wilson, 1995; Sperber et al., 2010) with particular focus on conversational implicatures. I argue that the prediction that hearers should not infer implicatures from uncooperative speakers can be derived from Grice's account whereas Relevance Theory and the Epistemic Vigilance account predict that hearers should derive implicatures and possibly reject their content. I then address the question of whether false implicatures can be considered lies (Meibauer, 2014). Lastly, I pave the way for the experimental studies presented in the following chapters by reviewing the recent experimental literature on Quantity implicatures.

In Chapter 4 I present a study which addresses the question of how Quantity implicatures are affected in a non-cooperative situation. Participants play the role of the receiver in a competitive signalling game. The study investigates their comprehension of Scalar implicatures and particularised quantity implicatures arising from the utterance they receive from their opponent. The results of this study indicate that hearers do derive implicatures arising from the utterance of an uncooperative speaker and they are more likely to consider false Scalar implicatures to be lies compared to particularised quantity implicatures.

In Chapter 5 I present a study investigating to what extent participants infer scalar implicatures from an uncooperative speaker and to what extent they accept the content of the implicature. Participants read a short story which presents a character as uncooperative (or cooperative) and then reports an utterance of the same character which can give rise to a scalar implicature. The results of this study indicate that hearers are likely to infer the implicatures of uncooperative speaker and then reject the content of the implicature as false, which is consistent with the predictions of the Epistemic Vigilance account (Sperber et al., 2010).

In Chapter 6 I present a study investigating the strategies used by speakers themselves in an uncooperative setting. Participants play the role of the signaller in a competitive signalling game and they complete utterances which serve as hints for their opponents. The results of this study indicate that uncooperative speakers tend to produce more lies and *ostensively* uninformative utterances and they also point to individual differences in their choice of strategy.

In Chapter 7 I present a study investigating whether the *exhaustivity* of the utterance as an answer to the QUD affects the choice of QUD. Participants interpret a non-linguistic utterance

which has different interpretations depending on the source of relevance or QUD (Roberts, 1996/2012) that the participants take the utterance to be addressing. The context of the utterance is manipulated so that the utterance is a more exhaustive to one QUD in some contexts and a more exhaustive answer to a different QUD in other contexts. The results of this study indicates that hearers tend to interpret utterances as addressing the QUD to which they provide exhaustive answers. As QUD can be seen as a formalization of the interlocutors' joint purpose in conversation (see section 3.1.4), this chapter addresses a question that is relevant to the investigation of how conversation works as a cooperative enterprise.

I chapter 8 I draw conclusions from the results of the four studies presented in this thesis and I outline directions for future research that would address open questions about the role of cooperation in communication.

2   Cooperation and communication

In this chapter I set out to find an appropriate notion of cooperation to discuss its relationship with communication. In doing this I will introduce different notions of cooperation and I will focus on the notions of shared intention and joint action.

2.1   What is cooperation?

Cooperation is an abstract term that is encountered in many different contexts. It is used both informally in everyday conversation and technically in the jargon of different disciplines. Cooperation may involve very simple organisms such as bacteria (e.g. in biology) or very complex entities such as firms or governments (e.g. in economics). My subject of interest is the cooperation in the context of human communication and therefore I am interested in the cooperation that happens between humans. The way in which bacteria cooperate is very different from the way humans cooperate. For example, human cooperation is intentional as I will discuss, while bacterial cooperation is not. Different fields (e.g. economics, biology) have different definitions of cooperation that are especially apt for capturing the kinds of phenomena they refer to. Although theories of communication sometimes rely on the notion of cooperation they tend not to define it and instead rely on the reader's intuitive notion of what it means for humans to cooperate. My aim is to determine a notion of cooperation that provides the appropriate conceptual tools for discussing the role of cooperation in intentional communication as described by Grice (1989), Relevance Theory (Sperber & Wilson, 1986/1995) and Clark (1996). The kind of cooperation that is relevant to these theories involves intentional agents, actions (physical or verbal), and goals. This seems to mirror the intuitive notion of cooperation that is described in dictionaries as involving people acting or working together to the same end (Cambridge Dictionary, 2017; OxfordDictionaries.com, 2017). In this section I will look at what theories and definitions best capture this notion. I will reach the conclusion that Joint Action is the right framework for the notion of cooperation following the fundamental assumption of Clark's (1996) theory of communication.

2.1.1   Biological cooperation

Biological definitions of cooperation tend to be broad enough to be able to apply to all forms of life: from bacteria to humans. In biology, cooperative behaviour is normally defined as a behaviour that is beneficial to a recipient (West, Griffin, Gardner, 2007). This definition includes cases of altruism, where the cooperative behaviour is costly to the actor, and mutual benefit, where the cooperative behaviour benefits the actor. This definition of cooperation as a behaviour fits the intuition that cooperation involves action: people *doing* something. However, when this definition is applied to human behaviours it turns out to be too broad and it also

covers instances of behaviour that that do not seem to be cases of cooperation after all. For example, imagine that my neighbour always sets her garden sprinkler in such a position that it regularly waters her lawn as well as mine so that I never have to water my garden. However, she does this accidentally and she never realised that she is also watering my garden. Although my neighbour's behaviour falls under the biological definition of cooperative behaviour in that it benefits me, intuitively this does not seem to count as cooperative behaviour for humans. The reason, as Tuomela (2011, p.69) points out, is that for humans "cooperation must on conceptual grounds be intentional". The neighbour did not intend to benefit me by the means of her behaviour.

Now imagine instead that my neighbour knows that she is watering my garden and she positions her sprinkler this way on purpose. However, I have never realised that someone else is watering my garden and in my ignorance I believe that my garden doesn't need water to thrive. Now the beneficial behaviour is intentional but it still does not seem that my neighbour and I are *cooperating* to keep my garden flourishing. The term cooperation is usually applied to activities that people do collectively and intentionally and in this example there is no 'collectivity'.

Reboul (2017) applies the biological definition of cooperation to human communication and argues that the definition, which in biology refers to costly and beneficial behaviour in terms of evolutionary fitness, can be applied to human communication if the costs and benefits are conceptualised in terms of the interlocutors' interests. Although a biological notions of cooperation based on costs and benefits is general enough that it can apply to intentional communication, of which Reboul provides an example, this notion does not capture the 'intentional' and 'collective' aspects of the intuitive notion that is used in discussing human communication. In what follows I will review some theories that will take us closer to a notion of cooperation that offers the right conceptual tools to discuss cooperation in communication, namely intentional agents and shared goals.

2.1.2 Shared Intentionality

I mentioned that the notion of cooperation applies to human behaviours that are both 'intentional' and 'collective'. It is intentional in the sense that it involves a purpose or an intention to do something. It has a 'collective' aspect in the sense that *all* the agents involved have an intention to do something. However, it is not sufficient for them to all have this intention, they must have this intention *collectively*. For example, imagine that ten people in the same city decide to sing happy birthday, but they all decide to do so independently and sing it in their own homes unaware that anybody else is signing the same song. Even if by chance they sing in perfect synchrony and in tune with each other they are doing so individually and they

cannot be said to be cooperating in this rendition of happy birthday. This is an example of multiple individual intentions to sing happy birthday but not an example of a 'collective' intention. Now imagine instead that the ten people know of themselves as well as of each other that they each intend to sing happy birthday at the same time. In this second case the ten singer are acting and thinking together in a way that is qualitatively different from the simple summation of individual actions or thoughts. Tomasello and Carpenter (2007) call the capacity for this kind of collective intentions *shared intentionality* (Gilbert, 1989; Searle, 1995; Tuomela, 1995) and they propose that this capacity *transforms* behaviour that is also found in other species into distinctively human behaviour. For instance, gaze following (i.e. looking where someone else is looking) is transformed by shared intentionality into joint attention, and group activities such as group hunting are transformed in collaboration or joint action.

Shared intentionality captures both the intentional aspect and the collective nature of the notion of cooperation that I am advocating. However, cooperation involves a behaviour and therefore it presupposes action. In the next section I will present how shared intentions provide the basis for actions to become *joint actions* and what are the core features of this notion of joint action.

2.1.3   Joint action

Joint action presupposes some kind of intention in the same way that individual action presupposes intention (Davidson, 1980). If my arm is moved by another person while I sleep, I did not perform an action because I did not *intend* to move my arm. Joint action presupposes *shared intention* or *we-intention*. The presence of a *shared intention* distinguishes multiple individual intentional actions from *joint action*. I will illustrate this distinction by adapting a famous example from Searle (1990). Imagine that while you are in a park it suddenly starts raining and a number of people around you sprint towards a gazebo in a coordinated and seemingly choreographed fashion. One possibility is that the people you saw were strangers to each other and their 'choreography' happened accidentally and unintentionally when they sought shelter from the rain. A second possibility is that they were performers and their actions were part of a planned choreography. I will refer to these two cases as the *strangers scenario* and the *performers scenario*. In both cases they were acting together, but only in the performers scenario this was a case of *joint action* as it involves the performers having a *shared intention* to perform the choreography.

There are multiple ways of defining shared intentions, and therefore multiple accounts of joint action. However there are three elements that are often considered central to the notion of

shared intention and therefore to joint action: a shared goal, coordination and common knowledge.

The shared goal is a basic feature of both joint action and shared intention. In the performers scenario the performers were involved in joint action because they had a shared goal to perform the choreography while the *strangers* performed the choreography by accident and they could not be said to have a shared goal, even if they had identical individual goals such as the goal to seek shelter from the rain. The peculiarity of a shared intention is that it is an intention towards a goal that an individual cannot accomplish alone. A single performer in the example above cannot have the intention 'I intend to perform the choreography' because she cannot perform the whole choreography on her own. The accomplishment of the shared goal depends on each performer doing their part. For this reason the various definitions of we-intention address the question of who is the subject of the shared intention by making reference to the intentions and beliefs of multiple agents.

Consider this definition of we-intention proposed by Tuomela and Miller (1988):

> *A member A of a collective G we-intends to do X if and only if*
> *(i) A intends to do her part of X*
> *(ii) A believes there are the conditions for success in doing X (namely that a sufficient number of members of G do their parts of X)*
> *(iii) A believes that there is mutual belief in G that there are the conditions for actually doing X*

If this definition is applied to the performers scenario, it is clear that for one the performers to *we-intend* to perform the choreography means that (i) she intends to do her part, which is to run towards the shelter, (ii) she believes that the other performers will also run towards the shelter and therefore the shared goal (i.e. the choreography) will be achieved and (iii) she believes that the other performers believe her and everyone else in the group to believe that the shared goal will be achieved. This definition features a shared goal (i.e. to do X) and an element of common knowledge (i.e. point iii). However, as Searle (1990) pointed out it fails to capture the feature of coordination which seems an essential feature of shared intention and joint action. In Clark's words (1996, p. 59): "What makes an action a joint one, ultimately, is the coordination of individual actions by two or more people. There is coordination of both *content*, what the participants intend to do, and *process*, the physical and mental systems they recruit in carrying out those actions." Searle provides the example of a group of graduates of a business school who all believe in Adam Smith's theory that they will help humanity by pursuing their own interests. Each graduate has the intention to help humanity by pursuing self-interest and they all have mutual belief to the effect and the success of this intention. Although this case fits Tuomela and Miller's definition, the business graduates pursue their shared goal by

individualistic means and not by coordinating their actions. Therefore Searle argues that they cannot be said to have a we-intention.

Searle (1990) also criticised Tuomela and Miller's account as reductionist in the sense that it does not present the *we-intention* as qualitatively different from individual intentions and instead it reduces it to the sum of individual intentions (i.e. A intends to do her part) and beliefs (i.e. A believes that there are conditions for success and mutual belief). Searle argued that the we-intention is indeed an intention held by an individual, as opposed to an intention of a group construed as *collective* agent as in Pettit's (2003) proposal, but it is a different *type* of intentional attitude than ordinary individual intentions, a *primitive* that cannot be analysed in terms of individual intentions or beliefs.

Bratman (1992, 1993) proposed an account of we-intention and joint action which introduces the idea of coordination, which lacked in Tuomela and Miller's (1988) account. Coordination takes the form of *interlocking* intentions and plans of the participants. In Bratman's account the agents in a shared intention have an intention of the form '*I intend that we perform the choreography'*. The use of a propositional intention or aim intention, where the content of the intention is a proposition ('we perform the choreography') instead of action intentions (i.e. the content of the intention is an action: 'I intend *to run*') solves the problem of intending things that are not our own actions. Bratman's (1993, p.106) definition of shared intention (i.e. we-intention) is as follows:

> *We intend to J if and only if*
> *1. (a) I intend that we J and (b) you intend that we J*
> *2. I intend that we J in accordance with and because of 1a, 1b and meshing subplans of 1a and 1b; you intend that we J in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b.*
> *3. 1 and 2 are common knowledge between us.*

To paraphrase, Bratman defines *shared intention* as involving (1) a shared goal, (2) an intention to coordinate and (3) mutual knowledge of the shared goal and the intention to coordinate. Coordination is stipulated in terms of meshing plans: participants in joint action have to plan their participatory action and they must pursue the shared goal with plans that are *co-realizable* (i.e. meshing). For example, if a friend and I intend to paint a house together but my friend intends that we paint it all red and I intend that we paint it all blue, our plans do not mesh because we cannot carry them out both at the same time. Although coordination happens by interlocking plans of the agents, the agents do not need to represent all the plans and how they interlock in advance of their action. Point 2 of the definition simply requires that participants *intend* to make their plans mesh. Indeed, in long term joint projects agents often do not have a

complete plan of their contribution, but they can nevertheless have a shared intention towards its completion.

Although Bratman postulates agents coordinating on *plans*, other authors have argued that joint action should be analysed as involving coordination at different levels of complexity. Pacherie (2012) argued that coordination happens at three different levels of intentions, which I will illustrate using again the performers scenario. At the most abstract and complex level, agents coordinate on their *shared distal intentions*, which are specified in *cognitive* terms of goals and means such as the overall goal to perform a choreography. At a lower level agents coordinate their *shared proximal intentions*, which are specified in terms of action-schemas and perceptual effects, such as the action of running towards the gazebo. At the lowest level agents coordinate their *coupled motor intentions,* which are specified in sensorimotor terms, such as the actual body movements involved in the act of running towards the gazebo. Pacherie argues that Bratman's plans refer only to the highest level of complexity, distal intentions, and that these need to be implemented into actions or proximal intentions and that actions need to be implemented into physical movements. Although Bratman does not directly address these lower levels of coordination, in giving an account of *shared cooperative activities*, which is *a* form of joint actions, he describes it as involving a shared intention and "mutual responsiveness of intention and in action*"* (Bratman, 1992, p.339), where *mutual responsiveness in action* may be interpreted as going in the direction proposed by Pacherie.

Knoblich, Butterfill and Sebanz (2011) propose a two level-analysis of the kind of coordination involved in joint action in the same spirit as Pacherie's proposal. At the more complex and voluntary level, *planned coordination* involves the participants planning their own action in relation to their representation of the shared goal, being aware that other participants will contribute and possibly also representing the plans and participatory actions that other participants will contribute. This kind of coordination involves participants representing other participants tasks (Sebanz, Knoblich & Prinz, 2005), what is jointly perceived by the participants and how another participant's perceptual access differs from one's own (Brennan & Hanna, 2009). At a lower subpersonal level, *emergent coordination* is spontaneous coordination that emerges in social interaction even when individuals are not engaged in any joint action. Emergent coordination is based on perception-action coupling and it has been studied in terms of entrainment, synchrony of movements (e.g. Shockley , Richardson & Dale, 2009) and mimicry (e.g. Chartrand & Bargh, 1999). These effects seem arise automatically and involuntarily. Knoblich and colleagues (2011, p.91) propose that "emergent coordination is likely the key to dealing with the real-time aspects of joint action". The mechanisms of emergent coordination are limited and not flexible and interactive as the mechanisms that support planned

coordination, but their automaticity makes them able to cope with the small timescale of joint actions. While emergent cooperation does not entail any shared representations, planned coordination requires that participants share some information about their shared goal, their participatory actions or their perceptual situation. This last aspect of planned coordination ties to the third feature of joint action besides the shared goal and coordination: common knowledge.

All the definitions of shared intention considered so far require participants to have mutual beliefs or common knowledge (Tuomela & Miller, 1988; Bratman, 1993). A proposition *p* is common knowledge (Lewis, 1969) for two agents A and B, or two agents can be said to have mutual belief of a proposition *p* if and only if: (i) A and B believe that *p*; (ii) A and B believe that *i*. However, as Tollefsen (2005) points out, the requirement of common knowledge would suggest that young children (before the 4th year of age), who do not have a robust theory of mind, should not able engage in joint action. As a matter of fact, children do engage in joint activities that are in many respects equivalent to joint actions in adults such as pretence play (Rakoczy & Tomasello, 2006; Rakoczy, 2008). Moreover, Sperber & Wilson (1986/1995) have argued that common knowledge is also psychologically implausible in adults because it leads to an infinite regress of intentional states (e.g. I know that you know that I know…). Therefore, in order to account for joint action with children (and with adults) common knowledge should probably be substituted with a leaner requirement. Sperber and Wilson (1986/1995) offer a solution to this problem by proposing to substitute common knowledge with *mutual manifestness*. A fact is mutually manifest to an individual if they are able to represent it and accept it as true or likely true. The set of facts that are manifest to an individual forms their *cognitive environment*, in the same way that the set of objects that is visible to an individual forms their visual environment. A *mutual cognitive environment* of two individuals is the set of facts that are manifest to both individuals and for which the fact that they are manifest to both is also manifest. All the facts in a *mutual cognitive environment* are *mutually manifest* to the individuals who share the *mutual cognitive environment*. As Sperber and Wilson argue, mutual manifestness is psychologically more plausible than common knowledge. However, in this thesis I will not focus on the benefits of one notion over the other and instead I use one notion or the other depending on the account of joint action or communication that I am discussing. This is because these notions play analogous roles in the accounts in which they are used, and the contentious differences between them are not directly relevant to the scope of this thesis.

In this section I reviewed three core features of shared intention which are therefore also central features of joint action: a shared goal, coordination, and common knowledge. Besides these feature that joint action has in virtue of being based on a shared intention, joint action has been

proposed to have the additional features of *commitment* and *mutual assistance*. In the next section I will review Bratman's (1992) account of *Shared Cooperative Activities* (SCA), which are a form of joint action, to illustrate these two features.

2.1.4   Commitment and Mutual assistance

Bratman (1992) argues that SCA share three general distinctive features that are not definitional but that seem to characterize SCA in general: (i) *Mutual responsiveness of intentions and actions* (i.e. coordination) (ii) *Commitment to the joint activity* (even if, ultimately, agents have different motives) (iii) *Commitment to mutual support* (i.e. agents will *help* each other in the pursuit of the shared goal). Since I discussed the notion of mutual responsiveness in the context of coordination I will focus on the other two features.

Commitment to the joint activity can be interpreted as the obligation for participants not to *abandon* the joint activity while it is in progress. This feature also echoes other accounts which developed the idea of *joint commitment* as a central element to joint action (Gilbert, 2006; Roth, 2004). Joint commitment in these accounts is a commitment of multiple people that is not reducible to individual commitments and that has a 'normative' flavour. It involves *obligations* and *rights* for the agents involved in the joint commitment. Another way in which participants can be *committed to the joint activity* that is more specific to Bratman's framework is that participants must be willing to forsake their *non-meshing plans* so that the shared goal can be accomplished. For example, if my friend and I want to paint a house all in one colour together but I intend to use blue paint and my friend intends to use red paint our individual goals do not *mesh*. We must be willing to abandon these individual plans or we will not be able to achieve our shared goal.

The requirement of mutual support means that participants in joint action must be willing to help each other at least in some circumstances. Bratman gives the examples of two *unhelpful singers* who have to sing a duet. They have a shared intention to sing the duet and to coordinate in singing it, but should one of them stumble on his notes, the other would not be willing to help him under any circumstances. In fact each singer wants the other to fail. Bratman argues that this would be a case of shared intention but not a shared cooperative activity. He argues that participants in a shared cooperative activity must be willing to provide mutual assistance with respect to their participative actions.

While commitment of the participants is a central feature of joint action under multiple accounts, the feature of *mutual assistance* is not present in all instances of joint actions and it has been proposed as a factor that allows to distinguish between *cooperative* joint action and *non-cooperative* joint action. Tuomela (1993) argued that joint actions can be cooperative or

non-cooperative. Cooperative joint actions are those joint actions where participants can help each other in the performance of their participatory actions, because all the participatory actions benefit all the participants (e.g. if you and I are painting a house together, I could help you paint your part). Non-cooperative joint actions are those where it is not possible or *prudent* to help other participants perform their participatory actions. The philosophical literature on joint action has largely focused on *cooperative joint actions* (Pacherie, 2012), however this does not mean that all joint actions are cooperative. The typical example of non-cooperative joint actions are competitive games. If a friend and I are playing a game of chess, I cannot suggest my friend's moves without going against my interest. Competitive games and economic exchanges are instances of such non-cooperative joint action. They nevertheless maintain a 'joint action base' in terms of coordination of the participants and their shared goal to perform the action. Therefore, the feature of *mutual assistance* does not hold for all joint actions, but only for cooperative ones.

In conclusion, cooperation is a vague term that encompasses different notions. The kind of cooperation that is relevant to intentional communication involves intentional agents working towards shared goals. Joint action provides the right conceptual tools to capture this notion. In the next chapter I will firstly review theories of intentional communication. The core features of joint action reviewed in this section will be instrumental in discussing that I have reviewed will be instrumental in discussing the notion of cooperation that these theories presuppose and what role they assign to it in the context of communication. Secondly I will discuss communication in non-cooperative contexts and the predictions that theories of communication make for this scenario.

3 Cooperation in communication

In the previous chapter I have discussed the notion of cooperation and I have argued that the notion of joint action offers the appropriate conceptual tools to discuss the role of cooperation in communication. In this chapter I will firstly introduce Grice's (1989), Clark's (1996) and Sperber and Wilson's (1986/1995) accounts of communication and the role that the notion of cooperation plays in these accounts. Secondly, I will address the question of what happens to communication and in particular conversational implicatures in cooperative contexts. In addressing this question I derive predictions from Grice's account and Relevance Theory (Sperber & Wilson, 1986/1995; Sperber et al. 2010) that will be important for the empirical studies presented in the following chapters.

3.1 Theories of communication and the role of cooperation

Like cooperation, communication is a vague notion. In its widest sense, the term communication could be applied to any transmission of information, be it between humans, animals, bacteria or machines. However, in this thesis I focus on a specific kind of communication: the intentional overt communication that happens between humans. I will review some accounts of intentional communication and I will review what role these theories assign to cooperation and what notion of cooperation they adopt. In particular, the focus of this thesis is on pragmatic inferences and their relationship with cooperation. In this regard I will pay special attention to Grice (1989), who attributed explicitly a fundamental role to cooperation in his account of conversation and conversational implicature.

3.1.1 Grice

3.1.1.1 *Non-Natural Meaning*

We can identify Grice's notion of communication (in a narrow sense) with his definition of Non-Natural meaning. "'A meant$_{NN}$ something by *x*' is (roughly) equivalent to 'A intended the utterance of *x* to produce some effect in an audience by means of the recognition of this intention'" (Grice, 1957, p.385). This *meaning-intention* is a reflexive intention (Récanati, 1979; Clark, 1996) because it makes reference to itself. Clark (1996, p. 130) rephrases Grice's meaning intention in a way that makes its reflexivity more evident: "In presenting [sentence] *s* to audience A, a speaker S means for A that *p* if and only if: (*i*) S intends in presenting *s* to A that A recognize that *p* in part by recognizing that *i*." Some commentators of Grice (Levinson, 1983; Sperber & Wilson, 1995) have analysed his meaning-intention as involving three separate intentions:

(i)   the intention to convey a piece of information (i.e. *informative intention*)

(ii)   the intention that the audience recognizes this informative intention (i.e. *communicative intention*)

(iii)   the intention that the basic layer of information (the informative intention) should not be available without the recognition of the communicative intention.

As a consequence, successful communication involves the communicative intention (ii) becoming mutual knowledge between communicator and audience (Levinson, 1983). This last requirement of mutual knowledge is implicitly present when the communicative intention is expressed in its reflexive form.

I will borrow and extend an example from Récanati (1979, p.176) to illustrate what falls under the definition of non-natural meaning and what does not. Imagine that I am playing poker and I have a good hand of cards. In this situation I can raise the stakes to take advantage of my good hand and make my victory more profitable. Therefore, my raising the stakes will *mean* to the other players that I have a good hand. This however is not an example of non-natural meaning because, assuming that the other players' inference is only a side effect of my strategy, I did not intend to communicate to the other players that I have a good hand. I had no informative intention and no communicative intention.

Now, imagine that I do not have a good hand but I still raise the stakes because I want to make the other players believe that I do without realizing that I intend them to think so. In this case I *intend* the other players to think that I have a good hand (i.e. I have an informative intention) but I want this intention to remain secret. In this case there is no communicative intention – more precisely it is intentionally hidden. Therefore, this is still not a case of non-natural meaning.

Let's imagine a third scenario: I do have a good hand and I want the other players to see it so I simply turn over my cards and show them to the other players. Here I have both an informative intention (for them to know that I have a good hand) and a communicative intention (for them to realize that I *intend* them to believe that I have a good hand). The other players would believe that I have a good hand because this fact is self-evident when I show my cards. They do not *need* to recognize that I *intend* them to think that I have a good hand in order to believe it. For this reason this also not a case of non-natural meaning because in non-natural meaning the recognition of the communicative intention is *necessary* for the recovery of the informative intention.

An example of non-natural meaning would be if I pointed to my cards with a smug smile without revealing them to the other players. In this case the other players cannot see *for themselves* that I have a good hand, they can only recover the content this information if they realise that I am trying to communicate it to them, otherwise they would just wonder why I am smiling smugly and pointing to my cards. Another, more conventional example of non-natural meaning would be for me to tell the other players "I have a good hand".

3.1.1.2 Cooperation in Grice and conversational implicatures

Grice sees conversation as a cooperative activity to which interlocutors contribute, he makes analogies with baking a cake together or mending a car together. He captures the way in which interlocutors do (and *should*) contribute to conversation by postulating a *Cooperative Principle* (CP): "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (Grice, 1989, p.26). He also postulates corollaries to this principle, the maxims (Grice, 1989, p.26-27):

Quantity
> *1. Make your contribution as informative as is required (for the current purposes of the exchange).*
> *2. Do not make your contribution more informative than is required.*

Quality: Try to make your contribution one that is true
> *1. Do not say what you believe to be false.*
> *2. Do not say that for which you lack adequate evidence.*

Relation: Be relevant

Manner: Supermaxim: Be perspicuous
> *1. Avoid obscurity of expression.*
> *2. Avoid ambiguity.*
> *3. Be brief (avoid unnecessary prolixity).*
> *4. Be orderly.*

According to Grice, some aspects of communication depend on the assumption of cooperation. In this thesis I will focus on conversational implicatures. The assumption of cooperation in Grice's theory applies to all aspects of meaning$_{NN}$ including communication of linguistically encoded content. However only for implicatures does Grice explicitly claim that this assumption has a role in the interpretation process. Grice does not explicitly say that the assumption of cooperation is necessary for decoding linguistic material or for context-depended elements of *what is said* (i.e. disambiguation and reference assignment). From the speaker's perspective, the CP and maxims are *norms* that guide how speakers contribute to the conversation (i.e. what they say). From the hearer's perspective, these norms are standards on which hearers base their expectations about how the speaker will contribute. Because knowledge of the maxims is mutually assumed by speaker and hearer, these norms can be exploited in order to convey more information than what the speaker actually encodes into words. This assumption is the foundation of Grice's theory of conversational implicature.

3.1.1.3 Quantity implicatures

Conversational implicatures are propositions that the speaker can communicate and the hearer recover by overt exploitation, violation or flouting of the maxims. In this thesis I will focus on Quantity implicatures, implicatures communicated by the overt violation of the first maxim of quantity. Imagine that while a friend of mine is making the shopping list I tell her that *John used some of the shampoo*. Since my friend would need to buy shampoo if there were none left, both she and I know that it would be relevant and more informative for her to know whether *John used all of the shampoo*. Unless she has some reason to believe that I am not being cooperative, she can assume that I am trying to make my utterance informative enough for the current purpose of our conversation, which may be to exchange information about what needs to be bought. Therefore she can infer that there is some other reason why I have not uttered the more informative statement that *John used all of the shampoo*, and this reason must be that I do not believe that this alternative statement is true. Assuming that I know whether John used all of the shampoo or not, she can conclude that I believe that *John did not use all of the shampoo*, which is the implicaure that I am communicating.

All conversational implicatures, according to Grice, must be able to be *worked out* with an explicit argument rather than simply being intuitively accessible. The first step in calculating a quantity implicature is what the speaker *said* or the truth-conditional meaning of the speaker's utterance, which according to Grice is arrived at by decoding the linguistic material, resolving the ambiguities contained in the utterance and assigning referents to indexicals. For example, to understand what the speaker said by uttering that *John used some of the shampoo*, the hearer needs to assign referents to *John* and *the shampoo*. The second step is for the speaker to recognise that there are relevant and more informative alternative statements that the speaker could have made. In these example these *alternatives* include the statement *John used all of the shampoo*. The following step is for the hearer to find a reason for why the speaker did not utter the relevant and more informative utterance which preserves the assumption that the speaker is being cooperative. In the cases of quantity implicatures this reason is a clash with the first maxim of quality. The speaker did not say that *John used all of the shampoo* because they did not want to say something false or something for which they *lacked adequate evidence*. At this point, if the hearer cannot make the assumption that the speaker is *competent* about whether the alternative statement is false they can derive an *ignorance inference* (Geurts, 2010). For example, my hearer could infer that I am communicating that I don't know whether John used all of the shampoo. If the hearer instead assumes that the speaker is competent they will make what is sometimes called the *epistemic step* (Sauerland, 2004; Breheny, Ferguson & Katsos, 2013) by which they conclude that the speaker believes the alternative statement to be false.

For example, if my hearer assumes that I am competent about John's shampoo consumption she can conclude that I believe that he did not use all of the shampoo. In the last step the hearer assumes that the speaker intended them to reason this way and they are communicating the quantity implicature that *John did not use all of the shampoo*.

The implicature in the example I used above is a *Generalised conversational implicature* (GCI) as it is tied to a particular expression which *generally* carries that implicature. Specifically, 'some' gives rise to the 'not all' implicature in most contexts. GCIs contrast with *Particularised conversational implicatures* (PCI) which are implicatures that arise because of the *special features of the context*. For example, imagine that a friend asks me about my two friends Ann and Beth and whether they are still playing the piano; and I reply that *Ann is still playing*. If my friend believes that I am competent about whether Ann and Beth play the piano, she can infer the implicature that Beth is not playing anymore. This is because I could have uttered the alternative statement that *both Ann and Beth are still playing*, but I did not. Although Grice (1989, p.37) introduced the distinction between GCIs and PCIs he does not discuss it in detail. The notion of GCIs was developed more in detail by neo-Gricean theories such as those proposed by Horn (1972, 1989) and Levinson (2000).

Horn (1972, 1989) expanded Grice's account of quantity implicatures by introducing the notion of semantic scales. Semantic scales are arrays of linguistic alternatives belonging to the same grammatical category which can be ordered by informativity (Levinson, 1983). For example, <all, most, many, some> is a scale where 'all' is the most informative element and 'some' the least informative. Semantic scales are ordered by entailment in the sense that a statement containing a member of the scale entails an equivalent statement containing a lower-ranking scale mate. For example, the statement that *John used all of the shampoo* entails the statement that *John used some of the shampoo*. When a low-ranking element of a scale is used (e.g. *some*) it can give rise to scalar implicature by negating higher-ranking scale mates that could have been used instead (e.g. not *all*). Particularised quantity implicatures are sometimes called Ad hoc scalar implicatures (Hirshberg, 1985) as they can also be described using Horn's semantic scales. However, while the scalar implicatures use *lexical scales* which are tied to a particular lexical item and are independent from the context of use (e.g., <free, cheap>, <will, may>, <succeed, try>, <always, sometimes>, <impossible, difficult>), Ad hoc implicatures use Ad hoc scales that are specific to the particular context of the utterance. In the example above, where saying that *Ann is still playing* generates the implicature that *Beth is not playing anymore*, the implicature may be said to arise because the lower element of the informativity scale <*Both Ann and Beth, Ann*> has been used and this motivated the inference that the same statement with the most informative element is false (i.e. it is not the case that *Both Ann and Beth are still playing*).

Horn (2004) assumes that scalar implicatures, like other GCIs, arise whenever the element of a lexical scale is used unless they are blocked by the context. However, like Grice, he assumes that scalar implicatures are calculated and that they are not mandatory (Horn, 2005). This means that in contexts where they are not supported these implicatures are not calculated at all. And similarly to Grice, Horn assumed that it is not possible to have a scalar implicature that is not *intended* by the speaker. This is not the case for Levinson (2000) who proposes still another development of Grice's notion of GCIs.

According to Levinson (2000) GCIs are default *mandatory* inferences which arise whenever a particular lexical item or trigger is used. For example, the quantifier *some* automatically triggers the implicature *not all*. These implicatures arise in all contexts and are then cancelled by the hearer if they are inconsistent with the context. Unlike Grice's and Horn's GCIs, Levinson's GCIs are not calculated, they are mandated by a set of inferential heuristics. For example, scalar implicatures are mandated by the Q-heuristic which states that *what is not said is not the case*. So if the speaker says that *John used some of the shampoo* and does not say that *John used all of the shampoo*, the latter is not the case. Levinson's GCI's have the peculiarity that they arise independently of whether the speaker intended to communicate them.

Other theoretical positions maintain that the neo-Gricean GCI-PCI distinction is misguided and that all conversational implicatures are *nonce* inferences that are mandated by the particular context of the utterance and that they are never automatic or mandatory (Sperber & Wilson, 1986/1995; Carston, 2002; Geurts, 2010, i.a.). In addition, the neo-Gricean account of the GCI-PCI distinction, at least in Levinson's (2000) sense of the distinction, has been disproven by experimental evidence which I will partly review in section 3.3. I do not commit to a particular theoretical position regarding the GCI-PCI distinction in this thesis and I will be using these terms strictly in Grice's sense. However, I am sympathetic to the view that a *typical* interpretation may be available for lexical scalar expressions such as quantifiers (Newstead & Collis, 1987; Geurts & van Tiel, 2013; van Tiel, 2014) while this would not be the case FOR expressions that give rise to Ad hoc implicatures, which are by definition specific to a particular context and presumably are not used often enough to be associated with a *typical* interpretation.

3.1.1.4   *Grice's idea of cooperation: a shared goal*

Grice sees *talking* as a *purposive* and *rational* behaviour. Purposive because interlocutors are pursuing a shared goal by the means of conversation, and rational because they do so in a rational way.

> *Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least,*

cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or a set of purposes, or at least a mutually accepted direction. (Grice, 1989, p.26)

In Grice's account, it is not the fact that interlocutors are helpful to each other that makes conversation an instance of cooperation, it is the fact that conversation is the pursuit of a *shared goal*. Grice was concerned with rationality rather than helpfulness (Davies, 2007). In fact, Grice's CP and maxims are not indications on how to be 'helpful' in conversation but instructions on how to be rational in pursuing a shared goal or how to perform one's participative action in the shared enterprise. Grice's account does not provide instructions to "help your interlocutor" or "give the information that your interlocutors needs". Grice's view is that contributions to the conversation are in pursuit of the shared goal rather than just for the benefit of the audience and therefore "The talker who is irrelevant or obscure has primarily let down not his audience but himself" (Grice, 1989, p.29).

3.1.1.5   *Grice's cooperation and joint action*

Grice presents conversation as cooperative enterprise because it involves a shared goal of participants. This suggests a parallelism between his account of conversation and joint action. The parallelism is even clearer where Grice argues that conversation belongs to a larger category of *cooperative transactions* which have three distinguishing features (Grice, 1989; p. 29): (i) "*Participants have some common immediate aim*", even if they ultimately have different motives or goals, (ii) "The contributions of the participants should be dovetailed, mutually dependent" and (iii) the transaction should continue until both parties agree to terminate it, "You don't just shove off or start doing something else". These features match exactly those of joint action discussed above: (i) a shared goal, (ii) coordination among participants or, as Bratman (1992) puts it, a requirement that participants pursue the shared goal in accordance with *meshing plans*, and (iii) commitment to the joint activity. Although the notion of commitment has been in some cases analysed as more complex than what Grice proposes (Gilbert, 2006; Roth, 2004), its essence is that participants should not abandon the joint activity while in progress and therefore it can be argued to match the feature proposed by Grice.

Besides these parallels, joint action and Grice's account of conversation also share an assumption of common knowledge. Grice's definition of Non Natural Meaning, which includes implicatures, presupposes that the meaning intention of the speaker is common knowledge between speaker and hearer as it applies to *overt* communication. Grice also talks of a *mutually accepted direction* of the talk exchange and his whole account of conversational implicatures rests on the assumption that interlocutors share the knowledge of the shared goal and of their intentions to perform their participative actions. The maxims and the CP are instructions or descriptions of how speakers should perform their participative actions in conversation. If the

interlocutors did not have mutual knowledge that the speaker intends to perform their participative action (i.e. abide by the CP and maxims) it would not be possible for the speaker to communicate implicatures or for the hearer to calculate them. In the example I used above where I say that *John used some of the shampoo,* it would not be possible for me to communicate the implicature that *John did not use all of it* if I did not assume that my hearer expects me to be informative and truthful; and it would not be possible for her to calculate this implicature if she did not assume that I was trying to be informative and truthful. Similarly, interlocutors must have mutual knowledge of the purpose of the conversation in order to communicate and infer implicature. In the *shampoo* example, if my friend and I did not have a mutual assumption that my utterance is relevant to the goal of deciding what goes on the shopping list, we would not be aware that I could have said something relevant and more informative (i.e. that *John used all of the shampoo*), and so we would not have been able to communicate and infer the implicature.

The fact that the fundamental features of what Grice calls the category of *cooperative transactions*, under which he classifies conversation, match those of joint action indicates that joint action embodies the notion of cooperation that Grice adopts in his account. Since Grice *identifies* conversation as an instance of cooperative transaction, the features presented above can be taken as constitutive of conversation. In other words, a minimal notion of cooperation for Grice has the features of a shared goal, coordination, commitment and common knowledge.

3.1.1.6   *Grice's conversation is a cooperative joint action?*

I discussed the proposal by Tuomela (1993) that joint actions can be divided in *cooperative joint actions* where it is beneficial to help others with their participative action and *non-cooperative joint actions* such as competitive games or economic transactions where other participants have conflicting subgoals that exclude mutual assistance. Since Grice seems to draw an equivalence between joint action and conversation, this distinction should also apply to conversation. The question then is whether Grice saw conversation as an 'all cooperative' phenomenon, a cooperative joint action. I cannot find a direct answer to this question in Grice's writing, however the following passage suggests that he did not have an 'all cooperative' view of conversation:

> *"we should recognize that within the dimension of voluntary exchanges (which are all that concern us) collaboration in achieving exchange of information or the institution of decisions may coexist with a high degree of reserve, hostility, and chicanery and with a high degree of diversity in the motivations underlying quite meagre common objectives."*
> (1989, p.369)

Therefore, for Grice *mutual assistance* can be a feature of conversation but it is not a necessary one. Although conversation necessarily involves a shared goal, it is not 'all cooperative' and it is also compatible with the presentence of conflicting goals.

3.1.1.7   *Grice's purposes of conversation: Locutionary and Perlocutionary*

Grice makes multiple references to the idea of a shared goal, however he gives very little information as to what may be the nature of this goal. It seems that he is not sure himself that he has defined this goal well enough:

> *I have stated my maxims as if this purpose were a maximally effective exchange of information; this specification is, of course, too narrow, and the scheme needs to be generalized to allow for such general purposes as influencing or directing the actions of others.* (1989, p. 28)

Attardo (1997) has argued that the notion of *purpose* of the conversation in Grice's Cooperative Principle is ambiguous: *"Is the purpose that of the talk (i.e. getting the conversation to function), or that of the speakers (i.e. allowing the speakers to achieve whatever goals they have, outside of the conversation)?"* (1997, p.759). He argues that Grice uses both of these interpretations but he does not make a distinction between them. He therefore tries to distinguish these two interpretations by positing two levels of cooperation between interlocutors:

- *Locutionary Cooperation* concerns the goal of *making conversation function*. This consists in the interlocutors cooperating to *maximize the transmission of information*.
- *Perlocutionary Cooperation* concerns participants cooperating to pursue goal outside conversation such as filling one's car tank.

Attardo argues that although Grice explicitly claims that the purpose of conversation is the exchange of information, many pragmatic inferences are based on the perlocutionary goals. For example, consider the following example from Grice (1989, p.32):

> A: I'm out of petrol.
> B: There is a garage round the corner.

Here B can be taken to imply that the garage is open and selling petrol. However, this implicature can only be derived if the current purpose of the conversation, to which B's utterance should be relevant, is to help A fill their tank. Thus, according to Attardo many of Grice's own examples show that implicature generation is based on perlocutionary goals.

Attardo proposes to resolve this ambiguity in Grice's account by proposing two cooperative principles that operate at the same time: a Locutionary Cooperation Principle, which is Grice's own CP, binging participants to a maximally efficient exchange of information; and a

Perlocutionary Cooperation Principle (1997, p. 766), which he formulates as a principle with three corollaries:

> *The Perlocutionary Cooperation Principle: Cooperate in whatever goals the speaker may have in initiating a conversational exchange, including any non-linguistic, practical goals.*
> 1. *If someone needs or wants something, give it to them.*
> 2. *If someone is doing something, help out.*
> 3. *Anticipate people's needs, i.e. provide them with what they need, even if they do not know that they need it.*

According to Attardo the Perlocutionary Cooperation Principle has a wider scope than the Locutionary Cooperation Principle and takes precedence over it. This means that if the speaker can ignore the cooperative principle (e.g. not provide the information requested) if she can cater to her interlocutor's perlocutionary needs (e.g. give them information they need but that they have not asked).

Yet Attardo's analysis faces one important problem deriving from the fact that he does not draw a distinction between *shared* perlocutionary goals, which are *mutually accepted* by interlocutors, and the interlocutors' *individual* perlocutionary goals. Grice's Cooperative Principle and maxims only apply to a purpose *after* the purpose has been accepted, and not on the decision to accept it or modify it. Grice repeatedly says that the purpose of the conversation is *common* or *mutually accepted* and so the CP describes how this shared goal is pursued, not how interlocutors settle on a shared goal. Attardo's Perlocutionary Cooperation Principle instead prescribes that speakers take on any individual goals that their interlocutors may have. As I argued, Grice's idea of conversation as a cooperative activity is linked to participants having a shared goal, and not to the fact that they should be helpful and altruistic with each other. In fact, Grice explicitly acknowledges that conversation often involves *hostility* and conflicting motivations. Although most cultures have social norms that prescribe altruistic behaviour such as being helpful and attending to other people's individual goals, which apply within conversation, it was not Grice's intention to give an account of these norms. It is true that most of Grice's examples involve interlocutors helping each other, but this is most likely due to the fact that examples where an individual is helping another are the easiest examples involving a shared goal. In sum, while Grice wants to account for the way interlocutors rationally pursue a shared goal through conversation; Attardo's Perlocutionary Cooperation Principle accounts for the way speakers accept their interlocutors' individual goals and in doing so, make them shared goals.

Nonetheless, Attardo makes two important points in his analysis. Firstly, he recognizes that there is a distinction between goals that are *internal* to conversation (i.e. the goal to exchange information in a maximally efficient way) and goals that involve specific activities in the world

such as a person filling their tank, or *perlocutionary* goals. Secondly, he points out that the shared goals that afford implicatures are often *perlocutionary* ones.

I will adopt the *locutionary-perlocutionary* distinction but I will cast it in slightly different terms than Attardo's. From this point onwards I will use the term *locutionary* cooperation to refer to the bare bones cooperation between speaker and hearer that is necessary for conversation to function: the speaker must be willing to invest some effort in producing utterances and the hearer must be willing to invest some effort in interpreting them. In other words, *locutionary* cooperation is related to the discharging and recognition of the meaning-intention (non-natural meaning). I will use *perlocutionary* cooperation, *on the other hand,* to refer to cooperation regarding any shared goals that go beyond the mere functioning of conversation.

In the same vein, Scott-Phillips (2014) proposed a distinction between three types of cooperation: *communicative* cooperation, *informative* cooperation and *material* cooperation. Communicative cooperation is very close to *locutionary* cooperation as it refers to willingness to follow linguistic conventions and take part in conversation. *Informative* cooperation refers to cooperation in providing honest (instead of deceitful) information. *Material* cooperation refers to cooperation in aiming to achieve *pro-social* goals through conversation, which is equivalent to *perlocutionary* cooperation. Scott-Phillips proposes that Grice's cooperation corresponds to what he calls *communicative* cooperation. However, as Attardo points out, Grice's own examples strongly suggest that the *purposes* mentioned in the CP and that among other things afford implicatures, are or can be *perlocutionary* or *material* goals. In Grice's account what Scott-Phillips dubs *informative* cooperation is simply a consequence of applying the maxims of Quality, and therefore it is part of doing one's part in the pursuit of the shared goal in a rational way. Furthermore, being truthful in what one says is necessarily and individual goal rather than a joint one. And therefore, since *informative* cooperation does not involve a shared goal but only an individual goal, it is not a form of *cooperation* at least in the sense of cooperation that I have adopted. Since Scott-Phillips's *communicative* and *material* cooperation are more or less equivalent to locutionary and perlocutionary cooperation and his i*nformative* cooperation does not seem to be a separate form of cooperation at all, the *communicative-material* cooperation distinction and the *locutionary-perlocutionary* distinction, in effect, are equivalent.

In conclusion, the distinction between *locutionary* and *perlocutionary* captures an important difference in how different goals interact with cooperation. On one hand, a single conversation may involve a multitude of perlocutionary goals and interlocutors may end up cooperating on some of them as shared and not on others. Grice mentions that a conversation can have a *set* of purposes at a given time, and the CP refers a *current* purpose of conversation which can

therefore change as a conversation unfolds. On the other hand, *locutionary* cooperation involves only the goal of making conversation function, and if interlocutors do not share this goal there can be no conversation. In other words, *locutionary* cooperation refers to a more fundamental level of cooperation than *perlocutionary* cooperation. In the next section I will move on from Grice and I will introduce Clark's (1996) account of *language use*.

3.1.2   Clark: Communication as Joint action

As Carston points out (1999), the fact that Clark (1996) uses the expression 'language use' is somewhat misleading as he aims to account for both linguistic and non-linguistic aspects of communication, which he claims cannot be analysed separately. Clark's theory fully embraces the idea that communication is a form of joint action, which is also the foundation of his account of *language use*. Clark's analysis of *language use* as joint action includes all discussed elements of joint action: shared goals or shared projects, coordination, commitment of the participants and common knowledge, which he recasts in terms of common ground. Clark's Common ground is fundamentally equivalent to Lewis's (1969) notion of common knowledge, after which it is modelled: *A piece of information p is common ground for members of a community C if and only if (i) every member of C has information that the state of affairs A holds, (ii) A indicates to every member of C that every member of C has information that A holds (iii) A indicates to members of C that p* (Clark, 1996, p. 94[1]).

Clark adopts Grice's (1989) Non-natural meaning as notion of communication but he recasts this notion in terms of a joint act, the *communicative act*. In this joint *communicative act* the communicator's participative action is to signal something and the addressee's participative action is to understand the signal. In his account, the speaker's meaning is jointly constructed by the interlocutors. Clark applies Austin's notion of uptake (1962) to meaning, whereby for a communicator to have meant that p, they have to be taken by the addressee to have meant that p. Communicator and addressee reach a joint construal of what an utterance means as the turns following an utterance provide incremental evidence for or against the joint construal. For example, imagine a friend and I are at a picnic and she tells me that *it's getting chilly* as a way of suggesting that we should leave. To this I reply that *we should play some badminton to warm up*. My reply may give her evidence that we have different construals of what she has meant by

---

[1] Clark also proposes a reflexive formulation of common ground in which a piece of information *p is common ground in a community of people if (i) members of the community believe that p and that i* (Clark, 1996, p. 95). However, he points out that the formulation based on Lewis's (1969) common knowledge highlights the important role of the state of affairs as a *shared basis* which justifies the assumption of common ground.

her utterance. However, if instead I reply that *I am also a bit tired* this may give her evidence in favour of our joint construal.

Clark's idea of the role of cooperation in communication is very similar to Grice's, even though the role of cooperation, intended in terms of joint action, is presented more systematically in Clark's account. Grice discusses the fundamental features of cooperative transactions in relation to conversation and not in relation to his notion of Non Natural Meaning, which in Grice's account is an individual intention rather than a joint action. In Clark's account conversation even the notion of Non Natural Meaning is recast as a joint action. As in Grice's account, communication is ultimately *identified* with joint action. Therefore the minimal notion of cooperation that is required for Clark's communication has all the fundamental features of joint action: a shared goal or *joint purpose* in Clark's terminology, common knowledge or *common ground*, coordination and commitment, which he discusses in terms of *joint commitments* (p.289). Similarly to Grice, Clark claims that conversation is not 'all cooperative', like other joint activities, and it can be adversarial or competitive when it involves conflicting interests. In other words communication can be a *cooperative* or a *non-cooperative* joint action.

In sum, Clark gives a detailed explanation of the role that cooperation, in terms of joint action, has in communication. He also reviews Grice's account of implicatures and how the CP is fundamental to their derivation according to Grice. Although Clark criticises some aspect of Grice's account of implicatures (1996, p.143) he does not propose a detailed alternative theory of how implicatures are calculated. This is because Clark is not interested in the interpretation process and what exactly the hearer infers the speaker to be communicating. Rather, he is focussed on what the interlocutors *agree has been said* and how this agreement arises. Since the focus of this thesis is on conversational implicatures, Clark's account will play a less crucial role compared to other accounts that provide more detailed accounts of implicatures (e.g., Grice, 1989; Sperber & Wilson, 1986/1995).

3.1.3   Relevance theory

3.1.3.1   *Ostensive inferential communication*

Sperber and Wilson (1995, p.63) offer a definition of overt intentional communication which they dub *ostensive inferential communication*: "the communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions". There are some common elements between this definition and Grice's definition of Meaning$_{NN}$. There is still an informative intention *to make manifest or more manifest a set of assumptions I* and a communicative intention *to make it mutually manifest that the*

communicator intends, by means of the stimulus, to convey the informative intention. An important difference between Meaning$_{NN}$ and *ostensive inferential communication* is that while Meaning$_{NN}$ requires that the recognition of the communicative intention is necessary to the fulfilment of the informative intention, this is not the case for *ostensive inferential communication*. A consequence of this difference is that instances of *showing* fall under the category of *ostensive inferential communication* but not under the category of Meaning$_{NN}$. This is because in cases of showing the recognition of the communicative intention is not *necessary* for the addressee to recover content of the informative intention. To use the same poker example as before, consider the situation in which I *show* my good hand of cards to the other players by flipping my cards towards them. The other players will know that I have a good hand and they will know that I intend them to know this. However, they can *see* that I have a good hand just by looking at my cards, and so the fact that they know that I *intend* them to know is not necessary for them to see that I have a good hand. Therefore, this would be an instance of *ostensive inferential communication* but not an instance of Meaning$_{NN}$.

According to Relevance Theory the hearer is guided by the assumption that the utterances of the speaker are relevant. Specifically, the hearer expects that the speaker's utterance is relevant enough to be worth their effort to process it and that it is the most relevant utterance that the speaker could have uttered given their *abilities and preferences* (Sperber & Wilson, 1995). Implicatures are simply part of the interpretive hypothesis that the hearer makes as part of their comprehension procedure. The Relevance Theoretic comprehension procedure involves the hearer *testing interpretive hypotheses in order of accessibility and stopping when their expectations of relevance are satisfied* (Sperber & Wilson, 2004, p. 613). This contrasts with Grice's account (1989) where hearers are guided by an assumption that the speaker is cooperative and this assumption of cooperation is pivotal for inferring implicatures. Relevance Theory and Grice's account differ in their aims. While Grice's account aims to be a partly descriptive and partly prescriptive account of communication, Relevance Theory's cognitive approach aims to capture the cognitive mechanisms and processes involved in communication.

3.1.3.2   *Epistemic Vigilance and cooperation*

Sperber and Wilson (1986/1995) explicitly acknowledge that Grice's account presupposes a higher degree of cooperation than their account. They argue that that the level of cooperation that Grice postulates is not actually expected in communication and that "it is possible to be optimally relevant without being 'as informative as is required' by the current purposes of the exchange (Grice's first maxim of quantity): for example by keeping secret something that it would be relevant to the audience to know." (Sperber & Wilson, 1986/1995, p.162).

In their view, the only shared goal of communicator and addressee is that the speaker's informative intention is recognised by their addressee. This minimal requirement of cooperation corresponds to what I called *locutionary* cooperation and it is necessary for deriving both the explicit content and the implicit content of an utterance. The account of Epistemic Vigilance (Sperber et al. 2010), which can be considered an extension of Relevance Theory, also postulates that a basic degree of cooperation is necessary for the speaker to be willing to produce utterances and for the addressee to invest in the effort of interpreting them. In other words *locutionary* cooperation is necessary for communication to function. However, Speber and colleagues assume that communication does not involve any cooperation beyond this basic *locutionary* layer. In fact they assume that the nature of communication is not cooperative but manipulative (Krebs & Dawkins, 1984). According to this view speakers and hearers have different goals in communication. The speaker's goal is to affect the hearer and the hearer' beliefs in a way that is beneficial for the speaker; whereas the hearer aims to receive some true and relevant information. This means that is it in the speaker's interest to provoke the intended effect in their hearer by any means necessary, including lying. However, if speakers routinely lied it would become advantageous for hearers to distrust what is communicated. For communication to be advantageous to hearers it has to remain mostly honest. Therefore, as hearers have a suite of cognitive mechanisms that Sperber and colleagues call Epistemic Vigilance, which is aimed at evaluating the trustworthiness of interlocutors and of the information that is communicated to them. This suite of mechanisms protects hearers from the risk of deception and it contributes to keeping communication honest.

In sum, Relevance Theory does assign a role to cooperation in communication but only to *locutionary* cooperation. Sperber and Wilson do not elaborate on this basic form of cooperation that is necessary to make communication function. This basic form of cooperation necessarily has the minimal features of a shared goal (i.e. for the informative intention to be recognized) and coordination of the participant's actions (i.e. the speaker's production of an utterance and the hearer's interpretation). Common knowledge or rather *mutual manifestness* comes into play as it is part of *ostensive inferential communication*. The shared goal to have the informative intention recognized must be mutually manifest to speaker and hearer. This is because the informative intention of the speaker can only be fulfilled if speaker and hearer share the goal that this intention is recognized. Since one can only intend what they believe to be possible (Davidson, 1980), the speaker must believe that the hearer shares their goal that the informative intention is recognized. The hearer must also believe that this information is shared for communication to happen, that is to say that they must recognize the speaker's *communicative* intention to have their informative intention recognized and they must themselves be willing to

recognise it. Therefore, the core features of joint action are also present in Sperber and Wilson's notion of cooperation in communication, except for commitment which they do not discuss. However, *locutionary* cooperation is the only form of cooperation Sperber and Wilson postulate. The presence of shared *perlocutionary* goals is only an incidental feature of the context of communication. Sperber and Wilson recognize that speaker and hearer often do have shared goals that go beyond the mere recognition of the informative intention, however they claim that "Knowledge of such a common purpose, when it exists, is one contextual factor among others, and it is only as such that is can play a role in comprehension" (1995, p.162).

3.1.4   Goals in conversation: source of relevance and QUD

I discussed how the *purpose* of conversation is fundamental in Grice's (1989) account. However the idea that utterances are always relevant to a particular *point* or *topic* plays an important role also in other theories of discourse. Here I will refer to this notion as the *source of relevance* for an utterance. I will treat this notion as theory neutral and I will discuss how different theories formulate it.

In Grice's account the notion of *source of relevance* takes the form of a *purpose* or a shared goal. This purpose gives substance to the maxims in Grice's account. As Russel (2012) argues, the speaker's contribution are expected to be informative and relevant with respects to one particular purpose. Without this purpose it would be impossible for utterances to be *informative enough* as there would be no standard against which to assess their informativity and there would always be an utterance that is more informative.

Roberts (1996/2012, 2004) proposed an influential theory of discourse where the source of relevance takes the double form of a question, or a shared goal (i.e. the goal to find an answer to the question together). Roberts proposes that discourse is a joint enquiry of interlocutors. The ultimate goal of this inquiry if for interlocutors to agree on how to answer the question of what the world is like. This goal cannot be tackled in a single conversation and therefore interlocutors tackle smaller goals, smaller questions, such as agreeing on what the weather is like in London or answering the question of what should go on the shopping list. Discourse is partly organised by the entailment relationships of the questions tackled by the interlocutors. A question B is entailed by a question A if any answer to question A also answers question B (Groenendijk and Stokhof, 1984). Therefore, any question in discourse is entailed by a more general question. For example, the question of what the weather in London today is like is entailed by the question of what the weather in London is like every day of this year, which is entailed by the question of what the weather is like every day of this year in every city of the UK, and so on. When interlocutors agree to address a question that question becomes the

Question Under Discussion (QUD). The QUD can be seen as a goal in the sense that it is the interlocutors' shared goal to answer it. The shared goals of the interlocutors are organized in a *stack* of questions which are ordered by their entailment relationships. In Robert's account QUDs or shared goals take the form of semantic questions. QUD can be explicitly asked in the conversation or, when they are not asked explicitly, interlocutors can infer what QUD an utterance is addressing. Besides Robert's theory, QUD is becoming an increasingly popular way to describe the source of relevance of an utterance (Cummins, 2017) and various theories of discourse and dialogue use it as an analytical tool (e.g. van Kuppevelt, 1996; Ginzburg, 1996).

Relevance Theory (Sperber and Wison, 1986/1995) does not capture the source of relevance for an utterance in terms of a question or a goal but in terms of a *context*, which is a set of contextual assumptions and contextual implications. Contextual implications are conclusions the hearer can draw from integrating the meaning of the utterance with the contextual assumptions. Sperber and Wilson introduce the idea that the source of relevance does not precede utterance interpretation, but the two are constructed together. While in Grice's account it seems that the hearer needs to recover implicatures that fit the assumption that the speaker is working towards a particular *purpose*, according to Relevance Theory hearers start the interpretation process with an expectation of relevance rather than a *purpose* and they look for a source of relevance that once integrated with the utterance meaning satisfies their expectation of relevance.

There is a growing experimental literature investigating the effects of manipulating the explicit or implicit QUD on utterance interpretation (Breheny, Katsos & Williams, 2006; Zondervan, 2009, 2010; Zondervan, Meroni & Gualmini, 2008; Clifton & Frazier, 2012; Politzer-Ahles & Fiorentino, 2013; Degen & Goodman, 2014; Cummins & Rohde, 2015; Kehler & Rohde, 2016). However, the question of how hearers recover the QUD of an utterance has received much less attention. The focus of the utterance has been proposed as an important clue to recover the question that the utterance addresses (van Kuppervelt, 1996; Roberts, 1996; Cummins & Rohde, 2015). Recently, Tian, Brenehy and Ferguson (2010) proposed that negation is another clue that hearers can use to recover the QUD. Besides these two lines of research, the factors that affect the recovery of an utterance's source of relevance are vastly understudied. The study I will present in Chapter 7 aims to address this gap in the literature.

3.1.5   Conclusion

While Grice (1989) and Clark (1996) explicitly identify conversation as a cooperative activity, Sperber and Wilson (1986/1995) assign a minor role to cooperation in communication. Shared *perlocutionary* goals can be argued to afford implicatures in Grice's account (Attardo, 1997). In contrast, Sperber and Wilson argue that *locutionary* cooperation is the only cooperation

necessary for communication, while shared *perlocutionary* goals can affect the interpretation just like other features of the context but they are not necessary in the derivation of implicatures. The identification between joint action and communication is the foundation of Clark's account. While this identification is not explicit in Grice, joint action embodies all the features of Grice's idea of conversation as an instance of a *cooperative transaction*. Sperber and Wilson do not identify communication as a cooperative activity but they claim that communication necessitates a basic form of *locutionary* cooperation. This notion is not discussed in detail but it can be assumed to have most of the fundamental features of joint action (i.e. shared goal, coordination, mutual manifestness). In conclusion, communication involves a basic form of cooperation (i.e. locutionary cooperation) under all accounts and joint action seems to embody the notion of cooperation that is relevant for intentional communication. The role of cooperation on perlocutionary goals varies depending on the account. As I argued in Chapter 1, an excellent way to study the role of cooperation in communication is to look at communication in non-cooperative situations. Therefore, in the next section I will address the question of how communication is affected in non-cooperative contexts.

3.2 Communication in non-cooperative contexts

I have presented the role of cooperation in Grice's (1989), Clark's (1996) and Sperber and Wilson's (1986/1995) accounts of communication, with particular focus on Grice's account of implicatures. Although Grice's and Clark's account give a fundamental role to cooperation, they almost never discuss communication in non-cooperative situations. The focus of most of this thesis is on communication and particularly conversational implicatures in non-cooperative situations. This is because the role of cooperation in communication cannot be studied by only looking at communication in cooperative situations. Observing how communication is affected in non-cooperative situations can provide key information about the relationship of cooperation and communication. Therefore, I will now elaborate on what it means for a situation to be non-cooperative and I will review what predictions can be derived from the theories I presented in the previous chapter about communication in non-cooperative contexts.

Conversation may involve several goals of the interlocutors at the same time and since interlocutors may cooperate on some of them and not others. Cooperation cannot be taken as an all-or-nothing feature of conversation. As discussed in section 3.1.1.7, one important distinction is that between a *perlocutionary* level of conversation and a *locutionary* level of cooperation, where *locutionary* cooperation is to be intended as interlocutors being willing to invest the effort of producing and interpreting utterances with the mere goal of making communication function. Naturally, if this fundamental layer of cooperation is not in place,

communication simply cannot happen. While Clark (1996) and Sperber et al. (2010) make specific reference to this layer of cooperation Grice does not. Since I am interested in cases where communication does happen, the kind of non-cooperative situations I will consider are situations where interlocutors are locutionarily cooperative.

In section 2.1.4, I presented a distinction between cooperative joint actions, where participants have no conflict of interest, and non-cooperative joint actions where participants have opposing goals such that it would be disadvantageous for one participant to help another. And in section 3.1.1.6 I argued that this distinction should also apply within conversation. Both Grice and Clark explicitly claimed that conversation can vary in how cooperative it is and that some conversations happen in competitive or adversarial settings. I will focus on this last kind of situations and from now on I will use the term non-cooperative to refer to situations where participants are engaged in conversation but have conflicting goals that make helping one another disadvantageous. In other words, I will consider situations where interlocutors are willing to communicate (i.e., *locutionarily* cooperative), but have conflicting interests with respect to some *perlocutionary* goals such that it is not in their interest to help one another with respect to these *perlocutionary* goals.

In Grice's account of conversational implicature conversational goals afford implicatures. As discussed in the section 3.1.1, a conversational goal can afford an implicature only if it is mutually known by the interlocutors. The question I will address in this chapter and experimentally in the following chapters is what happens to implicatures when they depend on a goal that is mutually known but that is not a shared goal. I will derive predictions from Grice's account and Relevance Theory about these scenarios. These predictions are not explicitly endorsed by Grice or Sperber and Wilson in their writings; they are my own constructions and they are speculative at least to some extent. I will not review Clark's account in relation to these scenarios because he does not offer a detailed account of implicature derivation that allows deriving specific predictions about implicatures in non-cooperative situations.

3.2.1   Grice

Grice (1989) mentions four ways in which a speaker can *fail to fulfil a maxim*: blatantly *flout* a maxim, violate a maxim because of a *clash*, *opt out* of a maxim or *covertly* violate a maxim. In violating a maxim by *flouting* or because of a *clash* the speaker is overtly obeying the CP. These violations are consistent with the speaker genuinely intending to contribute towards the shared goal. These two types of violations tend to give rise to implicatures. For example, speakers *flout* the maxim of first maxim of quality when they communicate a metaphor through a literally false statement such as "Mary is an Angel". In contrast, when the speaker is *opting out* or *covertly*

violating a maxim they do not intend to further the current purpose of the conversation. In the case of *opting out* their unwillingness to contribute is overt, whereas in the case of *covert* violations their unwillingness to contribute is covert. *Opting out* and *covert* violations do not give rise to implicatures. Covert violations are cases of deception and they are based on the speaker's assumption that the hearer is not aware of the speaker's unwillingness to obey the maxims or the CP. For example, the speaker could covertly violate the first maxim of Quality and say something they believe to be false (i.e. a lie). Or the speaker may say something uninformative and intend that their hearers do not realise that they have not given them enough information. For this reasons, covert violations do not generate implicatures.

By *opting out* a speaker overtly refuses to follow a particular maxim or the CP with respect to a particular *purpose*. For example, if at a conference someone asks me how many of the croissants at the morning refreshment were eaten by me I could say that I have *no comment* on that and so signal that I'm not willing to cooperate on that particular topic. A speaker can also opt out of a particular maxim and so cancel an implicature that might arise from violating that maxim. For example, when asked about my croissant consumption I could say that I had *some* and signal that this is all the information I'm willing to give on the topic. The potential implicature that I did not have *all* is cancelled by the fact that I am opting out of the maxim of quantity (Grice, 1989, p.39).

According to Grice, in order for the speaker to be *opting out* they must "say, indicate, or allow it to become apparent" (1989, p.30) that they are not willing to cooperate. In other words, the speaker can only opt out if it is mutually known among interlocutors that they are doing so. The speaker's opting out may be apparent because the speaker signals it (e.g. *no comment*) or it may be evident from the situation. For example, in a guessing game the guesser does not expect the person who know the answer to give them the information they need; both know that they are opting out in order not to spoil the game.

In situations where the speaker is *opting out* the hearers should not derive implicatures from their utterances. Implicatures depend on the assumption that the speaker is cooperating towards a particular goal and opting out rules out the possibility that this assumption holds. In the example above, where I say that *I ate some of the croissants at the buffet* and signal that I'm not willing to give any more information, it would have been relevant and more informative for me to say that *I ate all the croissants*. If my hearer believed that I am being cooperative, they could infer the implicature that *I did not eat all the croissants at the buffet*; however, if my hearer believes that I am opting out they should not infer this implicature.

Two objections can be raised against the Gricean idea that implicatures disappear when the speaker is opting out. The first objection, as raised by Fox (2014), is to claim that scalar implicatures (but not ignorance inferences) are available even in situations where it is apparent that the speaker is opting out of the maxim of quantity. Fox makes the example of a gameshow where money is hidden in five out of a hundred boxes and the host tells participants that *there is money in box 20 or 25*. Fox argues that the host's utterance give rise to the implicature that the money is not in both boxes even though it is apparent that the host does not intend to be *informative enough* for the purpose of finding the money. However, this objection is not particularly strong. This is because it is debatable whether in Fox's example the inference that the money is not in both boxes is a genuine scalar implicature. This inference may in fact be a non-communicated inference of the speaker based on common knowledge that game show hosts do not easily give away prizes. The second objection is that in some cases of opting out the speakers may be implicating that they reluctant to give more information. Sperber and Wilson (1995, p.273) consider Grice's example where Peter asks *Where Gerard lives* and Mary answers that he lives *somewhere in the south of France.* Here Sperber and Wilson argue that Mary could be implicating that she does not want to say exactly where Gerard lives. In other words she may be communicating that she is *opting out* through an implicature. This kind of implicature is incompatible with Grice's account as it requires the assumption that the speaker is unwilling to cooperate but they are not incompatible with the Relevance Theory (Sperber & Wilson, 1995; Carston, 1998; Mazzarella, 2015b).

In deriving psycholinguistic predictions from Grice's account we should also address the question of whether the mutually manifest assumption of cooperation (i.e. the speaker not opting out) which affords implicatures should be considered as *licensing* or a *cancelling* condition for implicatures (N. Katsos, personal communication, March 2, 2018). If cooperation is a *licensing* condition for implicatures this means that implicatures are not derived at all in situations where the speaker is opting out. If cooperation acts as a cancelling condition, the hearer should derive the implicature and then cancel it upon realizing that the speaker could not have intended it as they are opting out. I will not take a position on the choice between these two possible routes, but I do suspect that both routes may be possible depending on the context (e.g. a context where the hearer is unsure about the speaker's intentions may invite the second route).

A scenario that Grice does not discuss concerns situations where the hearer believes the speaker to be uncooperative and the speaker is not opting out. For instance, these cases could include foiled covert violations, where the speaker lies but the hearer recognizes their attempt at deception. In this type of scenario, even though the hearer knows the speaker to be

46

uncooperative (i.e. not obeying the CP or the maxims), this fact is not mutually manifest as in cases of opting out. It is important to note that from the point of view of the hearer's interpretation, foiled covert violations are not different from 'successful' lies. What changes is what the hearer does with the information they receive.

Besides lying, a speaker can also communicate a false implicature (Meibauer, 2014). A speaker communicates a false implicature when they implicate something and they believe the content of their implicature to be false. For example, I might say that *I ate some of the croissants* and intend that my interlocutor infers the implicature that I did not eat all of them when in fact I know that I ate all of them. The notion of a false implicature is compatible with Grice's account. In fact Grice himself points out that a false implicature can arise from a true statement "*since the truth of a conversational implicatum is not required by the truth of what is said*" (Grice, 1989, p.39).

In the kind of situation considered here, the hearer believes the speaker to be uncooperative with respect to the current purpose of the conversation. In the example above, my interlocutor may believe that I do not intend to help them figure out what happened to the croissants. Since implicatures are afforded by the assumption that the speaker is cooperative as discussed in section 3.1.1.3, Grice's account as formulated in *Logic and conversation* (1975) should predict that hearers do not derive implicatures (true or false) from a speaker that they believe to be uncooperative. This means that false implicatures in a context where the hearer believes the speaker to be uncooperative are incompatible with Grice's account. Notice instead that false implicature in situations where the hearer believes the speaker to be cooperative are perfectly compatible with Grice's account. In other words, Grice's account has no problem explaining cases where the hearer is actually deceived by the false implicature. Consider the example above where my hearer believes that I do not intend to help find out what happened to the croissants. When I say that *I ate some of the croissants* the hearer need not find a justification for why I used an under-informative statement. They already believe that I am uncooperative and therefore they should not derive an implicature from my utterance. However, if my hearer mistakenly believes that I am cooperative with respect to the goal of finding out what happened to the croissants they will justify my under-informative utterance and derive the false implicature that I did not eat all of the croissants.

However, a passage in Grice's later writings seems to make room for implicatures to be possible even in situations where the hearer believes the speaker to be ultimately uncooperative. Specifically, in his Retrospective epilogue Grice mentions a "secondary range of cases" (1989, p. 369-370) such as cross-examination "in which even the common objectives are spurious,

47

apparent rather than real; the joint enterprise is a simulation, rather than an instance, of even the most minimal conversational cooperation; but such exchanges honor the cooperative principle at least to the extent of aping its application". These scenario described by Grice, where the goals are apparent and cooperation is only simulated seem to included the cases I described above where the hearer believes the speaker to be uncooperative even though the speaker is not opting out. The fact that even in this situation the cooperative principle applies means that even in this situations interlocutors can communicate and infer implicatures.

I will make a distinction between the predictions of an early Gricean account based only on Logic and conversation (1975) and a later Gricean account which includes the discussion of the 'secondary range of cases' in the Retrospective epilogue (1989). Grice's early account is inconsistent with hearers inferring implicatures from a speaker that they believe to be uncooperative. However, Grice's later account can account for implicatures being inferred in these scenarios. It is possible to reformulate the expectations of the hearer in the terms of Grice's later account by saying that the hearer only expects the speaker to *appear* to be following the CP. Therefore, hearers can infer implicatures from uncooperative speakers who want to *appear* cooperative. For example, when I say that I ate *some* of the croissants at the refreshment, my interlocutor may assume that I only want to *appear* to be making an appropriate contribution for the purpose of finding out what happened to the croissants, when in fact I am working against this purpose and I am leading them to infer that I did not eat all the croissants, which they can infer to be false. In this thesis I will refer to Grice's predictions in terms of his earlier account.

In sum, Grice discusses cases where speakers opt out of a maxim or they covertly violate one. Opting out has to be apparent and the hearer should not draw implicatures from the utterance of a speaker who is opting out. With lying, which is a covert violation of the first maxim of Quality, speakers rely on the hearers not realising that they are being uncooperative. Besides lying, uncooperative speakers can also communicate false implicatures, which by virtue of being implicatures should only be derived if the hearer believes the speaker to be cooperative. Grice does not discuss situations in which the both speaker and hearer know that the speaker is uncooperative but this information is not mutually known. I interpreted Grice's account, at least in it's early version, to predict that in these situations hearers should not derive implicatures. A problematic phenomenon for Grice in these scenarios are foiled false implicatures, where the hearer derives implicatures from the speaker's utterances and infers that the content of the implicatures is false because they believe the speaker to be uncooperative.

3.2.2   Relevance Theory

According to Relevance Theory hearers are guided by an assumption that the speaker will be relevant, not by an assumption that they will be cooperative. The only kind of cooperation assumed is *locutionary* cooperation (Sperber & Wilson, 1986/1995; Sperber et al. 2010). In interpreting an utterance the hearer constructs the most relevant interpretation compatible with the speaker's *abilities and preferences*, which may include implicatures of the utterance, regardless of whether the speaker is cooperative with respect to the goals involved in the conversation. The speaker's *abilities* may relate to the speaker's competence on a particular topic and their ability to provide certain information. The speaker's *preferences* instead are related to the speaker's goals. For example a speaker may be unwilling to provide a piece of information because providing it would be against their interests. Therefore the provision that the interpretation has to be compatible with the speaker's *abilities and preferences* allows for scenarios that would fall under Grice's category of opting out. In other words, if the hearer believes the speaker to be reticent on a particular topic they may not infer implicatures from their under-informative utterances. For example, if a speaker says that they are *meeting someone for dinner* and the hearer believes the speaker to be reserved about this topic they may not infer implicatures from their under-informative utterance (e.g. that the speaker does not know whom they are meeting for dinner).

Wilson and Sperber (2002) take issue with Grice's first maxim of quantity and the requirement for truthfulness in *what is said*. They argue that this requirement poses problems for the ability of Grice's account to explain phenomena of loose uses of language (e.g. Holland is flat), which do not seem to violate the maxim nor abide it perfectly either. They however side with Grice's supermaxim of quality in assigning an important role to truth in *what is communicated*. They propose that "the notion of cognitive efficiency cannot be divorced from that of truth" (2002, p.263) and therefore relevance depends on genuine (i.e. true) improvements of the epistemological state of the individual. They introduce the notion of *positive cognitive effects* which contribute positively to cognitive goals, which according to Wilson and Sperber correspond to constructing the best (and true) representation of the world. The relevance of an assumption to an individual in a particular context depends on its *positive* cognitive effects. Cognitive effects that are not positive do not contribute to relevance even if the individual believes them to be positive. For example, a lie *seems* relevant because the cognitive effect it produces will seem *positive* to the individual being deceived. But the lie is not actually relevant because it does not advance the addressee's knowledge of the world. If told my interlocutor that *I wrote my PhD thesis in two days*, this piece of information may *seem* very relevant to them as it would give rise to a wealth of cognitive effects due to adjusting their prior beliefs about my

intelligence and writing speed. However, since my utterance would be a lie, the cognitive effects it would give rise to would not be *positive* cognitive effects and therefore it would only *seem* relevant.

According to Relevance Theory a communicator can be seen as having two distinct goals, the goal to be understood - their communicative intention - and the goal to be believed - their informative intention. While comprehension requires the communicative intention to be successful, the informative intention may be successful or not (Sperber & Wilson, 2004; Mazzarella, 2015a). For example, in situations where the hearer realizes that the speaker is lying, the speaker's communicative intention is successful but their informative intention probably not as the hearer is unlikely to believe them. Sperber et al. (2010) expand the framework of Relevance Theory by postulating a set of mechanisms which allow hearers to evaluate the trustworthiness of the speaker as a source of information and the trustworthiness of the communicated content. These processes determine whether the hearer will accept the communicated information, but they do no influence the interpretation process. Instead they run in parallel to interpretation (Sperber et al., 2010).

Sperber and colleagues therefore propose a model where interpretation and acceptance are distinct but parallel processes (Mazzarella, 2015a). This is in contrast with models that postulate that comprehension and acceptance are a single process followed by an optional process of 'disbelieving' (Gilbert, 1993; Millikan, 2004; Millikan, 2005) which Mazzarella calls the Spinozan models. And it is also in contrast with Cartesian models, where comprehension and acceptance are distinct but sequential processes, where epistemic evaluation follows comprehension (Mazzarella, 2015a). According to Sperber et al. (2010) judgements of trustworthiness do not affect the interpretation process in the sense that hearers interpret utterances *as if they were trustworthy*. This prediction is relevant to situations where the hearer believes the speaker to be uncooperative because in these situations the hearer is also likely to consider the uncooperative speaker an untrustworthy source of information. In this kind of scenario Sperber and colleagues predict that hearers *should* derive implicatures from the utterances of untrustworthy-uncooperative speakers and then possibly reject the content of the implicatures. The experiments presented in the following chapters will address the topic of a dissociation between comprehension and acceptance of the communicated content. However, as the paradigms use in these experiment do not investigate the online processes of comprehension and acceptance but only their end results, the results of these experiments will not be able to distinguish between the parallel model supported by Sperber and colleagues, Spinozan models and Cartesian models.

Although in this thesis I will focus on the dissociation between comprehension and acceptance which I presented as a feature of the Epistemic Vigilance account, there have been proposals which argued for a more interconnected relationship between the mechanisms of Epistemic Vigilance and the process of comprehension. Mazzarella (2015c, 2016) argued that the mechanisms of Epistemic Vigilance may have early effects on interpretation. Specifically, Mazzarella argued that Epistemic Vigilance towards the speaker determines the expectations of relevance of the speaker. For example, if the hearer believes that the speaker is not *competent* they will adjust their expectation of optimal relevance to an expectation of *attempted optimal relevance*. If they believe the speaker to be deceitful, their expectation will be adjusted to a *purported optimal relevance*. Mazzarella argues that these adjustments of the expectations of relevance caused by the Epistemic Vigilance mechanism may push the hearer to rule out interpretive hypotheses that are not compatible with the speaker's assumed *competence* and they may push them to consider interpretive hypotheses that the speaker already believes to be false (and therefore irrelevant) because if they are dealing with a deceptive speaker. These adjustments can therefore have a major influence on the interpretation process.

Since the prediction derived from Grice concerns *uncooperative* speakers and the predictions of the Epistemic Vigilance account concern *untrustworthy* speakers, by juxtaposing the two accounts I may appear to implicitly assume that these two categories of speakers are the same. I do not wish to make this assumption and I do not believe that this assumption is necessary to compare the predictions of the two accounts. While untrustworthiness only captures whether the speaker is likely to give accurate information, cooperation or its lack may capture to the speaker's motivations for giving accurate or inaccurate information. *Untrustworthiness* and *uncooperative-ness* do not necessarily coincide. For example, I may judge a cooperative speaker to be untrustworthy because I believe them to be misinformed. In this thesis however I will focus on cases where speakers are both uncooperative and untrustworthy and so the predictions of the two accounts with both apply.

In sum, Relevance Theory (Sperber & Wilson, 1986/1995; Sperber et al. 2010) is compatible both with speakers 'opting out' *à la* Grice and with hearers deriving implicatures from speakers that they consider uncooperative and untrustworthy. For the latter scenario Sperber et al. (2010) predict that the hearer should interpret the implicature of the speaker *as if the speaker was trustworthy* and then possibly reject the content of the implicature. This is in contrast to the prediction derived from Grice.

3.2.3 Lying and falsely implicating

An uncooperative speaker may lie or implicate something false. According to traditional definitions of lying, to lie is to *say* something believed to be false with the intent to deceive the hearer (Isenberg, 1973; Primoratz, 1984). A false implicature is to communicate something false through a conversational implicature (Meibauer, 2014). There is an ongoing debate on whether false implicatures should be considered to be lies (Meibauer, 2005, 2014) or not (Dynel, 2011; 2015). Coleman and Kay (1981) and Hardin (2010) asked participants to rate to what extent they thought an utterance giving rise to a false implicature was a lie. Both studies used the following example of a false implicature (Coleman & Kay, 1981, p.31):

> *John and Mary have recently started going together. Valentino is Mary's ex-boyfriend. One evening John asks Mary, 'Have you seen Valentino this week?' Mary answers, 'Valentino's been sick with mononucleosis for the past two weeks.' Valentino has in fact been sick with mononucleosis for the past two weeks, but it is also the case that Mary had a date with Valentino the night before. Did Mary lie?*

The ratings they collected were on average near the middle of the scale. The question of whether or to what extent false implicatures are lies is closely connected to the issue of whether implicatures can be incorporated into *what is said* or into the truth-conditional content of an utterance.

Grice sees 'what is said' to be primarily determined by the semantic import of an utterance: "I intend what someone has said to be closely related to the conventional meaning of the words (the sentence) he has uttered" (1989 p.25). Grice's *what is said* involves contextual elements only to the extent of resolving any ambiguities (semantic or syntactic) and saturating elements of the sentence that allow arriving at a truth-evaluable proposition (e.g. reference resolution). Therefore, according to Grice, implicatures (conventional or conversational) are not part of what is said. Theorists who have worked on the notion of what is said have diverged from Grice and argued either for a more minimal notion of what is said or for a more enriched/pragmatic notion than Grice's own notion. Proponents of a more minimal notion propose a theoretically motivated notion that is very close to the semantic meaning of the sentence and that may also be sub-propositional (e.g. Bach, 1994; 2001). Proponents of a more enriched notion have aimed for psychologically-motivated notions that are fully propositional and that may incorporate various types of pragmatic inferences (including implicatures) which contribute to the truth-conditional meaning of the utterance (e.g. Récanati, 1989; 2001; Sperber & Wilson, 1986/1995, Carston, 2002). This more enriched theoretical notion of *saying* is also closer to the intuitive notion of *saying*. Récanati proposes that "In deciding whether a pragmatically determined aspect of utterance meaning is part of what is said, that is, in making a decision concerning what

is said, we should always try to preserve our pretheoretic intuitions on the matter" (Récanati 1989, p. 310). He called this requirement for a notion of 'what is said' *The Availability Principle* (Récanati, 2001). Relevance theorists (Sperber & Wilson, 1986/1995, Carston, 2002) propose their own notion of *explicature*, which is "*a development of a logical form encoded by the utterance*" (Sperber & Wilson, 1995, p. 182), and do away with the notion of what is said altogether. Carston (2002, p.183) argues that a notion of *what is said a la* Grice is unsuited to an account that is aimed at capturing the cognitive processes involved in interpretation.

In this thesis I am mostly interested in the speakers' and hearer's intuitive and therefore pretheoretic notions of *lying* and *saying.* For this reason I will not discuss the theoretical debate on the notion of *what is said* in detail. I will rely on Grice's notions of *what is said* and *what is implicated* and I will use the expressions *explicit content* and *implicit content* of an utterance as equivalent to Grice's notions. This use of the expressions *explicit* and *implicit,* which are theoretically loaded in Relevance Theory (Carston, 2002, 2009) is simply a matter of convenience and it is not meant to reflect a theoretical stance. Similarly, I will not discuss the theoretical arguments in favour or against the inclusion of false implicatures under the category of lying. I will however point out choosing one notion of *what is said* over another may drastically affect the range of cases that notions *lying* based on the falsity of *what is said* apply to. Definitions of lying which are a based on the speaker *believing* that what they said is false will necessarily call for notions of *saying* that are psychologically valid and that aim to model the pre-theoretic notion of *saying*. Another terminological premise I should make is that throughout this thesis I talk about quantity implicatures (Ad hoc implicatures and Scalar implicatures in particular) and I sometimes discuss the possibility that these 'implicatures' may be integrated into the explicit or truth-conditional content of the utterance. In doing this I am using the term implicature loosely and I am sacrificing theoretical precision for the sake of using consistent terminology through the chapters. In fact, implicatures are by definition implicit and not part of the truth-conditional content of the utterance (Grice, 1989). Therefore in these cases it would be more exact to talk to talk of pragmatic inferences rather than implicatures.

Empirical research has addressed the question of whether implicatures enter the intuitive notion of *what is said* or affect the intuitive truth-conditions of an utterance. Nicolle and Clark (1999) asked participants to choose between two paraphrases of an utterance which was reported after a short context:

> *Peter and Mary were rock climbing; Mary was leading and Peter was standing at the foot of the cliff holding the ropes. All of a sudden Mary slipped and fell about 10 feet. Peter asked her if she could continue, but Mary said, 'I've twisted an ankle.'*
> *a) Mary has twisted her own ankle*

b) Mary can't carry on climbing

One option was a close periphrasis of what the utterance *said* while the other option was an implicature of the utterance which was made very accessible by the context. Depending on the condition they were in, participants were asked to select the paraphrase that best reflected either *what the sentence said, what the speaker's words meant* or *what the speaker wanted to communicate*. The results indicated that participants were more likely to select the implicature regardless of the condition they were in. Doran and colleagues (Doran, Baker, McNabb, Larson & Ward, 2009; Doran, Ward, Larson, McNabb & Baker, 2012) probed whether implicaures can be incorporated in the truth-conditional content of an utterance by asking participants to indicate whether sentences that could give rise to an implicature were true or false in the light of a fact that contradicted the implicature (Doran, Ward, Larson, McNabb & Baker, 2012, p. 139):

> *Irene: How much cake did Gus eat at his sister's birthday party?*
> *Sam: He ate most of the cake.*
> *FACT: By himself, Gus ate his sister's entire birthday cake.*

In the example above, responding that Sam's utterance is false would indicate that the scalar implicature of *most* has been incorporated into the truth-conditional meaning of the utterance. Doran and colleagues used a wide variety of implicatures and they found that participants who were asked to interpret the utterance *literally* incorporated the implicature in the truth-conditional content of the utterance 44% of the time. Participants who received no instructions as to how to interpret the utterance incorporated implicatures 50% of the time. Thus, both the study of Nicolle and Clark and that of Doran and colleagues suggest that implicatures are often incorporated into the intuitive notion of *what is said*.

Meibauer (2014) suggests that different types of implicatures may be perceived as closer or farther from the traditional notion of lying. According to Meibauer, whether false implicatures are perceived as 'lying' depends on the relevance of the information conveyed by the implicature. The more the false implicature is relevant to the purpose of the conversation the more it can be considered a lie. I agree with Meibauer that the relevance of the implicature determines whether a false implicature is perceived to be a lie. Besides relevance, two other factors may influence whether an implicature is considered a lie. The first factor is the strength with which an implicature is communicated, which is tightly connected to the hearer's perception of how strongly committed the speaker is to the truth of the implicated content. According to Sperber and Wilson, the strength of an implicature increases with the strength of the speaker's informative intention to communicate that assumption. "*The strongest possible implicatures are those fully determinate premises or conclusions […] which must actually be supplied if the interpretation is to be consistent with the principle of relevance, and for which the speaker takes full responsibility*" (Sperber & Wilson, 1995, p.199). The more the implicature is an

assumption supplied by the hearer without the encouragement of the speaker the weaker the implicature is. For example, if while ordering food in a pizzeria I say that *I do not like American pizza recipes* I may be strongly implying that I am not going to order Hawaiian pizza or barbecue chicken pizza. My interlocutor could also supply the assumption that people who do not like American pizza also do not like American fast food chains such as McDonald's and so infer that *I don't like McDonalds.* This last inference however, if it can be considered an implicature at all, it is a very weak implicature as it rests an assumption supplies by the hearer. Furthemore, when an inference is based on assumptions for which the hearer is responsible, the perceived *commitment* of the speaker to content of the inference is small (Morency, Oswald & de Saussure, 2008). The literature on commitment has mainly focused on the idea that explicit content is associated to a higher degree of speaker's commitment than implicit content (Morency et al., 2008; Reboul, 2017). Morency and colleagues argue that this is because implicatures are defeasible and some of the assumptions involved in their derivation are not communicated but supplied by the hearer.

A second factor that may influence how close to a lie a false implicature is perceived to be may be whether the implicature in question is a *generalised conversational implicatures (GCI)* or *particularised conversational implicatures* (PCI). Scalar implicatures, such as the potential implicature in *John used some of the shampoo*, are based on an informativity scale that is tied to a lexical item (e.g., some) and is independent of the particular context of use such as *<all, most, some>*. Ad hoc implicatures are based on a scale of informativity which is specific to a particular context. For example, someone might say that they *mowed the front garden* and implicate that they did not also mowed the back garden only in a context where <both front and back garden, front garden> form an informativity scale.

The distinction between GCIs and PCIs may be relevant to whether false implicatures are perceived to be lies is based on experimental results of Katsos (2009). Katsos asked child and adult participants to correct a speaker (a fictional character) who used under-informative utterances that could give rise to GCIs and PCIs in contexts were the content of the implicature was false (e.g. using the utterance *the elephant pushed some of the trucks* when in fact the elephant pushed all of them). Although participants corrected the speaker's false GCIs and false PCIs to the same extent, adult participants tended to judge utterances that could give rise to false GCIs as unambiguously wrong in the context, whereas they judged the utterances that could give rise to false PCIs as partly right or *technically* right. Katsos (2009) argues that although his results are overall consistent with other studies in favouring a unitary view of GCIs and PCIs (e.g., Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006), he also suggests that adult speakers saw under-informative GCI-utterances as a stronger violation of informativity than

false PCIs. This perceived difference in the strength of the violation suggests that false GCIs are perceived to be more of a lie than false PCIs.

In summary, the phenomenon of false implicatures is relevant both to the theoretical debate surrounding the notion of lying (e.g., Meibauer, 2005, 2014; Dynel, 2011, 2015) and the one surrounding the notion of *what is said* (e.g., Bach, 1994; 2001; Récanati, 1989, 2001; Carston, 2002). In this thesis I will not propose theoretical arguments for one position over others in either of these two debates. However, the empirical studies that I will present in the following chapters may provide some information about the pretheoretic notions of lying and saying and whether (false) implicatures are considered part of what speakers *say*.

3.2.4   Possible advantages of implicatures in non-cooperative contexts

So far I have discussed how non-cooperative settings may affect implicatures. Now I will discuss how implicatures may be useful in these settings. Although Grice's account of conversational implicatures is based on cooperative situations, there are arguments in favour of the idea that implicit communication offers advantages compared to explicit communication in contexts where the interests of the interlocutors are not or may not be aligned. I will review the proposals of Pinker and colleagues (Pinker, Nowak & Lee, 2008; Lee & Pinker, 2010) and of Reboul (2017).

Pinker and colleagues propose that the *plausible deniability* of implicit communication allows speakers to deal with situations of uncertainty where they want to communicate something that may incur them in a cost depending on the goals and interests of their interlocutors. Precisely because the content of the implicature is not part of what is said, as discussed in the previous section, it is not part of the truth-conditional content of the utterance and it is – by definition- always deniable. Pinker and colleagues give the example of offering a bribe, which may incur the person offering in a significant cost. Offering a bribe in an implicit way leaves open the possibility of denying that the person ever intended to offer a bribe. The same reasoning can be applied to cases of deception. Lying carries a cost in reputation for the person caught lying. However, if the speaker implies something false with the aim of deceiving the interlocutor they can later deny that they intended the implicature and so deny that they intended to deceive.

Reboul (2017) proposes that implicit communication gives the speaker a higher chance of bypassing the hearer's epistemic vigilance mechanisms compared to explicit communication. Firstly, she claims that hearers are more vigilant towards communicated content when they perceive that the speaker is strongly committed to it. Since speakers are perceived to be more committed to what they explicitly communicate than to what they communicate implicitly (Morency, Oswald & de Saussure, 2008), hearers should be less vigilant towards what the speaker communicates implicitly. Secondly, Reboul claims that hearers are less vigilant towards

beliefs that derive from their own inferences compared to beliefs that are communicated to them. She bases this claim on what Mercier (2009) describes as the *egocentric bias*, which is the preference for our own beliefs over those that are communicated to us. Since implicatures are the result of the hearer's own inferences to a larger extent than explicitly communicated content, hearers should be less vigilant towards belief communicated through implicatures.

3.3 Experimental evidence on Quantity implicatures

In the following three chapters of this thesis I will address the question of how uncooperative contexts affect quantity implicatures both in comprehension and production. I will therefore review some of the experimental literature on this phenomenon. I will focus on experimental results that offer indications of how these inferences are derived and some of the relevant factors that affect their derivation. Since most of the literature on quantity implicatures focuses on scalar inferences I will start from these.

An important segment of the empirical research on quantity implicatures has addressed the theoretical debate on whether scalar implicatures are *nonce* inference which are derived *only* when the context of the utterance supports their inference (Sperber & Wilson, 1986/1995; Carston, 1995, 1998; Geurts, 2010) or whether they are *default* inferences which hearers derives whenever they encounter scalar expressions (e.g. the quantifier *some* automatically mandates the inference *not all*) and which can be cancelled if they are inconsistent with the context (Levinson, 2000; Chierchia, 2004). While *nonce* inferences should not arise without contextual support and they should require some additional time and effort compared to simply decoding linguistic meaning, *default* inferences should be fast and automatic (Levinson, 2000). A wealth of experimental evidence has now accumulated showing that the computation of scalar implicature is not automatic in contexts that do not support the inference (Breheny, Katsos & Williams, 2006) and that it is time consuming and effortful (Noveck & Posada, 2003; Bott & Noveck, 2004; Huang & Snedeker, 2009, 2011; Bott, Bailey, & Grodner, 2012) and in particular it is taxing on the hearer's working memory resources (De Neys & Schaeken, 2007; Dieussaert et al., 2011; Marty & Chemla, 2013; Marty, Chemla & Spector, 2013). This evidence has been interpreted as disproving the hypothesis that scalar implicatures are *default automatic* inferences that are derived and then cancelled if inconsistent with the context and in favour of the idea that they are *nonce* inferences arising from context. This is also indicates that scalar implicatures are derived by the same mechanism as Ad hoc quantity implicatures, which are particularised implicatures and therefore *nonce* inferences by definition. Although the evidence I reviewed puts in serious question the validity of drawing a theoretical distinction between scalar and Ad hoc implicatures, some developmental studies indicate that there are differences in the acquisition of these two inferences as pre-school children find it harder to derive scalar

implicatures compared to *Ad hoc* implicatures (Bale, Brooks & Barner, 2010; Stiller, Goodman & Frank, 2011). However, results are mixed as others found evidence that children calculate these two types of implicatures to the same extent (Katsos, 2009).

Nevertheless, Katsos (2009) found evidence that hearers perceive the under-informativity of scalar expression to be a more serious violation of informativity than in the case of under-informative utterances which could give rise to Ad hoc implicatures.

Experimental investigations have uncovered a wide variety of factors that influence the derivation of scalar implicatures. As I mentioned working memory is one of them. Hearers are less likely to derive scalar implicatures if they are under working memory load (De Neys and Schaeken, 2007; see also Dieussaert et al., 2011; Marty & Chemla, 2013; Marty, Chemla and Spector, 2013). A related finding indicates that having a smaller working memory capacity makes people less likely to derive scalar implicatures (Antoniou, Cummins, Katsos, 2016). Focus and the QUD have also been found to affect the derivation of scalar implicatures. These two factors are closely related as focus indicates which questions an utterance may be an answer to (Jackendoff, 1972) and so it provides clues as to what QUD it is addressing (Roberts, 1996/2012). Experiments manipulating the explicit QUD of utterances which could give rise to scalar implicatures found that hearers are more likely to infer a scalar implicature if the salient QUD in the context inquires whether the stronger alternative is true (Breheny, Katsos & Williams, 2006; Zondervan, Meroni & Gualmini, 2008; Zondervan, 2010; Degen & Goodman, 2014). Studies that manipulated the QUD by manipulating the information structure of a sentence found that participants were more likely to infer implicaures if the scalar expression appeared in a focused position in the utterance (Breheny, Katsos & Williams, 2006; Nieuwland et al., 2010). Studies that manipulated intonational focus by placing sentence stress on a scalar expression (e.g. 'or') found that hearers were more likely to infer a scalar implicature if the scalar word bore intonational focus (Zondervan, 2010, p.69[2]; Schwarz, Clifton & Frazier, to appear; Cummins & Rohde, 2015).

There is a growing interest in how individual differences affect the derivation of scalar implicatrues. The results of Antoniou, Cummins and Katsos (2016) indicate that age and working memory capacity affect sensitivity to under-informative utterances. Nieuwland et al. (2010) found evidence that people with autistic traits are also less likely to infer implicatures. This is a promising line of research to explain the fact that experiments often have a group of participants who consistently choose the semantic interpretation and a group who consistently choose the pragmatically interpretation for the same utterance (Noveck, 2001; Bott & Noveck, 2004;

---

[2] Zondervan (2010) manipulated utterance context and sentence stress at the same time and therefore it is unclear how much of the effect he found is attributable to intonational focus.

Feeney et al., 2004; Guasti et al., 2005 i.a.). Further research is likely to uncover further factors that affect an individual's sensitivity to under-informative utterances or their pragmatic *tolerance* (Katsos and Bishop, 2011).

Particularly relevant to the studies I will present in this thesis is the fact that the hearer's assumptions about the speaker have been found to affect implicature derivation. Bergen and Grodner (2012) found evidence that the competence of the speaker about the stronger alternatives from which quantity implicatures is generated immediately affects whether hearers derive a scalar implicature or an ignorance inference. Breheny, Ferguson and Katsos (2013) found complementary evidence that also for Ad hoc quantity implicature the epistemic state of the speaker immediately affects the derivation of the inferences. Grodner and Sedivy's (2011) study, although it focuses on contrastive inferences, demonstrated that the hearer's assumptions about the speaker's reliability strongly affect whether they will draw inferences from their utterances. They asked participants to move objects in a display following the spoken instructions of a confederate. Participants were given multiple cues that the confederate did not use noun modification to generate relevant contrastive inferences, for example the confederate may ask the participant to move the *tall glass* even if there was only one glass in the display. They found that this caused participants not to draw contrastive inferences in interpreting the confederate's utterances even in scenarios there the contrastive inference would have been relevant (i.e. where there were both a *tall glass* and a *small glass* in the display).

Related to this issue is the finding that hearers seem to draw less scalar implicatures in face-threatening contexts (Bonnefon, Feeney & Villejoubert, 2009; Bonnefon, De Neys & Feeney, 2011; Feney & Bonnefon, 2013). For example, hearers are less likely to draw an implicatures from an utterance like 'some people hated your speech' (i.e. not everyone hated your speech) compared to 'some people loved your speech'. This finding was interpreted as evidence that face threatening contexts block implicature derivation, and it may be interpreted more generally as evidence that the speaker's assumptions about the speaker's goals and motivation influence whether implicatures are drawn. However, as Mazzarella and colleagues (Mazzarella, 2015a; 2015b; Mazzarella, Trouche, Mercier & Noveck, 2016) point out the experiments of Bonnefon and colleagues measure whether participants believe that the content of the implicature is true and not whether they draw the inference. As predicted by the Epistemic Vigilance account (Sperber et al. 2010) hearers may infer implicatures and then reject their content. Mazzarella and colleagues (2016) replicate the experiments of Bonnefon et al. and provide evidence that politeness makes the implicature less believable but not less available. Mazzarella and colleagues point out that there is a difference between measuring the comprehension and

acceptance of a pragmatic inference, and this remark is actually relevant to most experiments on implicatures and it is particularly important for the studies presented in this thesis.

I mentioned that much empirical research on scalar implicatures addressed the question of whether scalar implicatures are *nonce* or *default* inferences. Another theoretical debate that motivated empirical research concerns issue of whether the *exact* readings of numerals is due a scalar implicature or another semantic phenomenon. One position holds that numeral expressions have a semantic *at least* meaning, and that context can supply an upperbound *at most* interpretation which yields the *exact* interpretation (e.g., Horn, 1972; Gazdar 1979; Levinson, 1983; Levinson, 2000). According to this position the utterance "John ate three cookies" literally means that John ate *at least* three cookies, and in a context where it is relevant whether John ate more than three the utterance is pragmatically enriched to mean that "John ate *exactly* three cookies". An opposed position holds that the *exact* meaning of numeral expressions is part of their literal or truth-conditional meaning (Carston, 1998[3]; Breheny 2008; Kennedy 2015). Huang, Spelke and Snedeker (2013) found evidence showing that processing the *exact* interpretation of a numeral expression is as fast as processing the literal meaning of the quantifier *all*, while deriving the scalar implicature of the quantifier *some* is a slower process. Huang, Spelke and Snedeker (2013) as well as Papafragou and Musolino (2003) also show that pre-school children tend to give exact interpretations of numerals while they tend to give semantic interpretations of scalar expression (i.e., not calculating scalar inferences). Furthermore, Marty, Chemla and Spector (2013) found that while being under working memory load makes hearers derive less scalar implicatures, it causes them to assign more *exact* readings to numeral expressions. These pieces of evidence are in favour of the view that the *exact* meaning of numeral expressions is part of their semantic meaning rather than arise in context as an implicature.

Finally, to the best of my knowledge, the only experiment that investigated implicatures in non-cooperative context was conducted by Pryslopska (2013). She asked participants to play a competitive game with two confederates. One of the confederate was allied with the participant while the other was an opponent, playing against them. Each player was given a card displaying a set of coloured pebbles. Each players aimed to find out what was on the other players' cards and they could ask them questions about it. Players could not lie but they were allowed to be obscure or under-informative. While it was in the ally's interest to give informative answers to the participant about what their set contained the opponent benefited from being under-

---

[3] Carston (1998) actually argues that cardinals have an underspecified meaning and that whichever sense they assume in context (i.e. at least, at most or exactly) contributes to the truth conditional meaning of the utterance.

informative. After playing this game participants performed a computerised task where they heard the voices of the two confederates describing cards analogous to the one used in the game. In each trial of the experiment participant had to match the description with one of three card displayed on the screen. In critical items the descriptions contained the quantifier *some* (e.g., some of the pebbles are red) and the cards displayed included a pragmatic option where *some but not* all pebbles were red, a semantic option where *all* pebbles were red and a wrong option where *none* of the pebbles were red. In this computerised task participants were simply instructed to click on the card that best fitted the description. Pryslopska found that participants chose the pragmatic option on 88% of trials when the description was uttered by the ally and on 66% of trials when the description was uttered by the opponent. Participants gaze was also tracked during the computerised task. The eye-tracking data showed that when the description was uttered by the opponent participants were slower to fixate the *pragmatic* option compared to when the description was uttered by the ally. The eye-tracking data showed that the opponent's description made participants slower to fixate the target card even in control items where only one card matched the description (e.g. the description said that *all* the pebbles were red and only one option had all red pebbles). These results are interesting as they indicate that implicatures are affected when they arise from the utterances of a non-cooperative speaker. However, this experiment suffers from two problems. The first problem is that the testing phase is not in itself a non-cooperative scenario. The manipulation is based on a priming phase (i.e., the game) where the non-cooperative was an opponent. In the testing phase participants have no information about the confederate's intentions when the descriptions were recorded and whether the opponent benefited from them making mistakes in the task. The second problem is that, similarly to the experiments of Bonnefon and colleagues, this paradigm does not distinguish between comprehension and acceptance of the scalar inference (Mazzarella, 2015a, 2015b). Participants in this experiment may have been drawing implicatures from the utterances of the ally and the opponent at the same rate but then rejected the content of the opponent's implicature. In fact, it is possible that the delay shown by the eye-tracking data reflects the epistemic assessment processes that participants performed when faced with an untrustworthy opponent.

3.4   Conclusion

In this section I addressed the question of what happens to implicatures in non-cooperative contexts. The non-cooperative contexts I considered are situations where interlocutors are *locutionarily* cooperative but they do not cooperate on some *perlocutionary* goals. Grice (1989) accounts for cases where a speaker opts out of the CP and maxims or covertly violates a maxim, for example by lying. Grice does not consider scenarios where the hearer knows that the speaker

is uncooperative but this information is not mutually known with the speaker. In this type of scenario I derived the prediction from Grice's account that the hearer should not infer implicatures from the speaker's utterances. In contrast, the prediction I derived from Relevance Theory (Sperber & Wilson, 1986/1995) and the Epistemic Vigilance account (Sperber et al., 2010) is that hearers derive implicatures from the utterances of speakers that they consider uncooperative and untrustworthy but they may reject the content of the implicatures. The dissociation between the processes of comprehension and acceptance proposed by the Epistemic Vigilance account (Sperber et al. 2010; Mazzarella 2015a) makes this account particularly suited to capture situations where a hearer infers a false implicature (Meibauer, 2014) from the utterance of a speaker that they consider untrustworthy and they reject its content. I reviewed some possible advantages of implicatures over assertions in non-cooperative context as proposed by the Strategic Speaker theory (Pinker, Nowak & Lee, 2008; Lee & Pinker, 2010) and Reboul (2017). Finally, as the following chapters will present experiments investigating quantity implicatures in non-cooperative contexts I reviewed a relevant part of the experimental literature on quantity implicatures and the only existing study investigating implicatures in a non-cooperative context (Pryslopska, 2013).

Accounts of intentional communication, including the ones I introduced in this chapter, very rarely address the question of what happens to communication in non-cooperative contexts directly. From an experimental point of view the attention to this topic is equally scarce. In particular the question of what happens to implicatures in non-cooperative contexts remains to be investigated. As I argued in Chapter 1, investigating communication in non-cooperative contexts is essential for the aims of this thesis as it is not possible to study the role of cooperation in communication only by looking at communication in cooperative situations. Furthermore, looking at implicatures in non-cooperative contexts is particularly interesting as diverging predictions can be derived from Grice's (1989) account and Relevance Theory (Sperber & Wilson, 1986/1995) regarding how they are affected in non-cooperative contexts. The studies presented in the following chapters aim to partly fill this gap in the experimental literature.

4 Inferring Quantity implicatures in a competitive game

In the previous chapter I introduced open questions and issues regarding communication in uncooperative contexts. In particular, I presented predictions that can be derived from Grice's account (1989) and Relevance Theory (Sperber & Wilson, 1986/1995; Sperber et al., 2010) regarding inferring implicatures from uncooperative speakers. In this chapter I present a study that addresses these questions empirically by operationalising the uncooperative context as a competitive signalling game.

4.1 Introduction

The present experiment investigates what happens to Quantity implicatures in a controlled non-cooperative context. To the best of my knowledge, Pryslopska (2013) conducted the only existing study on the comprehension of quantity implicatures in a non-cooperative context. She operationalised the speaker's lack of cooperation with a priming phase where the participant and the speaker played a competitive game where the speaker had to give hints to the participant even though it was in their interest to reveal as little as possible. In a subsequent testing phase, Pryslopska's participants listened to under-informative utterances of the uncooperative speaker which could give rise to a scalar inference and they indicated their interpretation of these utterances. Pryslopska found that her participants inferred less scalar inferences from the under-informative utterances of an uncooperative speaker compared to the inferences they inferred from a control cooperative speaker. In the present experiment the non-cooperative context is operationalised as a competitive signalling game where signaller and receiver are opponents. In contrast with Pryslopska's paradigm, there are no separate priming and testing phase as the whole experiment is embedded in the competitive game. A signalling game is a game where a player, the signaller, sends a signal based on private information to a second player. The second player, the receiver, receives the signal and chooses an action. Crucially, the signaller cannot act, while the receiver has no direct access to the relevant information – he only has access to what the signaller communicates. The action of the receiver determines the gains for both players. Although this kind of game imposes a form of *locutionary* cooperation, as the signaller has to communicate something to the receiver; since signaller and receiver are opponents they have no shared *perlocutionary* goals. Crucially, this competitive game aims to recreate a context where the receiver knows that the signaller is uncooperative even though the signaller is not explicitly opting out.

In the signalling game used in this experiment, the participants were receivers. In each round of the game they saw two cards: a winning card and a losing card. Their goal was to click on the winning card and avoid the losing card although they did not know which one was the winning

card. The signaller was a virtual player who gave a brief description of the winning card for each round. The receivers could read the description provided by the signaller in each round and use it to decide which of the two cards in the round was the winning card. In control items the description given by the signaller was literally true of one card and false of the other. In experimental items the description matched both cards, but it could give rise to a quantity implicature that matched only one card. The implicatures used were scalar implicatures and particularised quantity implicatures. Therefore, participants could infer that a deceitful signaller was lying in control items and communicating false implicatures in experimental items.

In a competitive signalling game a rational receiver should not expect the signaller to abide the first maxim of quality and consistently tell the truth. However, they may also not expect the signaller to consistently lie, as they might consider this strategy too predictable. In other words the receiver may also expect the speaker to double-bluff and tell the truth expecting not to be believed. Since trying to anticipate the speaker's lying, double-bluffing, double-double-bluffing and so on is potentially an infinite regress, Hespanha, Ateskan and Kizilocak (2000) argue that if the receiver does not know the strategy of the signaller and the signaller has complete control over the information communicated, the receiver should just ignore the speaker's signals. In the game presented here this means that in control items participant should choose at random between the two cards.

The receiver should also not expect the signaller to abide the first maxim of Quantity and so, according to the prediction derived from Grice, they should not derive quantity implicatures. Not deriving implicatures in experimental items makes the signaller's description simply uninformative and therefore participants in this situation should choose at random between the two cards. Whereas in the view suggested by the Epistemic Vigilance account that the receiver first interprets the signaller's utterance as if it were trustworthy and then possibly reject the content, participants should derive implicatures from experimental items and then possibly reject the content of the implicature. However, the same rational strategy that applies to lying and double bluffing in control items applies to experimental items. Since participants do not know the strategy of the signaller they should ignore their signals, including implicatures, and choose at random between the two cards. This poses a problem for the goal of studying implicatures in uncooperative settings because, whether participants draw implicatures or not in experimental items, they should choose at random between the two cards.

In order to address the problem outlined above, the present experiment involves two competitive conditions: a 'competitive condition' as described above, where the signaller can lie, and a 'competitive-truthful' condition, where the signaller is not allowed to lie. The aim of

the competitive-truthful condition is to have an uncooperative setting where the first maxim of quality (though not the first maxim of quantity) is *artificially* kept in operation. This condition addresses the problem outlined above in so far as participants will consider false implicature to be lies to some extent. If participants draw implicatures in this condition but they understand the term "lying" as applying only to asserted content, the possibility of inferring false implicatures would still push them to choose at random between the two cards, which would not solve the problem.

However, participants may consider false implicatures to be lies to some extent and therefore false implicatures should to be ruled out by the no-lying rule in the competitive-truthful condition. This expectation is based on existing experimental evidence that the lay understanding of "saying" often includes implicatures. Nicolle & Clark (1999) asked their participants to select paraphrases for what was 'said' in utterances that were used in contexts where they communicated a single strong implicature. They found that participants often selected the content of the implicature as a paraphrasis for what the utterance 'said'. Doran et al. (2012) asked participants to evaluate sentences as true or false in light of information that contradicted the content of the implicatures that the sentences could give rise to when used in context. For example, their participants might be asked to evaluate the utterance "John ate most of the cookies" after being told that in fact John ate all of the cookies. They found that participants judged the utterance that gave rise to the implicature as false half the time, indicating that they often incorporated the implicatures in the truth-conditional meaning of the sentences. Since the notion of lying is based on saying something false (Fallis, 2009), if participants in the present experiment tend to incorporate implicatures into 'what is said' they should also tend to see false implicatures as lies. For this reason if participants draw implicatures in the 'competitive-truthful condition', as the Epistemic Vigilance account would predict, they should trust their content and choose the card described by the signaller in so far as they believe false implicatures to be ruled out by the no-lying rule. Furthermore, if participants consider false GCIs to be lies to a larger extent than false PCIs, there should be a difference between these two types of inferences in the extent to which participants trust their content. However, if participants do not draw implicatures at all because the signaller is uncooperative, in accordance to the prediction derived from Grice, they should choose at random between the two cards as in the competitive condition.

In addition to the two competitive conditions there is a cooperative condition where the signaller's goal is to help the receiver. In this condition by all accounts participants should infer and trust the implicatures communicated by the signaller. Whether participants in this condition do draw implicatures will depend on the availability of the inferences, therefore this condition

will serve as a useful reference for the competitive conditions. Lastly, in all conditions it was measured how quickly participants' responded to each item in the experiment. Although the predictions discussed are relevant to the choice data, exploring reaction time data may enrich the interpretation of the choice data.

This experiment involves several factors that may affect the participants' behaviour and the interpretation of the results of this study. To help clearly define what these factors are and to help clarify the rationale of the design I will list and group these factors:

- Factors relating to the rules of the game (i.e. the experimental manipulation):
  - Cooperative vs. uncooperative: In two conditions the game is presented as competitive whereas in one condition it is presented as cooperative.
  - Lying allowed vs. not allowed: in two conditions participants are told that the speaker is allowed to lie whereas in the Competitive-truthful condition participants are told that the speaker is not allowed to lie.
- Factors relating to the characteristics of different items:
  - Encoded vs implicated content: in control items the key information is linguistically encoded whereas in experimental items it is implicated.
  - Scalars (GCI) vs Ad hoc (PCI): In half of the experimental items the key information is communicated through a scalar implicature and in the other half through an ad hoc implicature.
- Factors relating to the interpretation of the speaker's behaviour:
  - Lies vs false implicatures: participants may or may not consider false implicatures to count as 'lies'. What they think in this respect will determine whether implicatures are affected by the no-lying rule.
  - Strategies: participants in the competitive conditions will believe that their opponent to follow one of different possible strategies (e.g. consistent strategies such as always lying and always being truthful/double-bluffing, or variable strategies such as lying only half of the time).

4.2   Methods

4.2.1   Design

The game used in this paradigm is a signalling game. In each round of the game, the participants saw two cards. They knew that one was a winning card and the other a losing card, but they did not know which was which. A short description of the winning card was presented together with the two cards. Participants were told that the description was written by another player (the describer) who could see both cards and who knew which one was the winning card.

Participants were assigned to one of three conditions: competitive condition, competitive-truthful condition or cooperative condition. In all conditions they were told that they scored points by clicking on the winning card. Participants in the cooperative condition were told that the game was cooperative and the describer also scored points when they clicked on winning cards. Those in the two competitive conditions were told that the game was competitive and the describer scored points when they clicked on the losing card. Participants in the cooperative and competitive conditions were told that the describer was allowed to lie. Those in the competitive-truthful were instead told that the describer was not allowed to lie. The full instructions for each of the conditions is reported in **Appendix A1**.

4.2.2 Materials

Materials included 36 items (i.e. 36 rounds of the game): 16 experimental items and 16 control items. In control items the description was true of one of the two cards (i.e. the description matching card) and false of the other card. In experimental items the instruction was true of both cards but it could give rise to an implicature which was true of one card (i.e. the implicature matching card) and false of the other card. **Table 1** shows an example for each type of item. Control items used description of the form *On the winning card none/neither of the objects is an X* and one of the cards displayed no objects of type X (i.e. the description matching card) while the other card displayed at least one object of type X.

Experimental items could give rise two types of quantity implicatures: scalar implicatures and *particularised* quantity implicatures. The scalar implicature experimental items used description of the form *On the winning card some/most of the objects are Xs*, which could give rise to the implicatature that *not all of the objects are Xs*. On one of the cards associated with these items not all of the objects were Xs (i.e. the implicature matching card) while on the other card all of the objects were Xs. The *particularised* implicature items used desctions of the form *On the winning card the top/bottom/big/small object is an X*. Each of the four modifiers used in these items could give rise to a different implicature. For example, the description *the top object is a fork* could give rise to the implicature *the bottom object is not a fork* (see **Table 1**). The length of the words used to refer to objects in the descriptions was kept between 4 and 7 characters and the frequency of the object words was kept between the Zipf score values (The British National Corpus, 2007) of 3 and 5. The full list of the items used in the experiment is provided in **Appendix A2**.

Table 1 *Examples for each type of item*

Item Type	Description	Matching card	Other card

Control Neither	On the winning card neither of the objects is a rocket		
Control None	none of the objects are dice		
Experimental Scalar	most of the objects are ribbons		
Experimental Ad hoc	the top object is a fork		
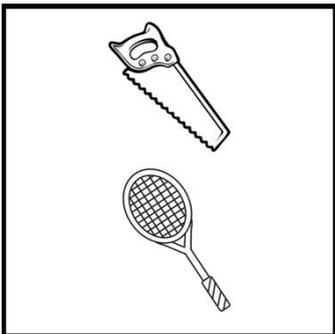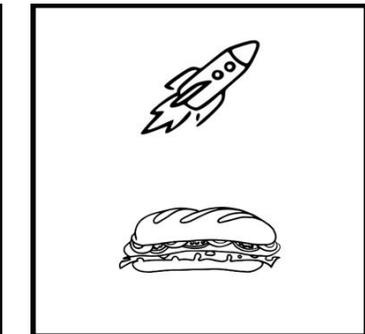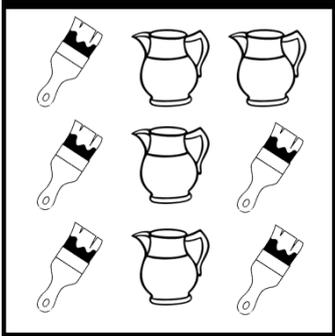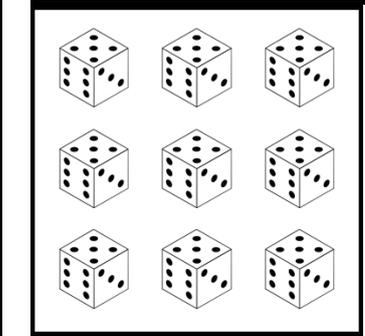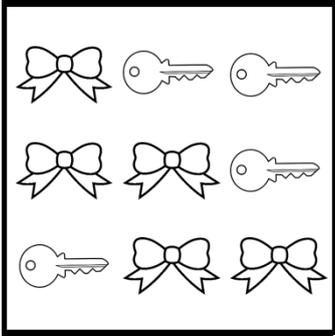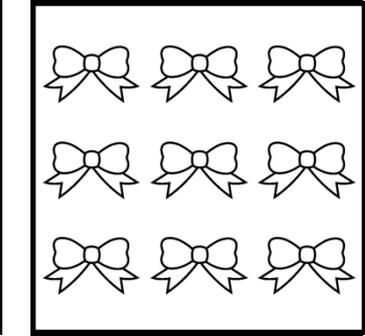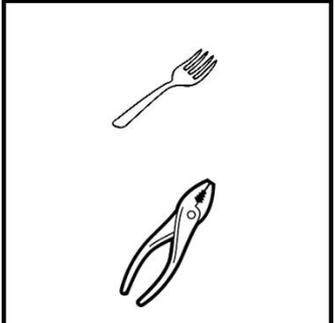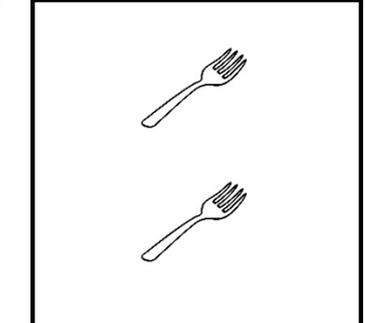
4.2.3   Participants and procedure

156 native English speakers (76 females, Mean age= 27.35) were recruited on the online crowdsourcing website Prolific.co.uk and they were directed to the Qualtrics website, where the experiment was hosted. The experiment was approved by the appropriate ethics committee at University College London and all participants gave their consent to take part in the research. Participants were told that they would play a guessing game with another player and that the study would take approximately five minutes. Participants were explained the rules of the game and they were shown an example of what the describer saw when writing the instruction (**Figure 1**)

On the winning card
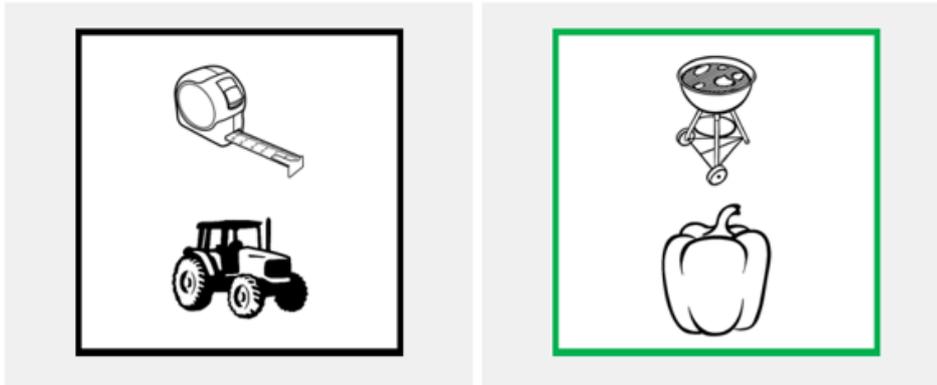
Figure 1 *Example of the describer view showed to participants*

After reading the instructions participants were asked two comprehension questions to check that they understood the instructions. Participants in the cooperative and competitive conditions were asked whether the describer's goal was to help them choose winning cards and whether their own goal was to click on winning cards. Participants in the competitive-truthful condition were asked whether the describer's goal was to help them choose winning cards and whether the describer was allowed to lie. Participants were presented all the items in a random order which was different for each participant. Each item or round in the game appeared to participants as a written description paired with two cards below the description. The position of the matching card changed randomly from left to right every time an item was viewed by a participant. In each round it was recorded whether the participant chose the description/implicature matching card or not and the time that elapsed between the presentation of the item and when the participant moved to the next item was measured using Qualtrics own web based reaction time measure. After the last round of the game participants were asked to rate on a 5-point scale (*never-sometimes-About half the time-most of the time-always*) the following two statements: (i) The describer was helpful; (ii) During the experiment I assumed that the describer was lying.

4.3 Results

4.3.1 Choice data

16 participants who answered incorrectly to the questions about the instructions were excluded from the analysis, which left the responses of 140 participants (68 females, Mean age= 27.51) for analysis. The percentages of description-matching responses to control items and implicature-matching responses to experimental items in each condition is shown in *Figure 2*.
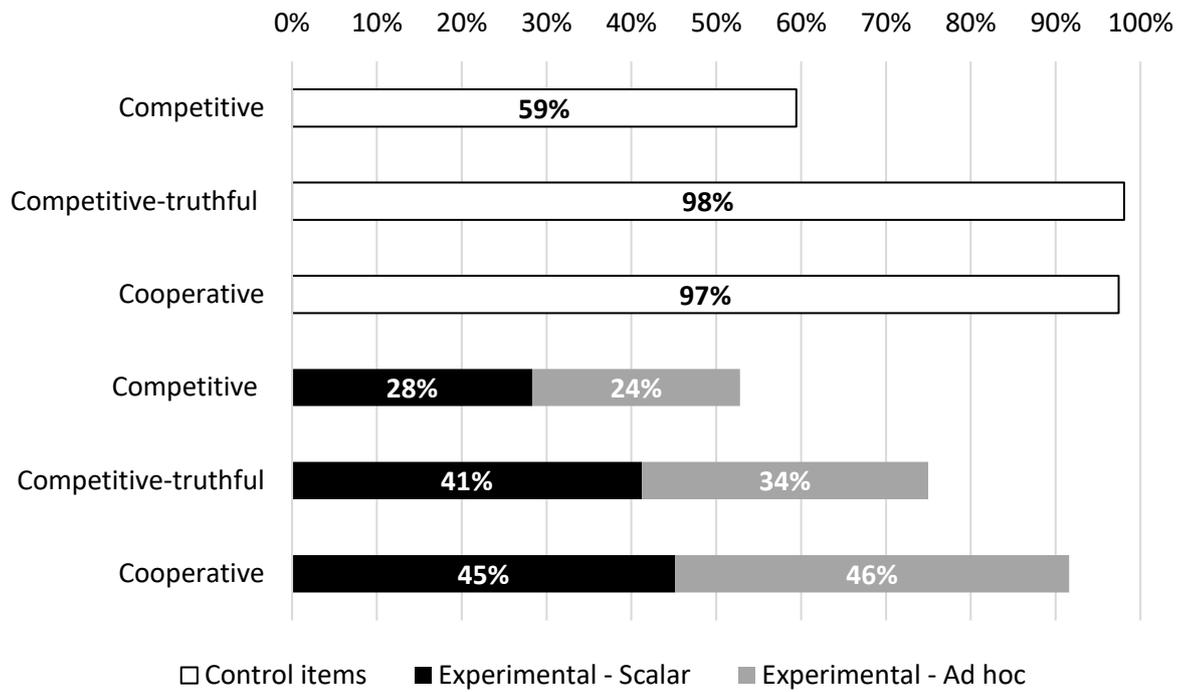
Figure 2. *Percentages of description-matching responses to control items and implicature-matching responses to experimental items in each condition*

In order to perform statistical analyses choice data was averaged by subjects and by items. The averages by subjects represent the likelihood of each participant to click on the description/implicature-matching card for each type of item (***Figure 3***). The averages by items represent the likelihood of each item to receive a description/implicature matching response for each condition. The tests reported below are Wilcoxon rank sum test for independent samples (W statistic) and Wilcoxon signed rank test for matched samples (V statistic).
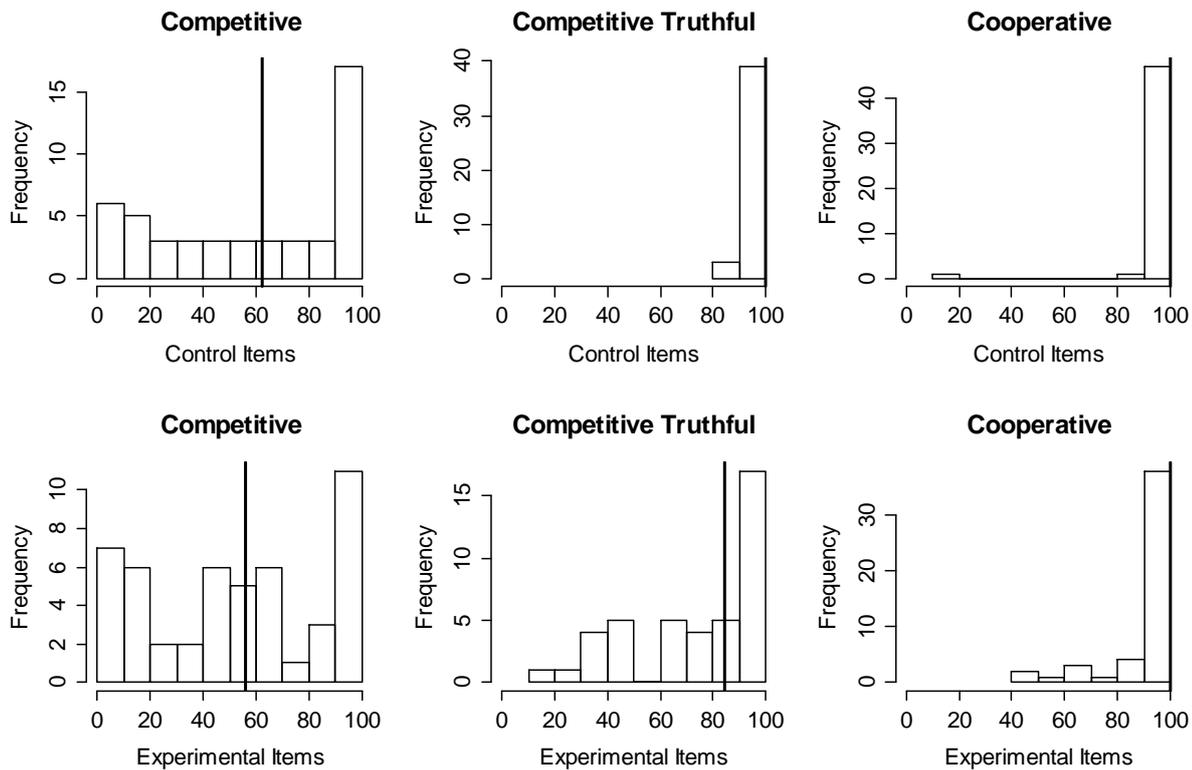
Figure 3. *Distribution of participants (histograms and medians) over the percentage of description-matching responses and implicature-matching responses given in the experiment by condition.*

Firstly I looked at control items and compared that rate of description-matching responses in each of the three conditions to chance level (p=0.5) to check whether participants were choosing at random between the two cards. The rate of description matching choices was significantly higher than chance both in the competitive-truthful condition ($W_{subjects}$=903, p<0.001; V_{items}=136, p<0.001) and in the cooperative condition ($W_{subjects}$=1223.5, p<0.001; V_{items}=136, p<0.001). For the competitive condition, the rate of description-matching responses was significantly different from chance by items (V=136, p<0.001) and the analysis by subjects showed a non-significant trend in the same direction (V=801, p=0.059). The distribution of individual participant performances on control items for the competitive condition in *Figure 3* shows that the two ends of the scale are the points of the scale with the highest frequency of participants. The two ends of the scale represent the two extreme strategies of either choosing the description-matching card on none of the trials or choosing it in all of the trials.

For control items, I checked whether there were any differences among the three conditions in the rate of description-matching responses. The rate of description matching responses to control items in the competitive condition was lower than in the cooperative condition ($W_{subjects}$=352.5, p<0.001; V_{items}=136, p<0.001), and in the competitive-truthful condition ($W_{subjects}$=1735.5, p<0.001; V_{items}=0, p<0.001). There was no significant difference between the

rate of description-matching responses to control items between the cooperative and the competitive-truthful conditions ($W_{subjects}=1122.5$, $p=0.269$; $V_{items}=52$, $p=0.436$).

Secondly, I looked at experimental items and again whether participants were choosing at random between the two cards by comparing the rate of implicature-matching responses in each of the three conditions to chance level ($p=0.5$). The rate of implicature-matching responses was significantly higher than chance in both the cooperative ($V_{subjects}=1221$, $p<0.001$; $V_{items}=136$, $p<0.001$) and competitive-truthful condition ($V_{subjects}=720.5$, $p<0.001$; $V_{items}=136$, $p<0.001$), but not in the competitive condition ($V_{subjects}=648$, $p=0.541$; $V_{items}=101$, $p=0.090$). Individual participant performances in the competitive condition for experimental items in *Figure 3* shows that although many participants cluster in the middle of the scale (i.e. roughly choosing the matching and non-matching cards in equal measure) many participants cluster at the two ends of the scale, indicating that they were adopting a consistent strategy.

I compared the rate of implicature-matiching responses in the competitive-truthful condition to the rates in the competitive and cooperative condition to see whether it was significantly different from either. The rate of implicature-matching responses in the competitive-truthful condition was significantly higher compared to the competitive condition ($W_{subjects}=1398$, $p=0.003$; $V_{items}=0$, $p<0.001$), and significantly lower compared to the cooperative condition ($W_{subjects}=601$, $p<0.001$; $V_{items}=136$, $p<0.001$).

Thirdly, within experimental items I compared the rates of implicature-matching responses of scalar items and Ad hoc items. In the cooperative condition, where a difference between the two rates may suggest a difference in the availability of the two types of inference, there was no significant difference between the two rates ($V_{items}=87$, $p=0.205$; $W_{subjects}=31$, $p=0.922$). In the competitive-truthful condition, where a differences between the two rates may suggest that the no-lying rule affected the two types of inferences differently, the rate of implicature-matching responses for scalar items was significantly higher than the rate of implicature-matching responses for the Ad hoc items ($V_{subjects}=285.5$, $p=0.004$; $W_{items}=62$, $p<0.001$).

Lastly, I looked at how participants in the three conditions responded to the final two questions about how often they thought that the other player had been helpful and how often they had lied (*Table 2*). Participants in the competitive condition thought that the describer was helpful significantly less often than participants in the competitive-truthful condition ($W=690$, $p=0.005$) and in the cooperative condition ($W=722.5$, $p<0.001$). There was no significant difference between the helpfulness ratings of the competitive-truthful and cooperative condition ($W=1192$, $p=161$). Participants in the competitive condition thought that the describer was lying significantly more often than participants in the competitive-truthful condition ($W=1745$,

p<0.001) and in the cooperative condition (W=2133.5, p<0.001). There was no difference between the lying ratings of the the competitive-truthful and cooperative condition (W=904, p=0.119).

Table 2. *Means and standard deviations for the helpfulness and lying ratings in the three conditions.*

	Competitive	Competitive-truthful	Cooperative
Describer was helpful	2.84 (SD=1.55)	3.76 (SD=0.75)	3.94 (SD=0.94)
Describer was lying	3.22 (SD=1.46)	1.38 (SD=0.82)	1.16 (SD=0.51)

It is interesting to note that in the distributions of responses over the scale points for the competitive condition for both the helpfulness rating and the lying rating (see **Figure 4**) participants cluster towards the ends of the scale rather than in the center. This distribution is consistent with the distribution participants over the frequency of matching-card responses in **Figure 3**.
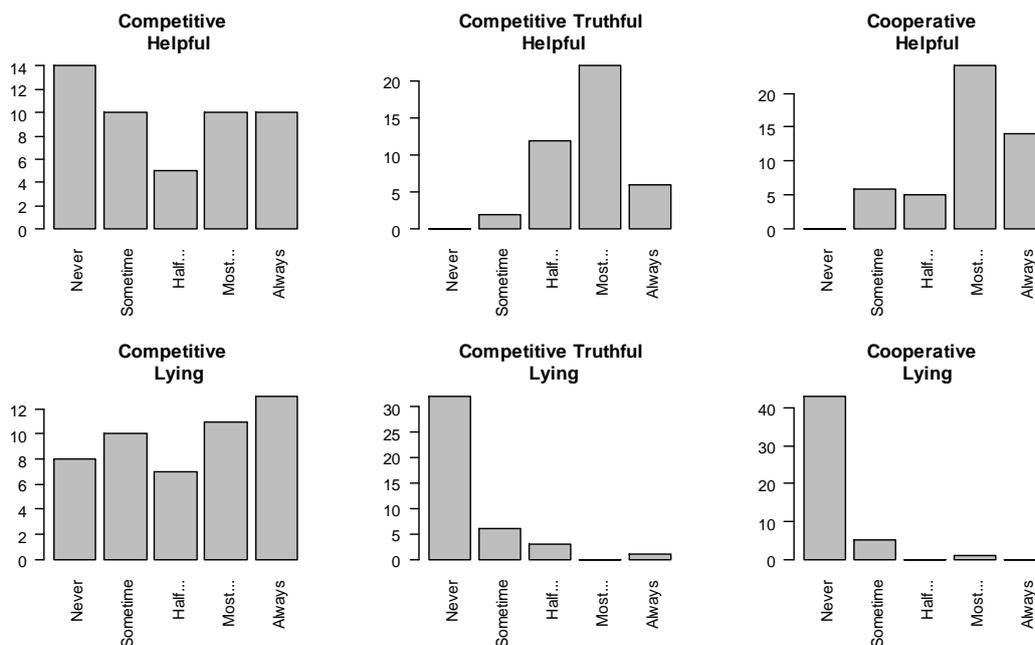
Figure 4. *Frequency of responses for each scale point of Helpfulness and Lying ratings by condition*

4.3.2 Reaction time data

The measure of reaction time is the time measured between item presentation and the time the participant moved to the next item. The raw reaction time data displayed large variance and a large number of outliers (**Figure 5Figure 5**).
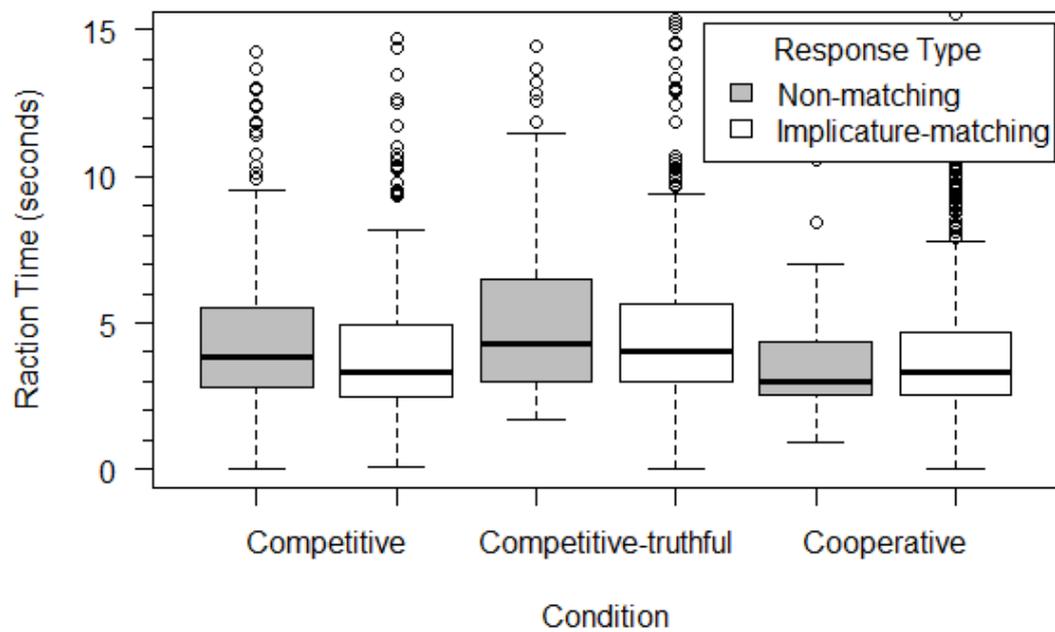
Figure 5. *Boxplot (median and quartiles) of experimental items reaction times by response type and condition (reaction times above 15 seconds are not displayed)*

For statistical analyses, I converted reaction times to their base ten logarithm in order to reduce the skewedness of the distribution. Extreme values were eliminated by removing values below or above two standard deviations from the overall mean reaction time.

The data was averaged by items and subjects separately for the two types or responses (i.e., description/implicature-matching response and non-matching response). This resulted in the subjects' dataset having a considerable amount of missing data because some subjects only gave one type of response for some categories of items. For example, in the cooperative condition many subjects consistently gave description-matching responses for all items and so they did not have reaction time data associated to a mismatch-response type. In conditions where the amount of missing data was low I used a method of mean imputation to perform the analysis. The method of choice is *mean substitution*, which consists in substituting the overall mean of all observations to the missing observations. This method was chosen because standard deletion methods (e.g. pairwise deletion) assume that the data is *missing at random* and this assumption does not hold in this case as missing data is associated specifically with participants who adopted extreme strategies (i.e. always choosing the matching card or always choosing the mismatching card). In conditions where the amount of missing data was high because many participants consistently gave the same type of response to all trials (e.g. in the cooperative condition) no comparison was performed. For all the comparisons presented, the homogeneity of variance assumption was checked using Levene's test.

Firstly, I compared the reaction time of description-matching responses (median raw reaction time = 3.42s) and non-matching responses (median raw reaction time = 4.03s) in control items

only in the competitive condition. In the competitive-truthful and cooperative conditions the rate of non-matching responses was too low to compare the reaction time of the two types of responses. For this comparison some participants lacked data points for one of the two response types. For instance, a participant who always chose the non-matching card would have no reaction time data associated with description-matching responses because they gave none. To remedy this, I substituted the overall mean reaction time to control items in the competitive condition (value = 0.58) to the missing values by subjects (16 data points out of 98). I compared the reaction times for the two responses with paired t-tests and found that participants were significantly faster in giving a description-matching response than in in giving non-matching responses by items (t(15)=3.37, p=0.004), but not by subjects (t(48)=0.25, p=0.8).

Secondly, I compared the reaction times for the two types of responses (i.e., implicature-matching and non-implicature-matching) in experimental items both in the competitive and in the competitive-truthful condition. In the cooperative condition the rate of non-matching responses was too low to compare the two types of responses. For these comparisons some subjects lacked data points for one of the two types responses for the same reason explained above. To remedy this, the mean reaction time to experimental items in the competitive condition (value = 0.57) was substituted to the 16 missing values in the competitive condition; and the mean reaction time to experimental items in the competitive-truthful condition (value=0.61) was substituted to the 13 missing values in the competitive-truthful condition. Participants in both the competitive and competitive truthful conditions were slightly faster in giving implicature-matching responses (Competitive condition Median raw reaction time = 3.35s; Competitive-truthful condition Median raw reaction time = 4.02s) compared to giving non-matching responses (Competitive condition Median raw reaction time = 3.82s; Competitive-truthful condition Median raw reaction time = 4.26s). For the competitive condition, that the reaction time to the experimental items was significantly faster for the implicature-matching responses compared to the non-matching responses by items (t(15)=4.53, p<0.001) but not by subjects (t(48)=0.94, p=0.353). For the competitive-truthful condition, there was no significant difference in reaction time between the two types of responses ($t_{subjects}$(41)=-0.28, p=0.777; t_{items}(15)=-0.15, p=0.880).

4.4   Discussion

The responses to control items allow to assess the effect of the experimental manipulation of the first maxim of quality. The aim of the manipulation was to suspend the first maxim of quality in the competitive condition but not in the other two conditions. Participants in all three conditions had a preference for the card matching the description. This preference is lower in the competitive condition than in the other two conditions and not as reliable due to the great

76

variability of strategies chosen by participants. The distribution of participants over the rates of matching card choices (**Figure 3**) shows participants clustering towards the two ends of the scale particularly in the higher end. This suggests that participants are not converging on a single strategy and in particular they do not seem to converge on the strategy of choosing at random between the two cards. Therefore the prediction that participants in the competitive condition would ignore the description and choose at random, following the strategy suggested by Hespanha, Ateskan and Kizilocak (2000), is not supported by the data. This result matches what participants indicated in their Lying ratings, namely that they did not think that the describer was lying half of the time. In fact the response that the describer was lying 'Half of the time' was the least chosen by participants in the competitive condition while the option that the describer was lying 'always' was the most popular. A considerable number of participants actually indicated that they thought the other player never lied or lied all the time. In contrast to this, most of the participants in the cooperative and competitive-truthful conditions indicated that they thought that the signaller never lied. In sum, it seems that although participants in the competitive condition did not cope with the possibility of deception by adopting the strategy expected, they mostly did not think that the maxim of quality was in operation. In contrast, participants in the both the competitive-truthful and cooperative conditions chose the description-matching on nearly 100% of trials. Therefore the hypothesis that participants in these two conditions believed that the signaller obeyed the first maxim of quality is supported by the data. In summary, it seems that the first maxim of quality was successfully suspended in the competitive condition whereas it was in operation in the competitive-truthful and cooperative condition.

Responses to experimental items in the cooperative condition are informative with respect to the availability of the implicatures used in the experiment. The preference for the implicature-matching in this condition was very high: roughly eight participants out of ten chose the card matching the implicature over 90% of the time. This indicates that the implicatures were highly accessible.

Responses to experimental items in the competitive condition should give an indication of what happens when the hearer does not believe that the speaker is obeying either the maxim of quality or the maxim of quantity. In this condition there was no evidence of an overall preference for either matching card or non-matching card. As discussed this result is compatible with two distinct possibilities: the possibility that participants not drawing implicatures at all and the possibility that participants are drawing implicatures but not trusting their content. However, if participants were not drawing implicatures at all their individual strategies should cluster unimodally around chance level (i.e., 50% in figure **Figure 3**). In other words, each individual

77

participant should choose at random between the two cards. Instead, the preferences of participants in the competitive condition seem to follow a three-modal distribution with a considerable number of participants who either consistently choose or consistently avoid the implicature-matching card. Therefore, the results of this condition suggest that at least some participants derived implicatures from the speaker's utterances and based their choices on them either by avoiding or choosing implicature-matching card. However, as discussed in the presentation of the study's design, the competitive-truthful condition should give a clearer indication of whether participants are drawing implicatures from the utterances of the speaker.

In the competitive-truthful participants should believe the uncooperative speaker is forced to obey the first maxim of quality but not the maxim of quantity. In this condition participants preferred the implicature-matching card. This in itself suggests that participants are actually drawing implicatures from the signaller's utterances. However, even if participants were drawing implicatures, the extent to which participants preferred an implicature-matching response in this condition ultimately depends on whether they chose to believe the content of the implicatures. As discussed, participants may to some extent incorporate the meaning of the implicature into the truth-conditional meaning of the descriptions (Doran et al., 2012). False implicatures that are incorporated into the truth-conditional meaning of the utterance should be seen as lies. This should push participants to trust the content of the implicatures in the competitive-truthful condition because of the no-lying rule attached to this condition. If participants completely equated false implicatures with lies, the preference for the implicature-matching card in the competitive-truthful and in the cooperative condition should be roughly equal, in the same way that in control items the rate of description-matching responses is roughly equal between these two conditions. The results of this experiment indicate that this is not the case as the preference for the implicature-matching card in the competitive-truthful condition was lower than in the cooperative condition. In summary, the results of the competitive-truthful condition suggest that participants inferred implicatures from the descriptions and the no-lying rule pushed them to trust the content of the implicatures more than participants in the competitive condition, but mostly they did not completely equate falsely implicating with lying.

The proposed interpretation of the results is therefore that hearers derive implicatures from the utterances of uncooperative speakers who are not explicitly opting out and then decide whether to trust the content of the implicatures. This conclusion is inconsistent with the prediction derived from Grice (1989) that hearers should not derive implicatures from uncooperative speakers. The implicatures that could arise from the signaller's utterances should have been afforded by the *perlocutionary* goal of having the receiver guess the correct card. Receivers in

the competitive conditions should not have believed the signaller to share the *perlocutionary* goal to have them guess the correct card. This interpretation rests on the assumption that the study was successful in creating a context where participants believed that the signaller was *perlocutionarily* uncooperative with respect to this goal even though they were forced to be *locutionarily* cooperative (and truthful in the case of the competitive-truthful condition) by the game. The proposed interpretation is instead consistent with the view that there is a dissociation between comprehension and acceptance (Sperber et al. 2010; Mazzarella, 2015a) such that hearers infer the implicatures of untrustworthy speakers independently and in parallel to their assessment of whether the speaker is trustworthy. As discussed in section 3.2.1, a slight modification could make Grice's account consistent with the present findings and their proposed interpretation. If the hearer's expectation was not that the speaker *does* observe the CP but merely that they *appear* to be observing the CP the hearer should derive implicatures from the utterances of uncooperative speakers who they believe want to *appear* cooperative. This modification would introduce the same kind of dissociation between comprehension and acceptance in Grice's account that is advocated by the account of Epistemic Vigilance (Sperber et al., 2010; Mazzarella, 2015a). Hearers would interpret utterances guided by an assumption of *apparent* cooperation (i.e. as if they were trustworthy) and this would make their epistemic assessment of utterances independent from their interpretation.

In the competitive-truthful condition the rate of implicature-matching responses for scalar items was higher than the rate for Ad hoc items. There was no evidence of this difference in the cooperative condition. This difference between the two types of inference can be interpreted as participants being more likely to consider false implicatures arising from lexicalised scales to be lies than false implicatures arising from Ad hoc scales. The fact that this difference is not present in the cooperative condition rules out the alternative explanation that this difference is due to the different availability of the two types of inference (i.e., due to scalar implicatures being more accessible than particularised implicatures in the items). This finding is consistent with the results of Doran et al. (2012) who found that some types of implicatures tend to be incorporated into the truth-conditional meaning of an utterance more than others. Doran and colleagues however do not directly compare implicatures arising from lexicalised scales and Ad hoc scales. Even more relevant to this finding of the present experiment is Katsos's (2009) finding that using an under-infomative utterance containing an element of a lexicalised scales (e.g. saying that *the elephant pushed some of the trucks* when the elephant pushed all of the trucks) can be considered a more severe violation of informativity than using an under-informative containing an element of an Ad hoc scale (e.g. saying *the dog painted the triangle* when the dog painted both the triangle and the heart). In fact, the interpretation of Katsos'

results may also be recast in terms of the relationship between false implicatures and lies. If his participants were inferring false implicatures from the under-informative utterances they were asked to judge, they may have judged false scalar implicatures more harshly than false Ad hoc implicatures because they perceived false scalar implicature to be closer to being actual lies than false Ad hoc implicatures. If the proposed interpretation is correct, whether an implicature is based on a lexicalised scale or an Ad hoc scale is a factor that can affect whether an implicature is perceived as a lie or not. Meibauer (2014) proposes that it is the relevance of the implicature in context that affects whether a false implicature is perceived to be a lie. However, in the present experiment scalar implicatures and Ad hoc implicatures were equally relevant and equally necessary for deciding which was the winning card. Therefore, if relevance was the only factor influencing whether an implicature is considered a lie, there should have been no difference between scalar implicatures and Ad hoc implicatures in this respect. In summary, the result that participants in the competitive-truthful condition were more likely to trust implicatures arising from lexicalised scales than implicatures arising from Ad hoc scales can be interpreted as indicating that false scalar implicatures are more likely to be lies than false Ad hoc implicatures independently of the relevance of the implicatures in the context.

The helpfulness and lying ratings showed that participants in the competitive condition were more likely to think that the participant was lying and not helpful compared to participants in the other two conditions. In both the competitive-truthful and the cooperative condition the average lying rating was very low and the average helpfulness rating high. There was no evidence of a difference in the ratings between the cooperative and competitive-truthful conditions. The pattern of results for the lying rating is consistent with the expectation that participants would not think that the signaller was lying in the competitive-truthful condition, where lying was not allowed, or in the cooperative condition, where the signaller had no motivation to deceive. In the competitive-truthful condition, even though the signaller was playing against them, participants indicated that they mostly thought that the signaller was helpful. This result may be due to the fact that question about the helpfulness of the describer invited an interpretation along the lines of "How often was the description produced by the describer helpful in choosing the winning card". If most participants interpreted the question this was, the very high rating of helpfulness in the competitive-truthful condition is not surprising.

Reaction time data was affected by extreme subject variance. In particular, the differences between participants in terms of how long they spent on each trial were extremely large, with some participants taking on average only two seconds to respond to an item and other participants taking on average over ten seconds. These extreme individual differences are most

likely due to the strategic element of the task. Different participants may have had very different approaches to the game in the experiment and while some of them may have wished to play 'quickly', others may have reflected on their moves for longer. In the same way that a chess player might think for hours or just a few seconds before making a move. Partly because of the large variance, the differences found were only significant by items. Therefore, the interpretation and value of the reaction time results must be taken with a pinch of salt.

In the competitive condition the description-matching responses to control items and implicature-matching responses to experimental items were faster than non-description-matching responses and non-implicature-matching responses respectively. One explanation for these two findings is that rejecting communicated content takes longer than accepting it as suggested in previous studies (Bott & Noveck, 2004; Mazzarella, Trouche, Mercier & Noveck, 2016). One consequence of this explanation is that participants must be inferring implicatures in the competitive condition for the delay of the non-implicature-matching response to be caused by the extra time taken to reject the content of the implicature. This conclusion is not problematic as it is consistent with the interpretation of the choice data proposed above. However, no firm conclusions from the reaction time data of the present experiment because of the very large variance. Furthermore, the reaction times in this type of paradigm are likely to reflect a variety of different processes related to the strategic element of the game. This may mean that reaction times are simply unsuited to investigating the phenomena that the present study addresses at least in combination with the type of paradigm used here.

In conclusion, I interpret the results of the present study to support two propositions. The first propositions is that listeners infer the implicatures arising from the utterances of an uncooperative speaker then decide whether to believe the content of the implicatures or reject it. This conclusion is in support of the views of Sperber et al. (2010) which is that hearers interpret utterances (including implicatures) of untrustworthy speakers as if they were trustworthy and then decided whether to believe the content of the implicature. This proposition is instead inconsistent the prediction derived from Grice (1989) that hearers should not draw implicatures from a speaker that they believe to be uncooperative. The second proposition is that false implicatures arising from lexicalised scales are perceived to be closer to lies than false implicatures asising from Ad hoc scales independently from the relevance of the implicatures in question. This conclusion is also consistent with previous research showing that some implicatures are more likely than others to be incorporated into the truth-conditional meaning of the utterance (Doran et al. 2012) and with the fact that under-informative lexicalised scalar expressions seem to be judged more harshly than under-informative Ad hoc scalar expressions (Katsos, 2009). The proposed interpretation of this study's results is therefore in

support of the view that there is a dissociation between comprehension and acceptance of implicatures which are supposed to be parallel and independent processes (Sperber et al., 2010; Mazzarella, 2015a). However this study does not investigate this dissociation directly as the paradigm employed does not measure the outcome of these two processes separately. In the next chapter I will present a paradigm that aims to investigate the outcomes of the hearer's interpretation and epistemic assessment separately.

5 Scalar implicatures in non-cooperative contexts: comprehension and acceptance

In the previous chapter I presented a study in which a non-cooperative setting was operationalised as a competitive signalling game. The results of the study suggested that listeners infer implicatures from uncooperative speakers. The proposed interpretation of the results is consistent with the view that there is a dissociation between comprehension and acceptance of implicatures (Sperber et al., 2010; Mazzarella, 2015a). However the study did not investigate this dissociation directly. In this chapter I present a study which also investigates whether listeners infer implicatures from uncooperative speakers and in particular the study focuses on the scalar implicature of *some*. The study presented here aims to measure the outcome of the hearer's interpretation and epistemic assessment separately.

5.1 Introduction

Communication is often regarded as a cooperative phenomenon (e.g. Grice, 1989; Clark, 1996). However, conflict is an important element of human interaction and conversation also takes place in non-cooperative contexts. As discussed in section 3.2, if interlocutors are not willing to invest the effort in at least producing and interpreting utterances, in other words if they are not at least *locutionarily* cooperative, communication cannot happen at all. So in order to study communication in non-cooperative contexts I will consider contexts where interlocutors are *locutionarily* cooperative but *perlocutionarily* uncooperative, that is to say they are not cooperating on goals outside of the conversation. In the study presented here the non-cooperative context is operationalised as a situation in which it is clear to the hearer that the speaker is not perlocutionarily cooperative.

5.1.1 Grice

In Grice's (1989) account, the inference of implicatures is based on the expectations that hearers have about how a rational cooperative speaker will behave in working towards this common goal. For example, imagine you are participating in a treasure hunt in which the team that finds the most prizes wins. The prizes are hidden in a garden, but you arrived late and missed the initial instructions so you do not know where you should look. A member of your own team tells you that *some of the prizes are hidden underground*. She could have been more informative by saying that *all* of the prizes are hidden underground, and you know she was present from the start so she probably knows whether this alternative statement is true. Since your team mate is working towards the same goal as you (i.e., for your team to find the most prizes) you can assume that she is not deliberately withholding information that would decrease your chances of your team winning (i.e., she is abiding by the first maxim of quantity). Therefore, the reason

why she did not say that all the prizes are hidden underground must be that it is not true. In order to derive quantity implicatures like the one in this example, the hearer must assume that the speaker is being cooperative and providing *enough* information for the purpose of the conversation.

In Grice's account implicatures often correspond to what the hearer needs to assume in order to preserve the belief that the speaker is being cooperative when they are violating one of the maxims. However, as discussed in section 3.2.1, the speaker can also explicitly *opt out* of a maxim or the CP altogether. In these situations the violation of a maxim on the part of the speaker will not result in the hearer inferring an implicature because the hearer does not need to preserve the assumption that the speaker is being cooperative. In these situations it is mutually manifest to both hearer and speaker that the speaker is not willing to cooperate.

In some situations the hearer may believe that the speaker is unwilling to cooperate with respect to a *perlocutionary* goal but they may not be explicitly opting out. As discussed in section 3.2.1, in these situations Grice's account should predict that in interpreting the speaker's utterances, the hearer does not infer implicatures that are afforded by the *perlocutionary* goals that the speaker is unwilling to cooperate on. Returning to the treasure-hunt example, imagine the same situation described above in which you are participating in a treasure hunt and you do not know where you should look for the prizes. This time a member of the opposing team, instead of a team mate, tells you that *some of the prizes are hidden underground*. Again, they could have been more informative and said that *all of the prizes are hidden underground*. However, the member of the opposing team is unlikely to share the goal that you know where to look for the prizes and so you cannot assume that they are being as informative as possible in pursuit of the goal that your team wins. You would derive the implicature that *not all prizes are hidden underground* if you needed to preserve the assumption that the speaker is cooperative even though they said something under-informative. Since you already assume that they are uncooperative, you have no reason to derive the quantity implicature.

Because of the prediction that hearers should not derive implicatures from uncooperative speakers, Grice's account is not suited to explaining situations in which a distrustful hearer infers that a speaker that they believe to be uncooperative is communicating a false implicature (see section 3.2.3). A false implicature is an attempt to communicate something false with the intention to deceive by means of a conversational implicature (Meibauer, 2014). In social settings speakers may have good reasons to prefer implicating something false rather than saying something false. For example, speakers may not wish to be *caught lying* and they may want to take advantage of the plausible deniability offered by implicatures. According to

Strategic speaker theory (Pinker, Nowak & Lee, 2008; Lee & Pinker, 2010) we are especially sensitive to the fact that we cannot deny having said something that was linguistically encoded in our utterance, while we may deny intending to communicate something that we merely implicated. In the treasure-hunt example above, the member of the opposing team may know that *all the prizes are hidden underground*, but they may choose to say that *some* of them are because they intend their hearer to infer the false implicature that *not all* of them are hidden underground. The prediction derived from Grice is that the hearer should not infer an implicature in this situation.

5.1.2 Relevance Theory and Epistemic Vigilance

In relevance theory the inference of implicature does not depend on an assumption of cooperation but on an assumption of relevance. Relevance Theory (Sperber and Wilson, 1995 p.270) the hearer assumes that the speaker's utterance is *optimally relevant,* which means that it is *at least relevant enough to be worth the hearer's effort to process it and it is the most relevant one compatible with the speaker's abilities and preferences*. This does not mean that whether the speaker is cooperative or uncooperative does not affect the derivation of implicatures, but only that the speaker's cooperation is a contextual factor among others rather than the fundamental assumption that affords implicatures (see section 2.2.3.2). For example, a hearer may not infer quantity implicatures from the utterance of a speaker who is being manifestly underinformative as these would cause the final interpretation of the utterance to be incompatible with the speaker's *preferences*.

Relevance Theory and the account of Epistemic Vigilance (Sperber et al., 2010) make different predictions than Grice regarding implicatures in situations where the hearer believes the speaker to be uncooperative even though the speaker has not explicitly opted out. Specifically, according to Sperber and colleagues hearers interpret the utterances of untrustworthy speakers as if they were trustworthy. This means that hearers will construct an optimally relevant interpretations of the untrustworthy speaker's utterances including any implicatures of the utterances. In parallel to this, hearers perform an epistemic assessment of the utterance content and of the speaker as a source of information. If the speaker is judged to be an untrustworthy source the hearer will reject what the speaker has communicated to them, including any implicatures. Therefore, Sperber and colleagues predict that hearers infer implicatures from the utterances of untrustworthy speakers and then possibly reject their content depending on the outcome of their epistemic assessment.

This view of the comprehension process and epistemic assessment as two parallel and independent processes (Sperber et al., 2010; Mazzarella, 2015a) makes the Epistemic Vigilance

account well suited to explaining how a distrustful hearer may infer false implicatures from the utterances of an untrustworthy speaker and no be deceived. Returning to the treasure-hunt example where a member of the opposing team says that *some of the prizes are hidden underground,* the hearer may construct an optimally relevant interpretation of the speaker's utterance which includes the implicature that *not all* of them are hidden underground. In parallel to this, the hearer is assessing the speaker's trustworthiness and if they judge the speaker to be untrustworthy they may choose not to believe the content of the speaker's implicature or the content of the whole utterance (if they think the speaker may be lying).

5.1.3   Distinguishing comprehension and acceptance of implicatures

In scenarios where the speaker is uncooperative and their unwillingness to cooperate is known by the hearer but not mutually manifest (i.e., the speaker is not explicitly opting out) Sperber and colleagues predict that the hearer should infer implicatures and then reject their content, while the prediction derived from Grice is that the hearer should not derive implicatures at all. Although the eventualities captured by these predictions are clearly different, conflating the measurement of the interpretation and epistemic assessment makes the difference between the two undetectable (Mazzarella, 2015b). Imagine that in the example where an uncooperative/distrustful speaker says that *some of the prizes are hidden underground* the hearer is asked whether they think it's possible that *all* of the prizes are hidden underground. A participant who has not drawn any implicature from the utterance will answer *yes* and a participant who has inferred the *not-all* implicature but rejected its content will answer in the same way.

As Mazzarella (2015b) points out, an example of this kind of conflation between interpretive process and epistemic assessment is found in a series of experiments on the effects of politeness on the inference of scalar implicatures (Bonnefon, Feeney & Villejoubert, 2009; Bonnefon, Feeney & De Neys 2011; Feeney & Bonnefon, 2013). Bonnefon and colleagues embedded scalar expression in face-threatening acts, which according to politeness theory are damaging for the positive self-image of the hearer (Brown and Levinson, 1987), and in face-boosting acts (i.e., boosting the hearer's positive self-image). Their experiments aimed to test the hypothesis that the scalar implicature of utterances containing the expression *some X-ed* (i.e. not all *X-ed*) would be less available if *X* was face-threatening for the hearer than if it was face-boosting. For example, in one of the scenarios used in their experiments (see **Table 3**) participants first read a context where they imagine that they just gave a speech and then a character tells them either that *some people hated your speech* (face-threatening scenario) or that *some people loved your speech* (face boosting scenario). Participants then answer a question on whether they thought that the stronger alternative (i.e., *that everybody hated/loved your speech*) was possible.

87

Bonnefon and colleagues found that participants were more likely respond in the affirmative to this question in the face-threatening condition and they interpreted this result in support of the hypothesis that face-threatening predicates made the scalar implicature less available. Bonnefon, Feeney and De Neys (2011) used a similar task and also measured how long it took participants to read the target sentence and answer the question, which were presented on the same page. They found that in face threatening contexts participants took longer to answer that the stronger alternative is possible (i.e. *yes* answer), which is incompatible with the content of the implicature. They interpreted this finding to indicate that politeness blocks the derivation of the scalar implicature and makes arriving at the semantic interpretation of the utterance slow and effortful.

Table 3 Speech scenario used in Bonnefon, Feeney & Villejoubert (2009)

Context	Imagine you gave a speech at a small political meeting. You are discussing your speech with Denise, who was also there. There are 6 other people in the audience that day. You tell Denise that you are thinking about giving the same speech to another group.
Target sentence	Hearing this Dense tells you that 'Some people hated [loved] your speech.'
Question	Given what Denise tells you, do you think it's possible that everybody hated [loved] your speech? [YES] [NO]

Mazzarella (2015b) took issue with the interpretation of these results given by the Bonnefon and colleagues. She argued that the face-threatening context may have affected the believability of the scalar inference and not its derivation. This is because the measures collected by Bonnefon and colleagues could not distinguish between the derivation and believability of the inference. The question asked in the experiments of Bonnefon and colleagues measured only the outcome epistemic assessment of the content of the implicature and the reaction time measured comprised both the derivation and the assessment of the scalar inference (target sentence and question were presented together). Mazzarella, Trouche, Mercier and Noveck (2016) conducted two experiments using one of the scenarios by Bonnefon and colleagues and the same paradigm, with the only difference that they presented the target sentence and the question sequentially, rather than at the same time. This adaptation allowed them to collect separate reaction times for the derivation of the scalar inference, which should happen when participants are interpreting the target utterance, and for the epistemic assessment, which should happen when participants answer the question. As in the experiments of Bonnefon and colleagues, they found that participants were more likely to say that the stronger alternative

was possible in the face-threatening condition compared to the face-boosting condition. In analysing the reaction times, they found that in the face-threatening condition participants were slower to say that the stronger alternative was possible (*yes* answer) than to say that it was not possible. However, they did not find that participants who answered *yes* and participants who answered *no* differed in the time it took them to read and interpret the utterance. They interpreted these results to indicate that face-threatening contexts do not affect or *block* the inference of implicatures, but they make the content of the implicatures more likely to be rejected. They propose that the latencies in reaction time found in their own experiment and in the experiments conducted by Bonnefon and colleagues are not due to the effort of *processing politeness* but they are due to participants taking longer to reject the content communicated in the implicature than to accept it.

5.1.4   The present experiment

The present experiment addressed the question of how talking to a speaker who is not cooperative and not explicitly opting out affects the inference and acceptance of the implicatures that may arise from the speaker's utterances. In particular this study focuses on the scalar implicature of *some*. The paradigm used in this study is similar to the one used by Bonnefon, Feeney and Villejoubert (2009) and by Mazzarella and colleagues (2016) where participants read a short vignette which introduces a character and then read a target sentence which reports an utterance of the character containing the quantifier *some* and that could give rise to a *not-all* implicature. However, while their experiments manipulated whether the target utterance was face-threatening or face-boosting, the present experiment manipulates whether the character who utters the target sentence is described as cooperative or competitive in the context preceding the target utterance. Therefore, in this manipulation the target utterance which can give rise to the scalar implicature is the same in the cooperative and competitive condition, while the context that precedes it is different. In the cooperative context the character shares the goal of the addressee while in the competitive context the character and the addressee have conflicting goals and the participant is told that the character wants to prevent them from achieving their goal. The information that the character is uncooperative is presented as private and not as part of the common ground. In both conditions, the content of the implicature which may arise from the target utterance is a crucial for the addressee's goal and part of an *optimally relevant* interpretation of the utterance.

Similarly to the experiment of Mazzarella and colleagues, the present experiment aims to distinguish between cases where the hearer does not infer any implicature from cases where they infer an implicature and reject its content. Therefore, in this experiment as in theirs, the utterance which may give rise to an implicature (i.e. some and *not all*) and the question asking

for an epistemic judgement on the content of the implicature (i.e. asking whether *all* may be the case) are presented on separate pages and separate reaction times are collected for the two pages. In addition, after the epistemic judgement question, the present experiment includes a question asking participants whether they thought that the speaker intended to communicate the implicature. This second question was added because it directly probes the participants' interpretation of the speaker's meaning.

As discussed in sections 5.1.1, for the kind of non-cooperative contexts used in this experiment Grice's account can be interpreted to predict that hearers do not derive implicatures from the uncooperative speaker's utterance. Relevance Theory instead predicts that hearers derive all the implicatures necessary to arrive at an optimally relevant interpretation of the utterance and then possibly reject their content if they consider the speaker to be untrustworthy (see section 5.1.2). Both accounts predict that when asked whether *all* may be the case (i.e. epistemic judgement), hearers will be more likely to answer that *all* is possible, which is inconsistent with the some and *not-all* implicature, if the speaker is uncooperative than if the speaker is cooperative. While for Grice this is due to the *not-all* implicature of the uncooperative speaker not being inferred at all, for Relevance Theory this is due to the *not-all* implicature being inferred and then rejected. Therefore, while Grice's account would predict that hearers are less likely to indicate that they thought that the speaker intended to communicate an implicature if the speaker was uncooperative than if they were cooperative, Relevance Theory predicts no such difference. Lastly, if hearers are inferring and rejecting the content of the implicature, they are also likely to think that the speaker is trying to mislead them.

It is important to note that Grice's account is compatible with scenarios where a hearer derives and afterwards rejects a false implicature only on the condition that the hearer believed the speaker to be cooperative when they inferred the implicature. For example, in the treasure-hunt example the hearer may infer an implicature from the speaker's utterance that *some of the prizes are hidden underground* and then reject the content of the implicature once they realise that the speaker is a member of the opposing team. In this particular case the predictions of Grice and Relevance Theory would not differ. In order to keep the predictions of the two accounts separate, in the present experiment participants encounter the utterance which could give rise to the scalar implicature only after reading the context which presents the speaker as either cooperative or uncooperative.

Lastly, the materials used in this experiment are constructed so that they rule out the possibility that the speaker may be lying. This is because the possibility of lying interacts with the epistemic assessment of the content of the implicature and complicates the study of the latter. For

example, if in the treasure hunt scenario the hearer thought that the player who said that *some of the prizes are hidden underground* lied, and actually *none* of the prizes are hidden underground, the content of the scalar implicature (i.e. that *not all* the prizes are hidden underground) is necessarily true. As a result, the epistemic assessment of the implicature becomes secondary to the epistemic assessment of *what is said*. In order to avoid this eventuality the present experiment uses scenarios where the information explicitly communicated by the target utterance is presupposed and not at issue in the context. In contrast, the information conveyed by the scalar implicature is crucial and not presupposed in the context.

5.2   Methods

5.2.1   Materials and design

The materials used in the experiment included five short stories (between 70 words and 100 words in length) in which the reader is given the perspective of a character. The reader's character needs to know a piece of information and they ask another character in the story. At the end of the story (in the target sentence) the other character in the story utters an under-informative utterance containing the expression *some* which may give rise to the implicature *some and not all*. Crucially, '*whether all'* is the relevant information that the participant's character needs. Each story had a cooperative version where both the character and the participant are interested in the participant acquiring the key information, and a non-cooperative version where the character would benefit from the participant being ignorant or misinformed about the key information. One of the five contexts used in the experiment is reported in **Table 4** as an example (see Appendix B1 for all the scenarios used in the experiment).

In all five stories, the target utterance is the same for both versions. Each story is followed by three yes-no questions: an *epistemic* question (i.e., whether the reader thinks that 'all' is possible given what the character has said), a *meaning* question (i.e., whether the reader thinks that the character meant to communicate 'not all') and a *deception* question (i.e., whether the reader thinks that the character was trying to mislead them).

Table 4 *'Company Exam' Story*

	Non-cooperative version	Cooperative version
Story	Imagine that you are about to sit a competitive exam for a particular position in your company. The exam has multiple choice and open answer questions. You don't remember whether you need to answer all the open answer questions in order to pass. You ask the person who has just sat the exam before you what she knows about the open questions. There is only one position opening and it's very well paid. Therefore you know that she probably hopes that you fail the exam so that she has a better chance of getting the job.	Imagine that you are about to sit an exam for a particular qualification in your company. The exam has multiple choice and open answer questions. You don't remember whether you need to answer all the open answer questions in order to pass. You ask the person who has just sat the exam before you what she knows about the open questions. The company has promised bonuses to all the employees if enough people pass the exam. Therefore she probably hopes that you both pass the exam.
Target Sentence	She says that 'Some of the open answer questions must be answered'.	
Epistemic Question	Given what she told you, do you think it's possible that all of the open questions must be answered?	
Meaning Question	Do you think she meant that you don't need to answer all the open question?	
Deception Question	Do you think she was trying to mislead you?	

5.2.2 Participants and Procedure

425 Native English speakers (236 females, Average age=28.64) were recruited on the crowdsourcing website Prolific Academic and directed them to the Qualtrics website, where the experiment was hosted. Participants were told that they would first read a short story and then they would be asked a few questions about the story.

Each participant read only one story and they were randomly allocated to read either the cooperative or the non-cooperative version of the story. When the participant finished reading the story, which was presented on one page, they clicked on the next page button and they read the target sentence. After reading the target sentence they moved to three other pages displaying the three yes-no questions. The experiment collected participants' responses to the three questions and measured the time it took each participant to read the target sentence and the time it took them to answer the epistemic question using Qualtrics' own web-based reaction time measure.

5.3   Results

The frequencies of yes responses to the three questions following the story was analysed using Chi squared tests (see data in Error! Reference source not found. and **Table 5**, see **Appendix B2** for a full breakdown of raw frequencies).
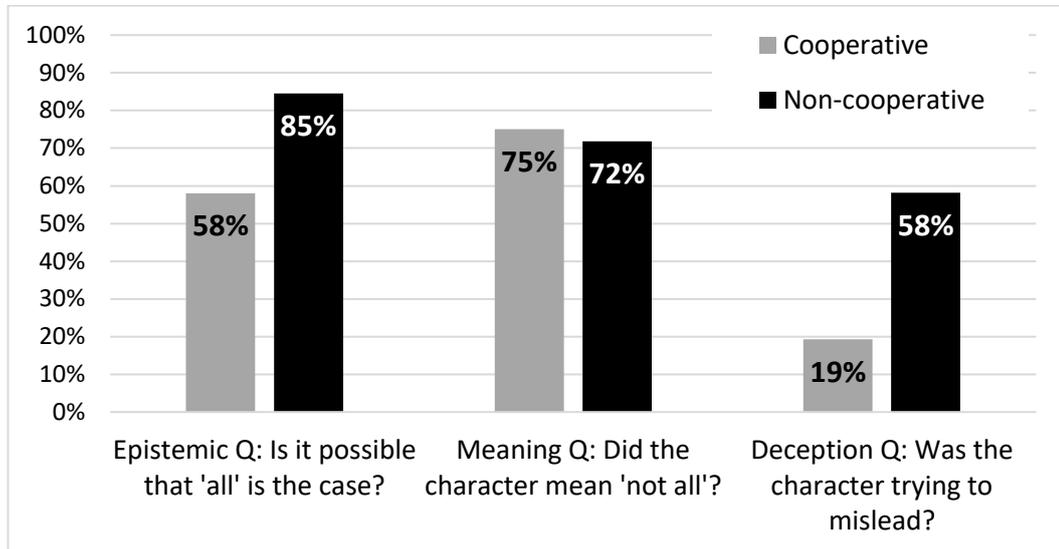
Figure 6 *Frequency of yes responses to each of the three follow up questions*

In the cooperative condition 58.02% of participants responded *yes* to the epistemic question, which is inconsistent with the content of the implicature. This number rises to 84.51% in the non-cooperative condition. Comparing the rate of *yes* answers to the epistemic question between the two conditions reveals that the *yes* responses were significantly more frequent in the non-cooperative condition than in the cooperative condition (X^2(1, N=425)=36.42, p<0.001). The same difference was found between the two versions of each story (at 0.01 Bonferroni-corrected significance level), except for the 'team retreat' and 'music performance' where the difference was only marginally significant (p=0.022 and p=0.021 respectively).

In the meaning question, a high percentage of participants in both the cooperative and non-cooperative conditions gave a *yes* response (75% and 71.83% respectively), indicating that they thought that the character intended to communicate the scalar implicature by uttering the target sentence. There was no significant difference in the rates of *yes* responses to the meaning question between the cooperative and non-cooperative condition (X^2(1, N=425)=0.55, p=0.46). There were also no significant differences in the frequency of *yes* responses between the two versions of each story (all ps>0.1).

In the deception question, 58.22% of participants in the non-cooperative condition gave a *yes* response, indicating that they thought the character was trying to mislead them. In contrast,

only 19.34% of participants in the cooperative condition gave a *yes* response to the deception question. Comparing the rates of *yes* responses to the deception question indicated that *yes* responses were significantly more frequent in the non-cooperative condition compared to the cooperative condition (X^2(1, N=425)=67.61, p<0.001). A significant difference in the rate of *yes* responses to the deception question was also found between the two versions of each story (at 0.01 Bonferroni-corrected significance level), except for the 'school editor' story which was only marginally significant (p=0.018) and 'music performance' which was not significant (p=0.063).

Table 5 *Frequencies of yes and no responses to the epistemic, meaning and deception question in each of the scenarios*

	Story	Epistemic Q.		Meaning Q.		Deception Q.	
		Do you think it's possible that 'all'?		Do you think the character meant 'not all'?		Do you think the character was trying to mislead you?	
		Yes	No	Yes	No	Yes	No
Cooperative	Company exam	64.29%	35.71%	71.43%	28.57%	9.52%	90.48%
	Team retreat	74.42%	25.58%	58.14%	41.86%	30.23%	69.77%
	Gameshow	55.81%	44.19%	88.37%	11.63%	11.63%	88.37%
	School editor	47.62%	52.38%	80.95%	19.05%	26.19%	73.81%
	Music	47.62%	52.38%	76.19%	23.81%	19.05%	80.95%
	Total	58.02%	41.98%	75.00%	25.00%	19.34%	80.66%
Competitive	Company exam	90.70%	9.30%	76.74%	23.26%	65.12%	34.88%
	Team retreat	92.86%	7.14%	52.38%	47.62%	69.05%	30.95%
	Gameshow	88.10%	11.90%	76.19%	23.81%	69.05%	30.95%
	School editor	79.07%	20.93%	74.42%	25.58%	51.16%	48.84%
	Music	72.09%	27.91%	79.07%	20.93%	37.21%	62.79%
	Total	84.51%	15.49%	71.83%	28.17%	58.22%	41.78%

5.3.1 Reaction time

The reaction times (the time spent on the page) for the target sentence and the epistemic question (see **Table 6**) were converted to their base ten logarithm to make their distribution more normal. Observations that were two standard deviations above or below the mean were removed.

94

Table 6 *Median raw reaction times (seconds) and means with standard deviations (base 10 logarithm) to the target sentence and epistemic question in each condition*

	Cooperative condition		Competitive condition	
	Median	Mean (SD)	Median	Mean (SD)
Target Sentence	3.80	0.56 (0.16)	3.53	0.54 (0.16)
Epistemic Q. – Yes response	6.82	0.85 (0.17)	6.02	0.80 (0.18)
Epistemic Q. – No response	7.99	0.89 (0.17)	6.68	0.83 (0.15)

The time participants spent reading and interpreting the target sentence was compared between the two conditions with a One Way ANOVA. Although participants in the cooperative condition were numerically slower than participants in the competitive condition they were not significantly slower ($F(1,411)=2.07$, $p=0.15$).

Reaction times to the *epistemic* question was analysed with a 2x2 factorial ANOVA with *answer type* (*yes/no*) and *condition* as factors. A Type III sum of squares was used in this analysis to compensate for the unbalanced group sizes. Neither the effect of answer type ($F(1,405)=0.78$, $p=0.377$), nor the effect of condition ($F(1,405)=2.87$, $p=0.091$), nor their interaction ($F(1,405)=0.04$, $p=0.846$) were significant.

Following Mazzarella and colleagues (2016), the reaction time to the target utterances of participants who answered *yes* to the epistemic question and that of participants who answered *no* only for the competitive condition was also compared. No significant difference was found between these two groups ($F(1, 203)=0.12$, $p=0.73$).

5.4   Discussion

This experiment addressed the question of whether hearers infer less implicatures from an uncooperative speaker, whose unwillingness to cooperate is not made common ground, than from a cooperative one. More participants in the non-cooperative condition than in the cooperative condition gave a *yes* response to the *epistemic* question, which is incompatible with the content of the implicature of the target sentence. Roughly three fourths of participants in both conditions answered *yes* to the *meaning* question, indicating that they thought the speaker intended to communicate a scalar implicature by their utterance. While 58% of participants in the non-cooperative condition answered *yes* to the *deception* question, indicating that they thought that the character was trying to mislead them, significantly less participants did so in the cooperative condition (only 19%).

The difference between the two conditions in the responses to the *epistemic* question indicates that the experimental manipulation was successful. However, this measure does not distinguish between the hypothesis that participants in the non-cooperative condition derived less implicatures than participants in the cooperative condition and the hypothesis that participants in non-cooperative condition infer implicatures to the same extent as participants in the cooperative condition but they are less likely to accept its content. If, consistently with the prediction derived from Grice (1989), the difference in responses to the *epistemic* question was due to participants in the non-cooperative condition inferring less implicatures than participants in the cooperative condition, participants in the non-cooperative condition should also be less likely than participants in the cooperative condition to indicate that the speaker intended to communicate the scalar implicature. Instead, the rate of *yes* answers to the *meaning* question was relatively high in both conditions and there was no significant difference between the two conditions in the responses to this question. Therefore, this pattern of results does not fit the prediction derived from Grice (1989) that hearers do not derive implicatures from an uncooperative speaker. In contrast, these findings fit the predictions of Relevance Theory and the Epistemic Vigilance account that hearers interpret the utterances of untrustworthy speakers as if they were trustworthy and then they may reject the content of what the speaker communicated. The responses to the *meaning* question indicate that the interpretation that participants in the two conditions assigned to the utterance did not differ, at least to the extent that it included a scalar implicature, whereas the responses to the *epistemic* question indicate that participants in the two conditions differed in the extent to which they trusted the content of the scalar implicature. The presence of a difference between the two conditions in the responses to the *epistemic* question but not in the *meaning* question is therefore in support of the predictions of the Epistemic Vigilance account and it highlights the dissociation between the processes of comprehension and acceptance that is advocated by this account (Sperber et al. 2010; Mazzarella, 2015a, 2015b).

The results to the *deception* question also fit the hypothesis that participants in the non-cooperative condition are inferring the scalar implicature and then rejecting its content. The majority of participants in the non-cooperative condition answered *yes* to the *deception* question, indicating that they thought that the character was trying to mislead them. In contrast, participants in the cooperative condition mostly did not think that the speaker was trying to mislead them. If participants in the non-cooperative condition thought that the speaker did not intend to communicate the scalar implicature, they should not consider the speaker to be misleading but simply under-informative. Instead, if participants in the non-cooperative condition are inferring and then rejecting the scalar implicature because of the speaker's

96

untrustworthiness, they should consider the speaker to be either misinformed or deceitful. Since the scenarios used in this experiment suggested that the speaker was competent with regard to the content of the implicature, participants should see the character as misleading rather than misinformed, which is exactly what the responses to the *deception* question suggest. This interpretation however does not fit all the participants who answered *yes* to the *deception* question. The interpretation of these responses hinges on what participants consider to be '*misleading',* which may vary across participants. For instance, 54 participants across both conditions indicated that the speaker was trying to mislead them despite having indicated that they did not think the speaker was trying to convey a scalar implicature. These participants may consider under-informative utterance to count as *misleading*.

One possible interpretation of the character's utterance was that the character was trying to communicate an ignorance inference instead of a scalar implicature (i.e. that they did not know whether *all* was the case). The predictions for this scenario are similar to the ones that apply to the scalar inference. The prediction derived from Grice that hearers should not infer implicatures from uncooperative speakers applies in this case as well. Relevance Theory and the Epistemic Vigilance account would predict that if the ignorance inference is part of the most relevant interpretation compatible with the speaker's *abilities and preferences* (e.g. the speaker is not competent enough to be communicating a scalar implicature) it should be derived and then possibly rejected depending on the outcome of the hearer's epistemic assessment. However, regardless of the theoretical predictions, if the ignorance inference were a popular interpretation among participants in this experiment, they would have indicated that they did not think that the speaker intended to communicate a scalar implicature. Instead most participants in both conditions answered yes to the *meaning* question, indicating that they inferred a scalar inference and not an ignorance inference from the character's utterance. This suggests that the stories were successful in conveying the assumption that the character was competent regarding the content of the scalar inference.

The analysis of the reaction time had the twofold aim of exploring how the non-cooperative context affected the speed of comprehension and epistemic assessment of the speaker's utterance and of providing a comparison to the findings of Bonnefon et al. (2009) and Mazzarella et al. (2016). Bonnefon and colleagues measured the total time participant took to comprehend and provide an epistemic evaluation of a face-threatening utterance. They found that participants who gave an epistemic assessment incompatible with the content of the implicature (a *yes* answer) took longer to read the utterance and answer the epistemic question. Using the same paradigm, Mazzarella and colleagues measures the time it took participants to read the utterance and to give an epistemic assessment of the implicature content separately. They

found that giving a *yes* answer to the epistemic question made the epistemic assessment slower but it did not affect the time participants took to read and comprehend the utterance. Following Mazzarella and colleagues, this experiment measured separate reaction times for utterance comprehension and epistemic assessment. As in the results of Mazzarella and colleagues, the answer to the epistemic question did not affect the time it took participants to read and comprehend the utterance. However, while they found that the answer to the epistemic question affected the reaction time to the epistemic question, the present experiment did not offer any evidence for this difference. The results of the present experiment also indicated that being in the cooperative or non-cooperative condition did not affect the time participants took to interpret the utterance or provide an epistemic judgement. Although I will not venture into an in depth discussion of the reaction time results, firstly because they are all null results and secondly because the main focus of this study was on the choice data, it can at least be pointed out that these results do not offer any counter evidence to the proposed interpretation of the choice data. For instance, if participants in the cooperative condition were drawing more implicatures than in the non-cooperative condition, a delay in the comprehension time for the target utterance in the cooperative condition due to the additional inference would have been expected (Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006), instead there was no evidence of this delay.

In conclusion, the findings of this study offer support to the predictions of Relevance Theory and the Epistemic Vigilance account that hearers of cooperative and non-cooperative follow the same interpretation procedure to an *optimally relevant* interpretation of the utterance (including any implicatures of the utterance) and then they may reject the information conveyed explicitly or implicitly by the speaker. Instead these findings are in contrast with the prediction derived from Grice's account that hearers of an uncooperative speaker should not infer implicatures. The pattern of results presented in this study is also a good illustration of the dissociation between utterance comprehension and epistemic assessment of the communicated content (Sperber et al. 2010; Mazzarella, 2015a). While the studies presented in this chapter and the previous one focused on how hearers comprehend the utterances produced by an uncooperative speaker, the next chapter will present a production study. In particular, it will address the question of how uncooperative speakers behave by asking participants to behave as uncooperative speakers.

6 Communicating strategically in a competitive game

In the previous two chapters I focused on how a non-cooperative context affects comprehension and in particular the inference of quantity implicature. In this chapter I will instead focus on production and present a study in which participants assume the role of speakers in a non-cooperative setting. As in the study presented in Chapter 4, the non-cooperative is operationalised as a competitive signalling game except that in the study presented here the participants are signallers instead of receivers. One aim of this study is to explore how participants communicate strategically in a non-cooperative situation. Furthermore, the materials of this study involve three different kind of quantity implicatures with the prospect that the similarities and differences in the way these three categories of implicature are used may reveal something about their nature and relationship.

6.1   Introduction

Grice (1989) presents conversation as a cooperative activity in which participants abide a *cooperative principle*, which binds them to make appropriate contributions to the conversation. From this principle follow more specific maxims such as the first maxim of quantity: "Make your contribution as informative as is required" (Grice, 1989, p. 45). Speakers can exploit the maxims in order to communicate implicit propositions (implicatures) of various types. For example, the speaker can violate the first maxim of quality to communicate a quantity implicature. If I say that *John used some of the shampoo* in a context where it would be relevant and more informative to know whether *John used all of the shampoo*, my hearer may infer that the reason why I am violating the first maxim of quantity is that the more informative statement is not true and therefore infer the implicature that *John did not use all of the shampoo*. The last two decades witnessed a wave of experimental investigation of how different types of quantity implicatures are processed and interpreted; and in harmony with Grice's account, these investigations have focused on situations where the cooperation and honesty of the speaker is taken for granted. However, conversation also takes place in non-cooperative or competitive situations, where the speaker may be deceptive or uninformative. Politicians are often good examples of unhelpful interlocutors. For instance, consider this evasive answer that Theresa May gave in 2016 when asked whether the UK should have access to the EU single market after Brexit: "What I want to see is the best possible deal for the United Kingdom in trade in goods and services" (Bull, 2016). The use and comprehension of implicatures in non-cooperative settings is a vastly understudied topic. To the best of my knowledge, the studies presented in this thesis and Pryslopska's (2013) study are the only existing comprehension studies on this topic and no published experiments investigate implicature production in uncooperative contexts. This study will address the question of how speakers use explicit and implicit communication strategically in a non-

cooperative scenario and how they use different types of implicature. Therefore, this study may offer a new perspective on the differences between well studied types of quantity implicatures.

A non-cooperative speaker may differ from a cooperative one in that they may be more likely to deceive or to be uninformative. Although Grice (1989) presents conversation as a cooperative effort, he contemplates both the possibility that speakers may be uninformative by *opting out* of the cooperative principle or of a maxim in an *overt* way, for example by saying "I can't tell you that", and the possibility that they may be deceitful by *covertly* violating a maxim. The paramount example of covert violations of maxims is lying, where the liar covertly violates the first maxim of quality (i.e. "Do not say what you believe to be false"; Grice, 1989, p. 46) and intends the audience to remain unaware of the violation. Besides lying, the realm of verbal deception includes falsely implicating. While to lie, at least according to traditional definitions (Isenberg, 1973; Primoratz, 1984), is to *say* something that the speaker believes to be false with the intention to deceive; to falsely implicate is to communicate something believed to be false by means of a *conversational implicature* (Meibauer, 2014). For example, if I said that *John used some of the shampoo* when in fact I believe that he used all of it, I could be falsely implicating that *John did not use all of the shampoo*. Although, there is an ongoing conceptual debate on whether false implicatures should be considered *lies* (Meibauer, 2005, 2014) or not (Dynel 2011, 2015) here I will treat them as separate for the purposes of experiment design and analysis. I will let the data speak about any potential difference between false implicatures and lies. Therefore the phenomena which may be expected in this study are uninformativity or *opting out*, lies and false implicatures.

It is reasonable to expect that explicit and implicit communication[4], which include lies and false implicatures respectively, are used differently in non-cooperative contexts. One reason for this expectation is that in *cooperative* conversation these two modes of communication are often not interchangeable and in a given context speakers usually have clear preferences as to whether a piece of information should be asserted or communicated implicitly. Consider the following examples adapted from Russel (2012):

1.   Careful! Some of the mushrooms are poisonous!
2.   #Careful! Not all of the mushrooms are poisonous!

The utterances in 1 and 2 carry the same content except that the explicit content of the utterance in 1 is implicit in the utterance in 2 and vice versa (see also van Tiel, 2014 for a

---

[4] I will assume that the explicit implicit distinction corresponds to the distinction between Grice's *what is said* and what is *implicated*. However, see Cartson (2002, 2009) and Recanati (2004) for different perspectives.

discussion of a similar example). Since the expression 'careful' suggests that both utterances are warnings, the key information that the speaker presumably wants to convey is that at least some of the mushrooms are poisonous, and not that some of them are harmless. Conveying the key information as implicit content rather than explicit content makes the utterance in 2 sound odd. Intuitively, this may be because in this context we would prefer to communicate key information through the relatively '*secure*' channel of explicit information as implicit communication is arguably more prone to misunderstanding (Reboul, 2017). Another reason why explicit and implicit communication may be used differently is that some features of implicit communication could be advantageous in non-cooperative contexts. For example, implicit communication offers the advantage of *plausible deniability* (Pinker, Nowak & Lee, 2010). Since implicatures are cancellable the communicator can deny having intended to communicate them. For example, after saying that *John used some of the shampoo,* I can claim that I did not mean to communicate that he did not finish it. I could not make the same claim if I explicitly said that he did not use all of it. This feature of implicit communication is useful in cases where the speaker wants to communicate something that may incur them some penalty, such as proposing a bribe or communicating false information. Reboul (2017) proposes that implicit communication may also offer another advantage in that it may be accepted more easily by the hearer than explicitly communicated content. Firstly, she claims that hearers are more vigilant towards content that the speaker is strongly committed to, and explicit content carries a higher degree of speaker commitment compared to implicit content (Morency, Oswald & de Saussure, 2008). Secondly, she claims that hearers are less vigilant towards content that is the fruit of their own inferences, which is the case for implicatures but not for asserted content.

The previous studies that are closest to the one presented here are recent studies looking at non-verbal deception in the context of signalling games where signallers have to give non-verbal hints (e.g., images, maps) to a receiver player who has to make choices based on the information provided in the hints. Crucially, in some cases the game is competitive and the signaller benefits from the receiver's wrong choices, which provides motivation to deceive. Signallers can give true hints, false hints, uninformative hints and misleading hints, which like false implicatures consist in conveying a true piece of information which leads the receiver to infer something false. Montague and colleagues (2011) found that their players preferred giving misleading hints rather than false hints. In their game the receiver did not know whether the signaller was cooperative or competitive and they could choose to check whether the hints were false and calibrate their trust accordingly, which was an incentive for the signaller not to be caught lying as it would have reputation consequences for the rest of the game. In a similar competitive game Ransom, Voorspoels, Perfors and Navarro (2017) gave their participants the options to give to

the receiver true, misleading or uninformative visual hints, but not false hints, and they manipulated the signaller's expectations regarding how suspicious or trustful the receiver would be. Because the receiver did not know whether their signaller was honest or deceitful, the signaller could pretend to be helping the receiver while in fact feeding them misleading information. They found that when signallers expected a trustful receiver they were more likely to mislead, whereas when they expected a suspicious receiver they were more likely to be uninformative.

Similarly to these studies, the study presented here employs a competitive signalling game. Furthermore, the types of deception these studies investigated follow the same fundamental mechanisms of the kinds of verbal deception which may be used in this study, which are to cause someone to have a false belief (Mahon, 2007) either by communicating something false (i.e. false hints, lies) or by communicating something true (i.e. misleading hints, false implicatures).

While the studies on non-verbal deception mentioned above drew a clear distinction between false and misleading hints; drawing a distinction between lies and false implicature is not straightforward. In two studies (Coleman & Kay, 1981; Hardin, 2010) where participants were asked to rate a false implicature on a scale that ranged from an utterance being a lie to an utterance not being a lie the average rating was near the middle of the scale. In parallel to these results, studies on the explicit-implicit distinction in comprehension found that lay people are likely to consider implicatures part of *what is said* under some circumstances (Nicolle & Clark, 1999; Doran, Baker, McNabb, Larson & Ward, 2009; Doran, Ward, Larson, McNabb & Baker, 2012). Doran and colleagues (2012) asked participants to judge whether sentences that could give rise to an implicature were true or false in the light of a fact that contradicted the implicature (e.g. judging whether the sentence *John used some of the shampoo* is true given that he used all of it). They found that participants incorporated scalar implicatures arising from quantifiers such as *some* and *most* into the truth conditional meaning of the sentence 32% of the time and implicature arising from cardinal numbers (e.g. *I have three cats* implicating that I don't have four) 53% of the time. Because different types of implicatures may differ in whether they are considered part of *what is said*, and therefore in whether they would be considered to be lies if used deceptively, the study presented here aims to gain a more comprehensive perspective by using three different types of implicatures.

The materials used in this study use three types of quantity implicatures: implicatures arising from numerals, the scalar implicature arising from the quantifier *most* and *Ad hoc* or *particularised* quantity implicatures. The quantity implicatures or upper-bound interpretations associated with scalar terms and numerals are drawn by negating an alternative utterance

where a stronger term on the same lexical scale as the scalar/numeral term is used. For example, the implicature of 'John used *some* of the shampoo' arises from negating the alternative that 'John used *all* of the shampoo'. And the implicature or upper-bound interpretation of 'John has *three* cats' arises from negating the alternative that 'John has *four* cats'. This contrasts with *Ad hoc* quantity implicatures where the stronger alternative can only arise from the context and not from the lexicon. For example, in the context where Ann and Rose have their birthdays together someone might say 'I bought a present for Ann' and implicate that they did not buy a present for Rose also. The stronger alternative negated is that they *bought a present for Ann and Rose* but it can only arise in this particular context. Both scalar terms and numerals have been at the centre of theoretical controversies concerning whether their upper-bound interpretation is an actual implicature or whether it is part of their semantic or *default* meaning (Levinson, 2000; Geurts, 2010).

Although some theorists have proposed that the upper-bound interpretation of scalar terms (e.g. *some and not all*) is their default meaning (e.g. Levinson, 2000; Chierchia, 2004), recent experimental evidence suggests that the scalar implicatures of quantifiers (i.e. some *and not all*) are derived in the same way as particularised quantity implicatures (see Katsos & Cummins, 2010 for a review). The distinction between scalar and Ad hoc implicatures has received particular attention in the acquisition literature, where experimental studies offer mixed results: some studies suggest that pre-school children have more trouble calculating scalar implicatures compared to *Ad hoc* implicatures (Bale, Brooks & Barner, 2010; Stiller, Goodman & Frank, 2011) and others suggesting that they calculate these two types of implicatures to the same extent (Katsos, 2009). Katsos (2009) asked participants to evaluate utterances that could give rise to either scalar or particularised implicatures in contexts where the content of the implicature is false. He found that both adults and children reject scalar implicature utterances and *Ad hoc* implicature utterances to the same extent, but adults consider an under-informative scalar-implicature utterance to be a more serious violation of informativity than an under-informative particularised implicature utterance.

With regards to numerals there are again two camps, with some theorists claiming that they have a lower-bound or *at least* meaning while the *exact* interpretation is supplied in context via implicature (e.g., Horn, 1972; Gazdar 1979; Levinson, 2000) and others claiming that the *exact* interpretation of numerals is not an implicature but part of their truth conditional meaning (Carston, 1998[5]; Breheny 2008; Kennedy 2015). Both the study of Papafragou and Musolino

---

[5] Carston (1998) actually argues that cardinals have an underspecified meaning and that whichever sense they assume in context (i.e. at least, at most or exactly) contributes to the truth conditional meaning of the utterance.

(2003) and of Huang, Spelke and Snedeker (2013) provide convincing evidence that numerals have an exact truth-conditional interpretation (i.e. three means 'exactly three') by showing that pre-school children, who are notoriously 'bad' at calculating scalar implicatures (Noveck, 2001; Chierchia et al, 2001; Gualmini, Crain, Meroni, Chierchia, & Guasti, 2001; Hurewitz, Papafragou, Gleitman & Gelman, 2006; Pouscoulous, Noveck, Politzer, Bastide, 2007) tend to give exact interpretations of numerals. Furthermore, Huang and Snedeker (2009) found that while processing the upper bound meaning scalar terms like *some* is slower for adults than processing the literal meaning of the quantifier *all*, processing the *exact* meaning of numerals is just as fast; suggesting that the former involves drawing a pragmatic inference and the second does not.

This study aims to explore how speakers use explicit and implicit communication in a non-cooperative context. Participants played a signalling game similar to game employed in the study presented in Chapter 3, except that participants played in the role of signallers. As in the study of Ransom and colleagues (2017), participants had to produce hints for a receiver either in a cooperative or in a competitive scenario. There are some fundamental differences between the game employed in the study and the one used by Ransom and colleagues. Firstly, in the present study participants are told that the receiver knows whether the game is cooperative or competitive. This removes the possibility for signallers to pretend that they are cooperative when their actual goal was to misinform the receiver. This feature of the competitive scenario eliminates the possibility of cooperation and, from a Gricean perspective, should push signallers towards the strategy of *opting out* and being uninformative. Secondly, participants in the present study gave linguistic hints by completing short descriptions. Since these hints are assertions, the false hints used in the games are lies, in the traditional sense, and the misleading hints give rise to false implicatures. Thirdly, the description templates that signallers completed constrained their hints into pre-determined categories. Half of the description templates pushed participants to convey the hint explicitly and half of them through an implicature – belonging to one of three types: Ad hoc quantity implicatures, linked to the scalar quantifier *most* or to the use of cardinals.

6.2 Methods

6.2.1 Materials and design

The game used in this paradigm is a signalling game with two players. Each round of the game has two cards such as the cards in *Figure 7*: a 'winning' card and a 'losing' card. The signaller knows which one is the winning card and they have to describe it. The receiver sees the same two cards but they don't know which one is the winning card. The receiver has to decide which one is the winning card with the help of the description made by the sender.
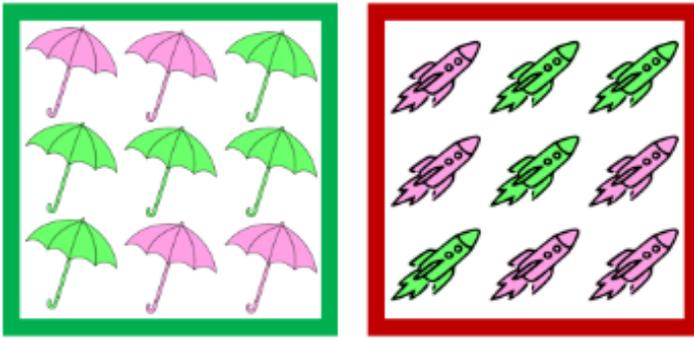
Figure 7 *winning card (green) and losing card (red) associated to the template description "On the winning card all of the objects are ___ "*

Participants only played the role of the describer, while the receiver player was a virtual player. Participants were assigned to one of two conditions: a cooperative condition or a competitive condition. In the cooperative condition participants were asked to help the receiver find as many winning cards as possible (a game of pure cooperation, in game theoretic terms), while in the competitive condition their goal was to make the receiver click on as many losing cards as possible (a so-called zero sum game).

Materials included in total 36 items (**Appendix C1**), which corresponded to 36 rounds of the game: 18 experimental items and 18 control items. Each item consisted of a template description and the two cards: the winning card, marked by a green outline, and the losing card, marked by a red outline (see *Figure 7*). Rather than write the whole description of the winning card, participants were asked to complete a pre-made description with only one word (e.g. see template description for *Figure 7*). All items were constructed in such a way that they had two obvious completions, referring either to the shape or the colour of the objects displayed in the cards (green vs. pink and rockets vs. umbrellas in the example in **Figure 1**). Control items used description templates containing either the quantifiers *all* or *none* and they had two obvious completions: a true assertion or a false assertion about the winning card. Experimental items used descriptions that could give rise to three types of quantity implicatures: exact interpretation of numerals, scalar implicatures associated with the quantifier *most* or particularised *Ad hoc* quantity implicatures. Experimental items were constructed in such a way that one of the two most accessible completions resulted in a true assertion giving rise to a true implicature while the other most obvious completion produced a true description giving rise to a false implicature. Each category of items, control and experimental, was counterbalanced for whether it was mentioning the colour or the shape of the object that gave rise to the false hint (false assertion for control items or to the false implicature for experimental items). The most accessible false assertions and false implicatures in each item were false of the winning card but

true of the losing card, so that they could be used to deceive the guesser into thinking that the losing card was actually the winning card. **Table 7** provides examples for each category of items.

Table 7 *Examples of each category of items*

Item Type	Description	True completion	False completion	Cards
Control (All)	On the winning card all of the objects are ___	umbrellas	Rockets	
Control (None)	On the winning card none of the objects are ___	blue	Green	
Experimental (Numeral)	On the winning card two of the objects are ___	blue	Mugs	
Experimental (Most)	On the winning card most of the objects are ___	lamps	Yellow	
Experimental (Ad hoc)	On the winning card the objects in the middle row are ___	green	Apples	
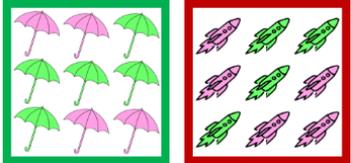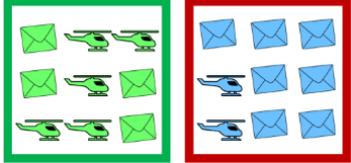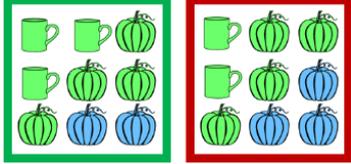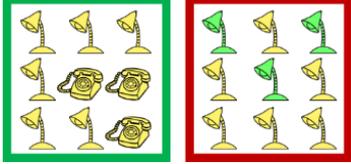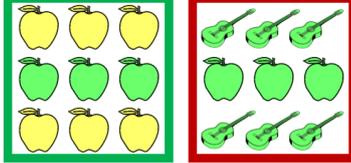
6.2.2 Participants and Procedure

103 native English speakers (66 females, Mean Age = 28.73) were recruited from the online crowdsourcing website prolific.co.uk and directed to the Qualtrics website where the experiment was hosted. Participants were told that they would play a game where they would have to complete descriptions with one word. Since the task required normal colour vision participants were screened for colour-blindness using two plates from the Ishihara colour-blindness test (Ishihara, 1917). Participants who failed the screening test were prevented from continuing the experiment. Each participant was randomly assigned to either a cooperative or a competitive condition in a between-subjects design. Participants in both conditions were told that the other player scored points by clicking on winning cards and that in each round he or she would read their description and use it to decide which card to click on. Participants in the cooperative condition were told that they themselves would score points when the guesser clicked on a winning card. Consequently, their goal was to help the guesser. Instead, participants

in the competitive condition were told that they would score points when the guesser clicked on the losing cards. Consequently, their goal was to make the guesser lose. Participants in the competitive condition were explicitly told that the other player knew that the person writing the descriptions was playing against them and vice versa for participants in the cooperative condition. This was to ensure that participants in the competitive condition knew that they could not pretend to be cooperative as the other player would expect them to be uncooperative.

Participants in both conditions were told that the other player did not know that they were completing the descriptions instead of writing them freely. This was to prevent participants from anticipating that the other player would think that the reason why they had not uttered a more informative description in experimental items was because the game prevented them, which would effectively block the derivation of quantity implicatures.

Participants in both conditions were told that they must complete the descriptions with only one word and they were explicitly told that they were allowed to lie. Before allowing participants to perform the actual task of the experiment they were asked four multiple choice questions to check their understanding of the game and participants who answered incorrectly to any of the four questions were prevented from continuing the study. Instructions for both conditions and comprehension questions are reported in **Appendix C2**.

Each participant saw all of the 36 items divided into two randomized blocks. Participants were not given feedback on the choices of the receiver as they believed that the receiver would play the game in a second phase. After the last item participants were asked to predict their performance by estimating on an 11-point scale ranging from 0% to 100% on how many rounds the guesser reading their descriptions would click on the winning card.

6.3   Results

Descriptions were automatically coded using an R script which classified each entry according to a predefined list of response types for each item. The list was constructed *a priori* and then adjusted after inspection of the data to accommodate spelling variants and unforeseen strategies. Two-word entries, which were explicitly forbidden in the instructions, were excluded from the analysis. Because of a counterbalancing error two out of the six *Ad Hoc* items had to be excluded from the analysis.

Each description was categorised in one of four categories: *true* and *false*, which corresponded to the two most obvious completions that were anticipated (see **Table 7**), as well as *uninformative* and *other*. Descriptions which could either apply to both the winning card and the losing card or to neither were classified as *uninformative*. For example, an uninformative

response for the *all* control item in **Table 7** was "On the winning card all of the objects are *quaint*", and for the *none* control item in ***Table 7***, in which both cards contain helicopters, an uninformative response was "On the winning card none of the objects are *helicopters*". Descriptions classified as *other* were descriptions that attempted to describe only one card but not through the two obvious completions expected (i.e. *true* and *false*). These description mainly fell in one of two categories of alternative strategies. One strategy consisted in mentioning an object associated only with one card even though this resulted in a statement that was false of both cards. An example of this strategy for the *most* item in **Table 7** was "On the winning card most of the objects are *telephones*", which refers to the winning card where telephones are present even though most of the objects are lamps. The other alternative strategy consisted in making reference to whether the objects singled out by the description were the *same* or *different* from the other objects in the card and relying on the preferred reading of the statement as referring to the shape of the objects rather than the colour. An example of this strategy for the *Ad hoc* item in **Table 7** was "On the winning card the objects in the middle row are *identical*". Because the amount of data in these two categories was relatively small they were grouped together under the category of *other strategies* for the purpose of this analysis.

In the cooperative condition (**Figure 8**) participants had an overwhelming preference for true descriptions in both types of control items (*all* and *none* items) and in items containing numerals. In the *most* and *Ad hoc* items participants had a preference for true descriptions but they also gave a considerable number of false descriptions. *Most* and particularly *Ad hoc* items also differed from the other item categories because of the high rate of *other* responses given by participants: 12% of the responses for *most* items and 30% for *Ad hoc* items. In the *predicted performance* question participants in the cooperative condition estimated that the receiver would click on the winning card 78.15% of the time.
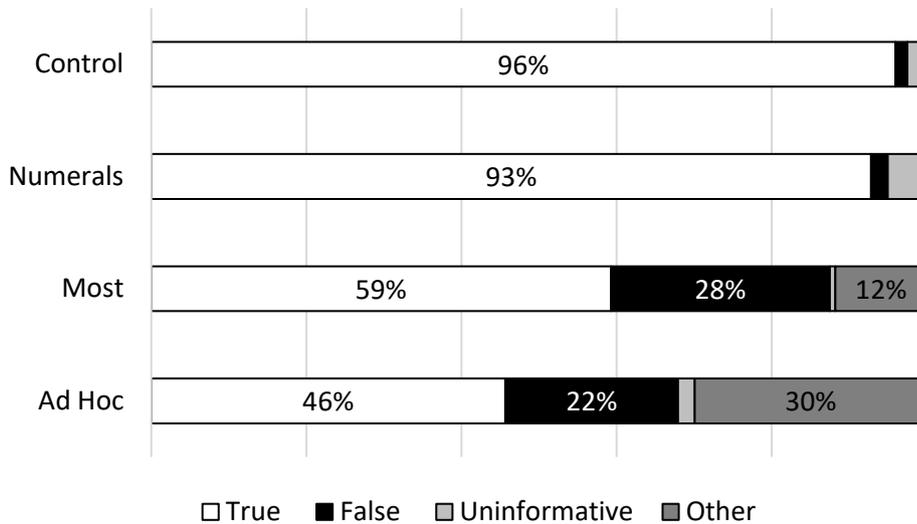
Figure 8 *Proportion of response types in the cooperative condition*

In the competitive condition (**Figure 9**) participants gave mostly equal numbers of true and false descriptions for all types of items. Participants also gave a considerable amount of uninformative descriptions in control items and items with numerals. The rate of uninformative descriptions was lower for *most* and *Ad hoc* Items. *Most* and *Ad hoc* items also exhibited a higher rate of *other* responses than other item categories but not as large as in the cooperative condition. In the predicted performance question participants in the competitive condition estimated that the receiver would click on the winning card 49.82% of the time, which was a significantly lower estimate than the one given by participants in the cooperative condition (t(98.81)= -8.351, p<0.001).
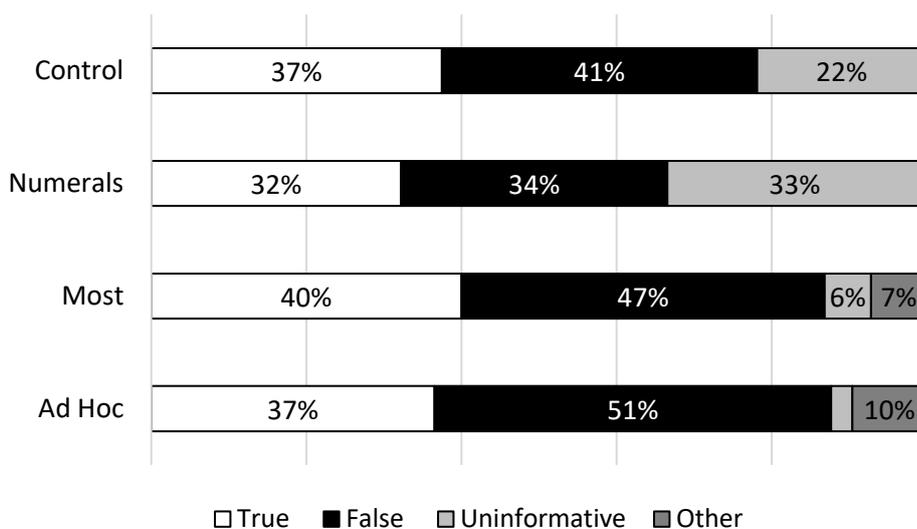
Figure 9 *Proportions of response types in the competitive condition*

The data was analysed with three binomial regression models. Each model regresses a different outcome variable on the same predictors: condition, item type and their interaction. Item type is a four-level dummy coded categorical predictor with control items as the reference level. Condition is a two-level dummy coded categorical predictor with the competitive condition as the reference level. Therefore, the simple effect of the cooperative condition represents the difference between conditions for control items; the simple effects of Numerals, Most and *Ad hoc* represent their difference from control items in the competitive condition, and their interactions with condition factor express how their difference from the control items changes in the cooperative condition.

Model 1 addresses the question of what factors affect the preference for a false description over a true description and the outcome variable was a binary variable where *true* descriptions were coded as 0 and *false* descriptions were coded as 1. The details of the analysis are summarised in **Table 8**. The three types of experimental items are not significantly different from control items in the competitive condition. The negative effect of the cooperative condition indicates that participants were less likely to give false descriptions for control items in the cooperative condition compared to the competitive condition. The significant interactions of *most* and *Ad hoc* indicate that the difference between these items from the control items in the cooperative condition is different from their difference in the competitive condition. There is no evidence that this was the case for numerals.

***Table 8**. Model 1*

	B	S.E.	Z	p-value
(intercept)	0.08	0.07	1.14	.254
Numerals	-0.03	0.15	-0.18	.854
Most	0.07	0.14	0.50	.613
Ad Hoc	0.24	0.16	1.48	.139
Cooperative	-4.22	0.29	-14.62	< .001
Numerals*Cooperative	0.39	0.52	0.74	.459
Most*Cooperative	3.32	0.34	9.77	< .001
Ad Hoc*Cooperative	3.19	0.37	8.53	< .001

Model 2 addresses the question of what factors affect the preference for an uninformative description over all other types of descriptions (true, false and other) and the outcome variable is a binary variable where uninformative descriptions were coded as 1 and all other responses were coded as 0. The details of the analysis are summarised in ***Table 9***. The negative simple effect of the cooperative condition indicates that participants were less likely to give uninformative descriptions for control items in the cooperative condition compared to the competitive condition. The negative simple effects of *most* and *Ad hoc* indicate that participants were less likely to give uninformative descriptions for these items compared to control items in

the competitive condition. On the contrary, the positive simple effect of numerals indicates that participants were more likely to give uninformative descriptions for these items compared to controls in the competitive condition. The significant interaction indicates that the difference between *Ad hoc* items and control items is different in the cooperative condition compared to the competitive condition. In fact, while the rates of uninformative descriptions for control items and *Ad hoc* items is roughly the same in the cooperative condition (i.e. 2%), they are considerably different in the competitive condition.

Table 9. *Model 2*

	B	S.E.	Z	p-value
(intercept)	-1.27	0.08	-16.68	< .001
Numerals	0.55	0.14	3.96	< .001
Most	-1.49	0.24	-6.14	< .001
Ad Hoc	-2.32	0.42	-5.52	< .001
Cooperative	-2.40	0.23	-10.27	< .001
Numerals*Cooperative	0.17	0.38	0.45	0.655434
Most*Cooperative	0.22	0.78	0.28	0.779232
Ad Hoc*Cooperative	2.39	0.66	3.64	< .001

Model 3 addresses the question of what factors pushed participants to resort to *other* descriptions instead of giving true, false or uninformative descriptions. The outcome variable for this model was a binary variable where *other* descriptions were coded as 1 and all other responses were coded as 0. The details of the analysis are summarised in **Table 10**. The only significant effects are the simple positive effects of *most* and *Ad hoc*, which indicate that participants were more likely to give *other* descriptions to these items than to control items in the competitive condition. The fact that their interactions are not significant means that there is no evidence that this trend was any different in the cooperative condition.

Table 10. *Model 3*

	β	S.E.	Z	p-value
(intercept)	-6.915	1	-6.911	< .001
Numerals	1.797	1.226	1.465	0.143
Most	4.35	1.023	4.253	< .001
Ad Hoc	4.747	1.024	4.634	< .001
Cooperative	-13.651	609.583	-0.022	0.982
Numerals*Cooperative	-1.797	1219.166	-0.001	0.999
Most*Cooperative	14.262	609.58	0.02	0.981
Ad Hoc*Cooperative	14.962	609.58	0.02	0.98

Although the overall proportions of true and false descriptions in the competitive condition seem roughly equal for each category of items (see **Figure 9**), these overall proportions were the result of different, sometimes opposed, individual strategies. For example, some

participants consistently gave false descriptions while others consistently gave true descriptions. These strategies were reflected in the performance predictions that participants gave about the receiver who would read their descriptions. In order to find out whether these individual strategies could be classified under a number of meaningful strategy profiles, the responses of participants in the competitive condition were analysed with a clustering analysis. A k-means clustering analysis was performed in R using four variables for each participant: overall proportion of true descriptions, overall proportion of false descriptions, overall proportion of uninformative descriptions and expected performance of the receiver. Using the average silhouette methodology (Rousseeuw, 1987), which allows to visually compare the quality of different clustering solutions in terms of tightness and separation of the clusters in each solution, a three-cluster solution was determined to be the clustering solution that best summarised the data. The three clusters are summarised in **Table 11** which reports each cluster's average values of the four variables used in the analysis (i.e. expected performance, overall proportions of true, false and uninformative descriptions; in the table as *clustering variables*) together with each cluster's average proportions of description types for each category of items.

Table 11. *Clustering variables and proportions of response types for each item category by clustering groups*

Cluster and size	Item type	Response type				Performance
		True	*False*	*Uninformative*	*Other strategies*	
Cluster 1	Clustering var.	59%	17%	16%		69%
N 18	Control	74%	07%	19%	01%	
	Numeral	55%	03%	37%	06%	
	Most	48%	49%	02%	01%	
	Ad Hoc	49%	42%	1%	8%	
Cluster 2	Clustering var.	24%	57%	11%		38%
N 32	Control	21%	66%	13%	0%	
	Numeral	24%	54%	22%	0%	
	Most	37%	51%	3%	9%	
	Ad Hoc	31%	55%	2%	12%	
Cluster 3	Clustering var.	16%	15%	59%		55%
N 6	Control	16%	6%	79%	0%	
	Numeral	6%	19%	75%	0%	
	Most	31%	19%	33%	17%	
	Ad Hoc	21%	50%	12%	17%	

Cluster 1 is characterised by a high rate of true descriptions and a high expected success rate (performance) of the receiver. Participants in this cluster were therefore mostly playing the

game as if their goal was to help the receiver as a high success rate of the receiver in the competitive condition corresponds to a low performance of the signaller, who caused the signaller to make only a few mistakes. Cluster 2, the most numerous, is characterised by a high rate of false descriptions and by the lowest expected performance of the three clusters. Participants in this cluster were mostly lying or falsely implicating and they expected their strategy to cause the receiver to perform worse than chance. In other words, those who believed to perform well as deceptive senders (and better than chance) are exactly those who used misleading implicatures. Cluster 3 is characterised by the highest rate of uninformative descriptions and an expected performance near chance. Although the rate of uninformative descriptions that these participants gave for *most* and *Ad hoc* items is still relatively high compared to the other clusters it is lower than for control items and numerals as participants seem to rely more on other strategies and on true and false responses.

6.4 Discussion

Participants in the present experiment played a signalling game in which they were either helping or competing against a receiver. Their task was to complete descriptions that could help the receiver choose the wining card out of each pair of cards. Some items pushed signallers to convey the hint via assertion and others via implicature (numerals, *most,* and *Ad hoc*). The hints used by participants were categorised into four types: true hints, which could be either true assertions or true implicatures, false hints, uninformative hints and *other*, where participants used ways of referring to one of the cards that was not expected. The items were constructed in such a way that the expected descriptions were the obvious completions for the description templates, therefore it is interesting that participants resorted to other strategies for completing the descriptions.

In the cooperative condition participants overwhelmingly chose true completions for the control items and the items containing numerals, with very few uninformative or *other* descriptions. This was expected given that their aim was to help the receiver find the winning card. Participants gave a considerable number of false descriptions for *most* and *Ad hoc* items in the cooperative condition. This is in contrast to the goal of helping the receiver and the most likely explanation for the high rate of false descriptions is that in some cases the potential implicatures of the descriptions were not available to participants and they randomly chose between the two most obvious completions. Furthermore, *most* and especially *Ad hoc* items elicited a considerable rate of *other* descriptions. This is interesting as these alternative strategies were not obviously available. One possible explanation is that participants anticipated the potential implicatures in these items and preferred to choose other strategies for communicating the key information in the description rather than trust the relatively unreliable channel of implicit

communication (Reboul, 2017). An alternative explanation, which is also consistent with the high rate of false descriptions, is that this behaviour is also caused by participants not seeing the potential implicatures of the two obvious descriptions, which without the implicatures are simply uninformative for the receiver. And in order to avoid giving an uninformative hint participants may have preferred resorting to other strategies. Previous production studies have investigated situations where speakers needed to communicate information through an inference rather than by asserting it. They found that speakers often, but not always, express themselves in a way that allow the hearer to draw an informative inference. In a study by Davies and Katsos (2010) participants needed to refer to objects in situations where using a bare noun would be under-informative (e.g. 'pass me the apple' in a situation where there are two apples) and their adult participants used expressions that allowed the hearer to draw a contrastive inference (e.g. pass me the *red* apple) almost 80% of the time. In a study by Degen, Franke and Jäger (2013), participants played a signalling game where they could only send messages that did not convey the key information unambiguously. Among the four messages they could choose from, only one conveyed the key information through an inference while the others were ambiguous or incorrect. They found that participants sent the target message on roughly 80% of the trials if the inference was simple and on 50% of the trials if the inference was complex. In the present study, participants in the cooperative condition are therefore towards the low end of the spectrum as they expressed themselves in a way that would allow a receiver to infer a true inference on 59% for trials for *most* items and 46% of trials for *Ad hoc* items. These rates were probably affected by the characteristics of the items and the complexity of the task.

The competitive condition differed from the cooperative condition mainly in the rates of false and uninformative descriptions. Participants in this condition were more likely to give false hints (i.e. to lie) in control items: they gave false and true descriptions at roughly the same rate. There was no evidence that the ratio of false to true descriptions was different for any of the other item types in the competitive condition. Items containing numerals, like control items, showed a large increase in the rate of false description in the competitive condition compared to the cooperative condition. The relative increase in the rate of false descriptions was significantly smaller for *most* and *Ad hoc* items as these items elicited a considerable amount of false descriptions in the cooperative condition as well. Participants were also more likely to produce uninformative descriptions for control items compared to the cooperative condition. Items containing numerals also elicited more uninformative descriptions in the competitive condition, in fact they elicited even more than control items. For *most* and *Ad hoc* items the uninformative descriptions were very few and significantly less than for control items. The fact that control items elicited a higher rate of false and uninformative descriptions in the competitive condition

suggests that the experimental manipulation had an effect as participants were aiming to cause the receiver to make mistakes either by lying or by being uninformative. The fact that participants relied either on uninformative hints and on equal ratios of true and false hints suggests that they did not expect to be able to cause the receiver to do worse than chance. This is also consistent with participants in the competitive condition indicating that they expected the receiver to click on the winning card roughly 50% of the time. In a similar non-verbal signalling game, Ransom, Voorspoels, Perfors and Navarro (2017) found that when the signaller expected a distrustful receiver, their participants gave uninformative hints roughly 75% of the time and only a few misleading or helpful hints. Although the participants in the present study were also probably expecting a distrustful receiver, the rate of uninformative hints they gave was much lower than what Ransom and colleagues found. This is most likely due to the fact that while participants in their study could not give a false hint, participants in the present study were allowed to lie and therefore they could take advantage of the fact that receivers would not know if informative hints were true or false. The competitive condition of the study presented in Chapter 4 employed a competitive signalling game very similar to the one used in the present experiment where participants played role of receiver and knew that the signaller was allowed to lie. The results of that condition indicated that both assertions and implicatures communicated by the signaller were interpreted as false half of the time and true half of the time. Therefore, the expectations of participants in that condition seem to match the behaviour of the signallers in the present study, who used true and false hints in roughly equal measure both in assertions and implicatures.

One interesting aspect of the way the three types of experimental items were used by participants is the fact that items containing numerals patterned with control items rather than with the other two categories of implicature items: *Ad hoc* and *most*. There was no evidence that numerals were used any differently than control items in terms of preference for false descriptions over true descriptions or in terms of the rate of *other* descriptions in either of the two conditions. In contrast to *most* and *Ad hoc* items, items containing numerals did not elicit more false hints than control items in the cooperative condition; which suggests that if the *exact* interpretation of numerals is an inference in this study it was as available as the semantic meaning of the quantifiers *all* and *none*. Another difference from *most* and *Ad hoc* items is that items with numerals did not elicit more *other* responses than control items in the competitive condition; which may be due either to the availability of *other* strategies for numeral items or due to the motivation to seek alternative strategies for these items. Items with numerals did differ from control items in eliciting more false descriptions in the competitive condition. However, this difference was in the opposite direction as *most* and *Ad hoc* items, which elicited

less uninformative descriptions compared to control items in the competitive condition. Overall, I interpret this pattern of results to indicate that participants used numeral items in a way that was closer to the control items than to the implicature items. This can be further taken to suggest that the *exact* interpretation of numerals is part of their truth-conditional meaning (Carston, 1998; Breheny 2008) and not an implicature (Horn, 1972; Gazdar, 1979; Levinson, 2000).

As mentioned, *most* and *Ad hoc* items were used differently than control and numeral items. They elicited a higher number of false descriptions than controls in the cooperative condition, which may be due to a lower availability of their upper bound interpretation. These items were also less likely to elicit uninformative descriptions compared to control items in the competitive condition. This may also be attributed to a lower availability of the upper bound interpretation, as the lower bound interpretation of *most* and *Ad hoc* description resulted in a description that was as unhelpful and indeed equivalent to an uninformative. These differences can be attributed to the fact that the key information was conveyed through assertion in control and numeral items, and through an implicature in *most* and *Ad hoc* items. The similarity in the way participants used *most* and *Ad hoc* items instead suggests that there was no difference in the way participants perceived these two types of implicatures. These results however cannot give an indication of how these implicatures are processed and therefore it cannot lend support either to the view that scalar implicatures and particularised implicatures are computed in the same way (Sperber & Wilson, 1986/1995; Carston, 2002; Breheny, Katsos & Williams, 2006; Geurts, 2010) or to the view that implicatures arising from lexicalised scales are default meanings computed differently from implicatures arising from *Ad hoc* scales (e.g. Levinson, 2000; Chierchia, 2004).

Although in the competitive condition the rates of true and false responses are overall equal, the cluster analysis suggests that this is actually the result of different opposing strategies that participants tended towards. One tendency was for participants to give more true hints, at least in control and numeral items, and expect the receiver to have a better performance as a result. The simplest explanation for why some participants chose this strategy, which is in contrast with their goal in the competitive condition, is that they were not following the instructions in this respect. In contrast, the largest group of participants tended to give more false descriptions and they expected the receiver to perform worse than chance as a result. These participants gave a higher rate of false descriptions for control items as well as for experimental items, suggesting that they were expecting the receiver to infer implicatures from their utterances and to trust their content. Although this is only a numerical observation as there were not enough data to perform meaningful statistical tests on this subgroup of participants, it would be interesting to

investigate this preference further. From a Gricean point of view, if participants did expect their receiver to draw false implicatures they must have expected the receiver to see them not as opting out but as cooperative enough to be communicating implicatures. Instead from the point of view of Relevance Theory and the account of Epistemic Vigilance (Sperber & Wilson, 1986/1995; Sperber et al., 2010), all participants may have expected the receiver's interpretation of their utterance to include the implicatures, but only participants who relied on false implicatures were expecting to be seen as trustworthy enough for the content of their implicatures to be trusted. A third smaller group of participants tended to give more uninformative descriptions. An interesting feature of this strategy is that by giving an uninformative hint participants made the unhelpfulness of their descriptions manifest to the other player. In Gricean terms, while giving false responses might be a case of a *covert* violation of the maxim of quality, giving uninformative responses signals that the speaker is *opting out* of the cooperative principle. On one hand, this strategy might be a calculated way of forcing the receiver to choose at random. On the other hand, some participants might prefer to be seen as opting out because they have an aversion to lying. In fact, multiple studies have found that people have an aversion to lying even in economic games where they would benefit from deceiving their interlocutor (Lundquist, Ellingsen, Gribbe & Johannesson, 2009; Gneezy, Rockenbach & Serra-Garcia, 2013). In other words, participants may have given uninformative hints in order to be honest about the fact that they were being unhelpful.

In conclusion, the results of this study indicate that uncooperative speakers tend to be more uninformative and to lie more than if their goal was shared with their interlocutor, at least in the kind of competitive scenario used in this study. Knowing that their interlocutor could be completely distrustful seems to push speakers towards the strategy of telling as many truths as lies, which seems to match the expectations that hearers had in the same situation in the competitive condition of the study presented in Chapter 4. Similarly to the competitive receivers in the study from Chapter 4, participants in this study were not uniform in their strategy and a large group of participants used a higher rate of lies and false implicatures, which suggests that they expected their interlocutors to infer and accept the content of their implicatures. The results of this study also suggest that the *exact* interpretation of numeral expressions is part of their truth conditional meaning and that particularised implicatures and the implicatures of scalar expressions are used in a similar way. This study explored how speakers communicate strategically in a non-cooperative situation where they expect the hearer to be distrustful. The way participants used utterances that could give rise scalar implicatures and *Ad hoc* implicatures in this setting was also compared to previous research on these types of implicatures and it was used to argue for a distinction between the *exact* reading of numeral expression on one hand

and the scalar implicature of *most* and *Ad hoc* implicatures on the other hand. In the next chapter I will investigate a different type of pragmatic inference altogether, namely the inference of the source of relevance for an utterance, which is also commonly sometimes formalised as *Question Under Discussion*.

7   Inferring the source of relevance

The previous three chapters, and in particular Chapter 4 and Chapter 5, presented studies that focussed on quantity implicatures and how this type of pragmatic inferences are affected by a non-cooperative versus a cooperative context. This chapter instead will present a study investigating a different type of pragmatic inference. Specifically, this chapter will present two experiments investigating how hearers infer the source of relevance for an utterance, which is often formalised as Question Under Discussion. The topic of Question Under Discussion ties in with the general theme of the thesis as this notion can be regarded as a formalization of Grice's purpose of conversation (see section 3.1.4). From the perspective of the parallel between cooperation in conversation and joint action it is therefore relevant to investigate how interlocutors individuate the shared goal or purpose of their joint enterprise which plays a fundamental role in conversation (see section 3.1.1.5). Experiment 1 will investigate whether the exhaustivity of the utterance with respect to different possible sources of relevance affects the choice of which source of relevance is used to interpret the utterance. Experiment 2 will explore how the effect of exhaustivity on the choice of a source of relevance interacts with the speaker's intention to deceive.

7.1   Introduction

On rare occasions, it happens that someone in a conversation says something that, despite being perfectly clear in its linguistic content, leaves the hearer confused and wondering '*what are they talking about?*' or '*why did they say that?*' or '*What is their point?*'. What is important however is that this is indeed a *rare occasion*. Most of the time hearers know why interlocutors say what they are saying and what point or topic they are addressing. This means that for every utterance hearers interpret, they are able to find a context or a topic, to which the utterance is relevant. I will call this the *source of relevance* for an utterance. As discussed in section 3.1.4, there are many ways of characterizing the notion of source of relevance, which is itself a way of describing how hearers integrate the utterance they hear with the context. Sperber and Wilson (1986/1995) deal with the notion of source of relevance in terms of a set of contextual assumptions and contextual implications that ensue from integrating the content of an utterance with the contextual assumptions. In Grice (1989) the source of relevance for an utterance takes the form of a *purpose* that the conversation is aimed at achieving. In Roberts (1996/2012) it takes the form of a question, which the utterance addresses. In what follows I will present these views in more detail.

Grice (1989) sees conversation as a form of purposeful action and in his account, the current *purpose* of the conversation directs the hearer's expectations about what the speaker will say

and the interpretation of their utterances (see section 3.1.1 for a detailed discussion of Grice's account). This purpose gives sense to a speaker's utterances in the same way knowing I am observing someone baking a cake gives sense to the actions involved in the process. The same actions seen by someone who doesn't know what baking a cake involves would leave the observer utterly confused. Since in Grice's view conversation is a *concerted enterprise*, this purpose is a goal which participants in the conversation share and are pursuing by means of their conversation. Knowing this purpose is essential in inferring the intentions of interlocutors and therefore the interpretation of their utterances. For example, if I need to reach the train station and my interlocutor tells me that "The station is somewhere in the city centre" I will infer that they mean that they do not know its exact location. This is because the purpose of our conversation is to help me reach the station and I expect them to try to give as much information as is required to achieve the purpose.

Russel (2012) discusses how, in Grice's account, the hearer's expectations are tied to a particular source of relevance. Relevance or informativity are not general properties of utterances: utterances are relevant or informative with respect to a particular point. Consider for example the expectations related to the first maxim of quantity. Russel argues that if speakers were expected to be maximally informative with their utterance in general (i.e. not relatively to a particular point) there would always be a more informative utterance that they could have made (e.g. speakers can always add more information to an utterance by conjunction). However, "Speakers are simply expected to make a contribution that is maximally informative *about the point they are making*" (Russel, 2012, p. 108). In sum, the source of relevance takes the form of a shared goal in Grice's account and it plays an important role in utterance interpretation as it guides both the speaker's behaviour and the hearer's expectations about the speaker, in particular the purpose of the conversation is fundamental for the inference of implicatures (see section 3.1.1).

As discussed in section 3.1.4, Roberts' theory of discourse (1996/2012, 2004) represents the source of relevance for an utterance as its Question Under discussion (QUD), which is at the same time a question and a goal in the sense that once interlocutors accept a QUD it is their shared goal to answer it. QUDs are semantic questions which can be explicitly asked by an interlocutor of inferred from the context. In what follows I will use the expression *source of relevance* interchangeably with *QUD*, which has gained much currency in recent years. However, in using this term I do not wish to commit to Robert's theory of discourse. I also remain open to the possibility that questions may not be a good way of characterising the source of relevance as Breheny (to appear) suggests. As a semantic object, a question is intended as a partition of the logical space of possibilities given by their possible answers (Groenendijk and Stokhof, 1984).

For example, the polar question of *whether Rome is bigger than London* is represented semantically as dividing the space of possibilities in two: cases were Rome is bigger than London and cases where Rome is not bigger than London. If this partition view is an accurate model of language processing, hearers should activate both these states to the same degree when processing a polar question. Instead, Tian & Breheny (2016) found evidence that in processing polar questions, representations related to the positive state (i.e. *that Rome is bigger than London*) becomes more activated than representations related to the negative state. Breheny (to appear) takes this result to support an alternative view where the hearer reconstructs the state of inquiry which would have prompted the interlocutor to ask the question that the utterance seems to address. For example, the utterance '*Rome is not bigger than London*' would make the hearer reconstruct a state of inquiry where it is relevant to know that *Rome is bigger than London*. Incidentally, this is also exactly the state that would prompt the question of *whether Rome is bigger than London*, which in Robert's terms is the QUD for the original utterance. In sum, Breheny's proposed notion of state of inquiry seems closely related to the notions of source of relevance even though Breheny does not does not explicitly draw an equivalence between these two notions.

Determining the source of relevance of an utterance is important for arriving at the right interpretation; experimental research on the role of QUD in interpretation provides evidence to support this claim. Zondervan, Meroni and Gualmini (2008) provide evidence that in interpreting sentences that contain scope ambiguities and scalar implicature triggers, hearers tend to choose an interpretation of the sentence that addresses the current QUD over one that does not address it. In a similar vein, Degen and Goodman (2014) found that in reading a sentence containing a scalar implicature trigger such as 'some' after a QUD inquiring whether the stronger alternative is the case, readers are much more likely to calculate the implicature compared to reading the same sentence after a QUD inquiring whether a weaker alternative is the case. For example, if one reads the sentence Jimmy ate *some* of the cookies after the question inquiring whether Jimmy ate *all* of the cookies, they are much more likely to arrive at the interpretation that he ate *some but not all* of them, than if the question were inquiring on whether Jimmy ate *any* of the cookies. In the same vein, Kehler and Rohde (2016) discovered that expectations regarding what QUD an utterance containing pronouns is addressing affects what reference is assigned to those pronouns. Finally, Clifton and Frazier (2012) demonstrated that in the course of reading a sentence readers find it more difficult to read a sentence completion that does not address the current QUD compared to one that does, which indicates that QUD also affects the readers' expectations regarding the content an utterance. All of these studies indicate that the

choice of QUD has a large impact on the interpretation of the utterance and especially regarding the pragmatic inferences involved in the interpretation.

Given the importance of the source of relevance in utterance interpretation, it is interesting to address the question of how hearers recover the source of relevance. There is both theoretical and experimental work about how an utterance's QUD, when it is not an explicit question, can be recovered from the linguistic material and information structure of the utterance itself. In Roberts' (1996/2012) theory prosodic focus acts as a clue to the QUD. Each focus marked constituent of a declarative sentence presupposes a question. For example, the sentence '*John cooked a PASTA',* with sentence stress falling on 'pasta', presupposes the question '*What did John cook?'* Conversely, '*JOHN$_F$ cooked pasta*' presupposes the question '*Who cooked pasta?*' Therefore, the focus structure of an utterance is itself an indication to what QUD the utterance may be addressing. Another clue to the QUD of an utterance, according to Tian, Breheny and Ferguson (2010), is negation. More specifically, Tian and colleagues propose that an assertion containing sentential negation usually accommodates the question of whether its positive counterpart is the case. For example, the sentence *John didn't cook pasta* can accommodate the QUD of *whether John cooked pasta*. Tian and colleagues also provided experimental evidence for the hypothesis that negative sentences involve accommodating a positive QUD. They found that after reading a sentence like *John didn't cook pasta,* participants were faster to recognize an image that was inconsistent with the sentence but consistent with the positive argument of the negation (e.g. cooked pasta) compared to an image that was consistent with the overall meaning of the sentence (e.g. uncooked pasta). They interpreted this result as an indication that participants were recovering the QUD as a step in the process of interpreting the utterance. Kehler and Rohde (2016) unveiled further clues are used to recover the likely QUD of an utterance by asking participants to fill in an utterance in discourses containing different types of implicit causality verbs such as "John infuriated Bob. _____ " and "John scolded Bob. _____ ". They also varied whether the utterance to complete began with an ambiguous pronoun (i.e. "John scolded Bob. He _____ ") or not. They found that participants were more likely to produce utterances addressing QUDs regarding what the subject of the first sentence did (e.g. John infuriated Bob. *He broke Bob's favourite mug.*) if the first sentence contained an implicit causality verb that pointed to the subject as the cause of the event (e.g. *John infuriated Bob*) or if the second sentence began with an ambiguous pronoun. This suggests that both the likely reference of a pronoun in the preceding discourse as well as semantic information related to the actions described in the preceding utterances (i.e. implicit causality) affect what QUD hearers anticipate an utterance to be addressing.

Linguistic clues are not the only resource used to identify the source of relevance for an utterance. Although the utterance itself and the preceding discourse can be important indicators of what question is being addressed there are many other ways in which hearers keep track of the context and the source of relevance. For example, non-linguistic utterances do not offer any linguistic clues to their sources of relevance yet addressee are usually able to interpret them even if these utterance do not follow a discourse that indicates what question they are addressing. In drawing this parallel between the interpretation of linguistic and non-linguistic utterances I side with Sperber and Wilson (1986/1995, 2015) and their proposal of a continuum between the two modes of communication, both of which fall under the umbrella of *ostensive-inferential communication*. Sperber and Wilson discuss the issue of identifying an utterance's source of relevance in terms of selecting a context for an utterance out of many possible alternatives. They argue that the hearer presumes that the utterance is optimally relevant and then looks for a source of relevance which justifies their assumption (Sperber & Wilson, 1995, p. 142). The interpretation at which the hearer arrives, and therefore the source of relevance which affords this interpretation, must be compatible with the speaker's *abilities and preferences*. This process applies to both linguistic and non-linguistic utterances alike.

Recovering the source of relevance is something that hearers normally do without even noticing. However, there are situations where it's not easy to hone in on a single source of relevance among the competing alternatives even with the help of linguistic material. Instances of pointing can provide good examples of these situations as they do not offer linguistic clues to what QUD they are addressing. Imagine that a friend and I are walking down a street dotted with restaurants. We have plans to have dinner together, and we agreed pick a place while we walk. At the same time we are discussing where I should host my birthday party that is coming up. After a pause, my friend points to a nice pub on the street. In this situation, my friend's utterance could be relevant to the question of where we might go for dinner or to the question of where I should host my birthday party. Normally either the preceding discourse or the relative importance of these to questions can help me decide what topic my friend is addressing, but it is easy to imagine that I might have trouble determining what my friend meant and I might want to ask for clarification.

This study addresses the question of what factors affect the hearer's choice of a source of relevance among alternatives in situations where this choice is not straightforward. Experiment 1 will address the question of whether how exhaustively an utterance addresses its QUD can affect the choice of a source of relevance for that particular utterance. This experiment will test the hypothesis that given two competing QUDs for the same utterance, hearers choose to interpret the utterance as addressing the question to which it can give the most exhaustive

answer. Experiment 2 will tie this question to the leitmotif of the experimental studies presented in this thesis which has been to investigate communication in non-cooperative contexts. Indeed, Experiment 2 will investigate how the effect of exhaustivity on the choice of QUD interacts with the possibility that the speaker is deceptive. As in these experiments the aim is to create a scenario where two different QUDs are plausible for the interpretation of an utterance, the utterance presented in the experiment is a pointing gesture, which does not give any linguistic cues as to which QUD it may be addressing. Another peculiarity of the paradigm used in these experiments is the use of an aversive stimulus (the image of a spider). As it will become clear in the explanation of the paradigm below, the aversive stimulus was used to manipulate the salience of the two QUDs with the aim of making them equally salient.

7.2   Experiment 1

In order to test the hypothesis of whether exhaustivity with respect to the QUD can affect the choice of a source of relevance for the utterance, this experiment employs a task in which participants have to interpret a non-linguistic utterance (a pointing gesture), which could be addressing either of two standing questions under discussion. This paradigm involves different scenarios and it manipulates whether the utterance is an exhaustive answer to the two QUDs in each scenario.

The task is presented as a box guessing game. In each scenario there is a set of boxes: some of the boxes contain a spider and some are empty. Participants have to choose an empty box in order to win. A character who is introduced before the game gives them a piece of information by pointing to one of the boxes. The character's pointing gesture could address the question of which box contains a spider, and so be interpreted as pointing to a box containing a spider, or the question of which box is empty and so be interpreted as pointing to an empty box. The number of boxes containing a spider varies so that in some scenarios the character's gesture is an exhaustive answer to the QUD of which box contains a spider and a non-exhaustive answer to the QUD of which box is empty, and vice versa in other scenarios. For example, if the scenario has three boxes of which only one contains a spider, pointing to one box can give an exhaustive answer to the question of which box contains a spider but not to the question of which box is empty. The opposite is true in a scenario with three boxes of which two contain a spider and only one is empty.

The experiment involves a spider, which is an aversive stimulus, on purpose. In the scenario presented to participants, it is quite counter-intuitive for the communicator to point to what they do not need or the box they should not choose. This intuitive bias works against the experimental design goal to make the two questions of which boxes are empty and which boxes

contain a spider equally salient a priory of the experimental manipulation. The choice of an aversive stimulus has the aim of increasing the salience of the question of which box contain a spider.

According to Roberts (1996/2012) QUDs are organised by their entailment relationships. The two QUDs of *which box contains a spider* and *which box is empty* are entailed by the question of *what each box contains.* This is because any answer to the question of *what each box contains* will also answer the two sub-questions of *which box contains a spider* and *which box is empty*. In each scenario, the *most exhaustive interpretation* of the utterance (i.e. the interpretation that gives the most exhaustive answer to its QUD) is also the interpretation which goes the furthest in answering the question of what each box contains. For example, a set of three boxes of which only one contains a spider has three possible states: the spider could be in the first, second or third box. Pointing to the box containing the spider rules out two of the three possible states, while pointing to an empty box rules out only one (e.g. pointing to the first box rules out only the state in which the spider is in the first box, but it does indicate whether the spider is in the second or third box).

The prediction tested in this experiment is that the exhaustivity of the utterance with respect to each QUD will push participants towards the more exhaustive interpretation in each scenario. Given that the two QUDs of *which boxes contain a spider* and *which boxes are empty* may not have equal salience a priori of the manipulation, despite the use of an aversive stimulus, each scenario will be compared to a baseline where the two interpretations of the utterance are equally exhaustive by design.

7.2.1   Methods

7.2.1.1   Materials and design

Participants played a guessing game where they were shown one of the scenarios in **Table 12** and they were asked to choose a box. They were told that if they chose an empty box they would win and if they chose a box containing a spider they would lose. Before they made their choice a character gave them *a piece of information* by pointing to one of the boxes in the set. In the context of the game, the character's utterance could either be addressing the question of *which box is empty/safe* or the question of *which box contains a spider*. Depending on which source of relevance participants choose, the character can be interpreted as pointing to a box containing a spider or to an empty box. For convenience, I will refer to these two interpretations of the utterance as *spider-interpretation* and *empty-interpretation*. The scenario was manipulated so that the spider-interpretation was the most exhaustive choice in half of the scenarios and the empty-interpretation in the other half, except for the baseline scenario, where both

interpretations are equally exhaustive. The baseline scenario gives an indication of the salience of the two interpretations without experimental manipulation. The most exhaustive interpretation in each scenario is also the one which goes further in answering the overarching question of what each box contains, and so the interpretation that rules out the highest number of possible states of the boxes (see **Table 12**).

Table 12 *Scenarios used in the experiment*

Scenario	Open box view	Clue view	States ruled out by spider-intepretation	States ruled out by empty-interpretation
Baseline			1 out of 2	1 out of 2
1 spider 2 empty			2 out of 3 (more exh.)	1 out of 3
2 spider 1 empty			1 out of 3	2 out of 3 (more exh.)
2 spider 8 empty			36 out of 45 (more exh.)	9 out of 45
8 spider 2 empty			9 out of 45	36 out of 45 (more exh.)
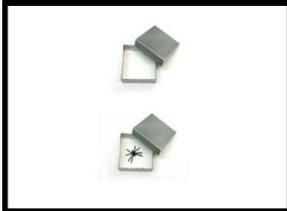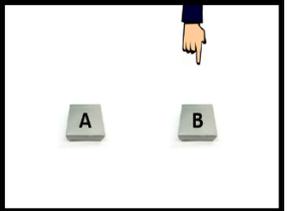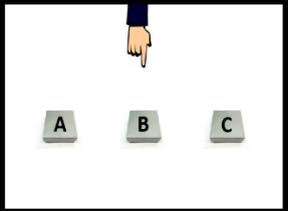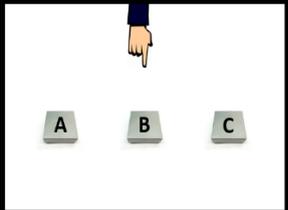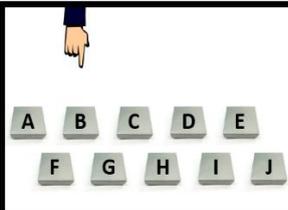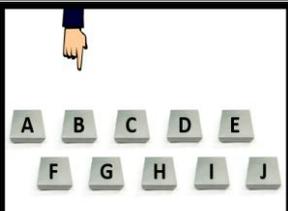
299 native English speakers were recruited online (163 females, Mean age = 29.83). Participants were told that they would see a set of boxes and they would be asked to choose one. Before the task participants were introduced to a fictional character (David) and they were shown a picture of the character (see ***Figure 10***). They were told that "*David is an honest guy. When it's time for you to choose a box, David will give you a piece of information. David can't talk but he can point.*"

Figure 10 *The image of David shown to participants*

Participants were randomly assigned to one of the five scenarios in **Table 12** and they were told that if they chose an empty box they would *win* and if they chose a box containing a spider they would *lose*. Each participant saw only one scenario and they were unaware of the existence of other scenarios. Each scenario was seen by approximately 60 participants. Participants were shown the set of open boxes, some of which contained a spider. The sets of boxes were digitally assembled photographs of the same box which contained a fake plastic spider in some of the pictures (see Open box view in **Table 12**). Accompanying the picture of the open boxes were the following instruction:

> *In this task you will see 2 boxes and you will have to choose one.*
> *As you can see below:*
> > *1 box is empty*
> > *1 box contains a spider*
> *If you choose the empty box you win*
> *If you choose the box containing a spider you lose*

The instructions reported above are for the baseline scenario, the numbers and the number agreement were different in each scenario depending on the number of empty boxes and boxes

containing a spider. Participants were then shown the set with the boxes closed and they were told that "*Now we shuffled the boxes. David knows what's in each box and on the next page he'll give you a piece of information.*" On the same page they were asked if they remembered how many of the boxes contained a spider; if they answered incorrectly they were prevented from continuing the task. On the following page participants were shown a picture of the boxes closed and labelled with letters where David's hand was pointing to box B (see clue view in **Table 12**). On the same page they were asked to choose a box by selecting the corresponding letter from a drop down list (e.g. list: A, B, C). Participants were not asked to click on the box to avoid the possible bias of matching the gesture of the character's hand and clicking on box B. For the full set of instructions see **Appendix D**.

After they made their choice they were taken to a new page where they were asked to answer the following multiple choice questions:

- What do you think David meant by his clue? (This box has a spider – Avoid this box / This box is empty- choose this box / I wasn't sure)
- In choosing a box you mostly… (based your decision on David's clue / ignored David's clue and chose at random)
- Did you think that David was trying to mislead you? (yes/no)

7.2.2   Results

For each participant the experiment collected their choice of box (see *Figure 11*), which was coded as *pointed* (i.e. choosing *the* box that was pointed to) or not *non-pointed* (i.e. choosing *a* box that was not pointed to), and their answers to the three follow up questions (*Table 13*).
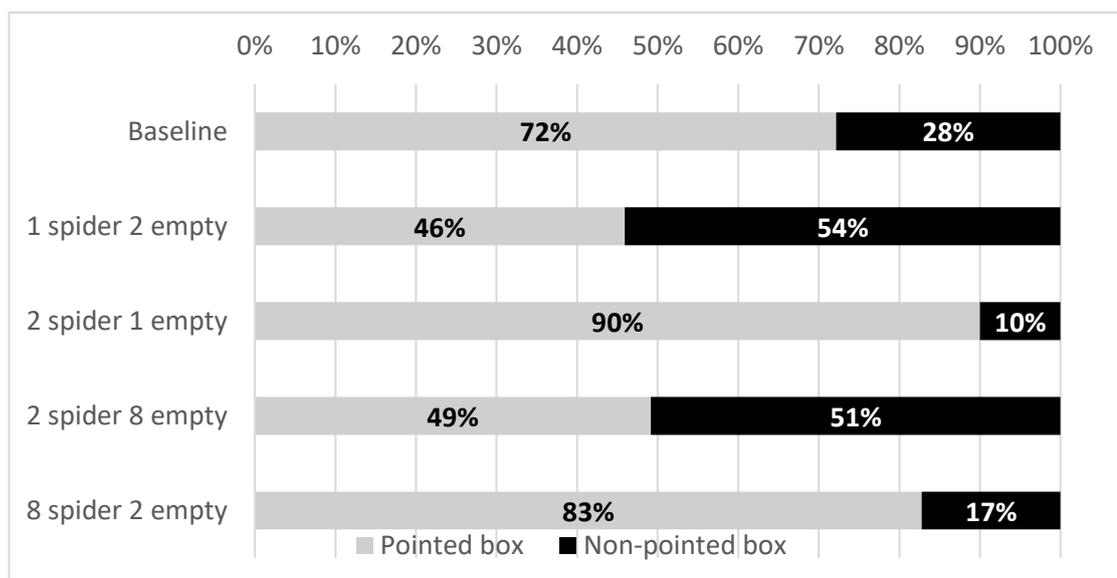
Figure 11. *Box choices in each scenario.*

Table 13. *Frequencies of answers to the Meaning, Clue use and Mislead questions in each scenario*

Scenario	Meaning question			Clue use		Mislead	
	Spider	empty	Not sure	Used clue	Chose random	yes	no
1 spider 1 empty	17.92%	73.58%	8.49%	91.80%	8.20%	9.84%	90.16%
1 spider 2 empty	27.88%	40.38%	31.73%	83.61%	16.39%	16.39%	83.61%
2 spider 1 empty	2.38%	76.19%	21.43%	86.67%	13.33%	15.00%	85.00%
2 spider 8 empty	15.79%	47.37%	36.84%	69.49%	30.51%	10.17%	89.83%
8 spider 2 empty	5.04%	72.27%	22.69%	79.31%	20.69%	17.24%	82.76%
Total	13.18%	62.57%	24.25%	82.27%	17.73%	13.71%	86.29%

The answers of participants to the Clue use question and the Mislead question (**Table 13**) indicate that the vast majority of participants reported that they used the character's clue in choosing a box and that they did not believe that the character was trying to mislead them.

Although the baseline scenario was designed with the aim of making the two interpretations of the character's utterance equally salient, participants' box choices and answers to the meaning question for that scenario indicate that this is not the case. The higher rates of non-pointed box response in the box choice and empty-interpretation responses in the meaning question suggest that the empty-interpretation was more salient than the spider-interpretation in the baseline scenario. In fact, comparing the box choices in the baseline scenario with chance level reveals that the box choices are significantly different from chance (X^2(1, n=61)=11.95, P<0.001).

In order to address the question of whether the exhaustivity manipulation had an effect on box choice and the interpretation of the utterance, the distribution of box choices and the distribution of responses to the meaning question of each scenario was compared with the baseline scenario. The box choices were not compared with chance level because the chance of choosing the cued box at random is determined by the number of boxes in each scenario rather than by the relative salience of the two interpretations, and therefore a comparison with chance would not address the question of how the experimental manipulation affects the salience of the two interpretations. For the responses to the meaning question only the responses that indicated one of the two interpretations were analysed and not the 'not sure' responses because including them may increase the chance of finding a difference in the distribution of responses even though no predictions hinge on how the manipulation would affect the rate of participants who failed to infer a meaning.

In the 2-spider-1-empty scenario and in the 8-spider-2-empty, where the *empty*-interpretation is the most informative interpretation 18% and 10% more participants respectively chose the

pointed box compared to the baseline scenario. The distribution of box choices in these two scenario was significantly different from baseline for the 2-spider-1-empty scenario (X^2(1, n=60)=9.53, P=0.002) but not for the 8-spider-2-empty (X^2(1, n=58)=3.25, P=0.07). For these two scenarios, the shift in the responses to the meaning with respect to the baseline was more in terms of avoidance of the non-exhaustive interpretation than preference for the spider-interpretation, with an increase in the rate of participants choosing the 'not sure' option compared to baseline. 15% less participants in the 2-spider-1-empty scenario and 13% in the 8-spider-2-empty scenario chose the empty-interpretation. The distribution of responses to the meaning question was significantly different from baseline in both the 2-spider-1-empty scenario (X^2(1, n=51)=16.72, P<0.001) and in the 8-spider-2-empty (X^2(1, n=49)=9.36, P=0.002).

In the 1-spider-2-empty scenario and in the 2-spider-8-empty scenario, where *spider-*interpretation is the most informative interpretation, more participants chose the non-pointed box compared to the baseline (26% and 23% more respectively). The distribution of box choices was significantly different from the baseline both for the 1-spider-2-empty scenario (X^2(1, n=61)=20.88, P<0.001) and for the 2-spider-8-empty scenario (X^2(1, n=59)=15.49, P<0.001). As for the other two scenarios, the main difference in the responses to the meaning with respect to the baseline was the decrease of participants choosing non-exhaustive interpretation: 33% less and 26% less participants chose the empty-interpretation respectively. The distribution of responses to the meaning question was significantly different from baseline in the 1-spider-2-empty scenario (X^2(1, n=60)=14.46, P<0.001), but not in the 2-spider-8-empty scenario (X^2(1, n=45)=1.07, P=0.3).

Given that some participants (17.73%) indicated in their answers to the clue use question that they ignored the clue and chose a box at random, the box choice responses were re-analysed after excluding these participants (***Figure 12***).
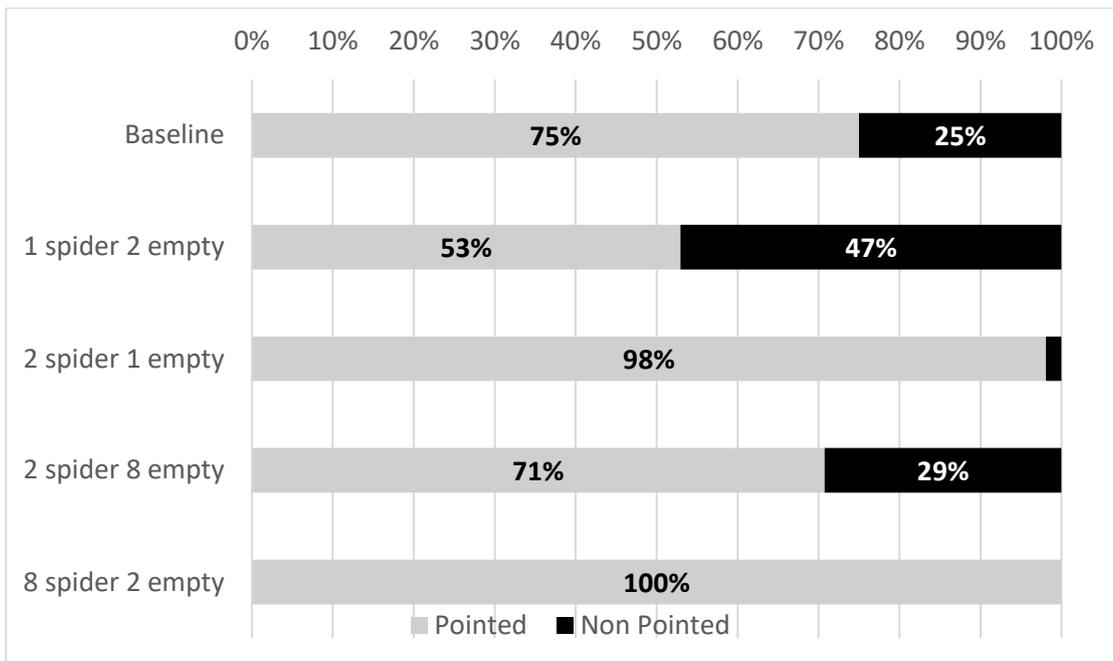
Figure 12. *Box choices of participants who indicated having used the clue in the Clue Use question*

In the scenarios where the *empty*-interpretation is the most informative interpretation, re-analysis confirmed that the distribution of box choices was significantly different from baseline for the 2-spider-1-empty scenario (X^2(1, n=53)=15.10, P<0.001) and, in contrast with the first analysis, found that the 8-spider-2-empty was also significantly different from the baseline (X^2(1, n=46)=15.33, P<0.0001). As for the scenarios where the where *spider*-interpretation is the most informative, re-analysis confirmed that the distribution of box choices was significantly different from the baseline for the 1-spider-2-empty scenario (X^2(1, n=51)=13.23, P<0.001) and, in contrast to the first analysis, it found no evidence that the 2-spider-8-empty scenario was different from the baseline (X^2(1, n=41)=0.4, P<0.528).

7.2.3   Discussion

The design of baseline scenario was not successful in making the two questions of *which box contains a spider* and *which box is empty* equally salient. It seems that the question of *which box is empty* was more salient a priory and led participants to prefer the empty-interpretation. The rationale of using an aversive stimulus such as the spider in order to increase the salience of the question of *which box contains a spider* was based on the *negativity bias*. Previous research has repeatedly shown an attentional negativity bias that leads us to pay more attention to negative stimuli compared to positive ones (Eastwood, Smilek & Merikle, 2001; Carretié, Mercado, Tapia & Hinojosa, 2001; Dijksterhuis & Aarts, 2003; Smith, Cacioppo, Larsen & Chartrand, 2003). It seems that the attentional 'boost' of the aversive stimulus negativity the question of *which box is empty* was still the most salient *a priori*. Many factors are at play in determining the salience

of a QUD in a particular context and in the last part of this discussion I will speculate on what factors may have caused this initial bias towards the question of *which box is empty*.

In spite of this unexpected bias in the baseline scenario, the results suggest that the exhaustivity manipulation had an effect. In the scenarios where the empty-interpretation was the most exhaustive (2-spider-1-empty, 8-spider-2-empty) participants were less likely to choose the spider-interpretation compared to baseline and they were more likely to choose the box that was pointed to, at least considering the re-analysis of the box choice data. The pattern of box choices in the 8-spider-2-empty scenario also showed a shift towards the box that was pointed also in the first analysis but the distribution was not significantly different from the baseline scenario. In the 1-spider-2-empty scenario, where the spider-interpretation was the most exhaustive, participants were more likely to choose the box that was not pointed to compared to the baseline scenario and they were less likely to choose the empty-interpretation compared to the baseline scenario. In the 2-spider-8-empty scenario, where the spider-interpretation was again the more exhaustive, there were also small differences in the expected direction for both box-choices and meaning question responses, but they were not significantly different from the baseline scenario, at least considering the re-analysis of the box-choice responses.

For the interpretation of box-choices I will focus on the second analysis of the data in which participants who indicated that they ignored the character's clue and chose a box at random are excluded. Including these participants would actually obfuscate the results as is suggested by the fact that excluding them brings the box-choices more in line with the responses to the meaning question. The re-analysis shows a clear difference between the 8-spider-2-empty scenario and the baseline, which matches the difference in meaning-question responses, and a lack of significant difference between the 2-spider-8-empty scenario and the baseline scenario, which matches the lack of a difference in the responses to the meaning question.

In the 2-spider-8-empty scenario, which does not differ from the baseline scenario either for box-choices or for responses to the meaning question, 37% of participants indicated that they were not sure of the interpretation of the utterance. This rate of 'not sure responses' is higher than in any other scenario and a four-fold increase compared to the baseline. One notable feature of this scenario is that, in contrast to the other scenarios, pointing to the spider is not a 'safe' move for the communicator as it does not guarantee that the other player will not choose a box containing a spider. Pointing to an empty box is always a 'safe' move and pointing to the spider when there is only one spider is also 'safe' as the other player, provided that they interpret the utterance as intended, will be able to avoid the spider. Pointing to a spider when there is more than one spider, as in the 2-spider-8-empty scenario, leaves the other player to

133

choose among boxes which carry a probability, however low, of containing a spider. If participants assumed that the communicator had a preference for 'safe' strategies, which is quite likely, the effect of assumption would be in direct contrast with the intended effect of the exhaustivity manipulation. This contrast where the more exhaustive interpretation corresponded to a 'risky' strategy on the part of the communicator probably contributed to the high rate of the 'not sure' responses and the lack of significant differences with the baseline in the 2-spider-8-empty scenario.

In the responses to the meaning question, the decrease of participants selecting the non-exhaustive interpretation in each scenario compared to the baseline did not correspond to a comparable increase in the number of participants choosing the exhaustive interpretation, while the number of 'not sure' responses often increased considerably. It seems that the exhaustivity manipulation was more successful in causing avoidance the non-exhaustive interpretations than in creating preference for the exhaustive interpretation. A simple explanation for this feature of the results might be that the higher number of boxes in the experimental scenarios made the task slightly more complicated and contributed to participants' uncertainty. An alternative explanation may be related to the initial bias for the empty-interpretation, that is evident in the baseline scenario, and that may have repercussions in the other scenarios.

Overall these results indicate that exhaustivity is a factor which affects the hearer's choice of QUD for the utterance they are interpreting. These results are not presented as evidence for or against Grice's and Roberts's accounts as they do not discuss which factors affect the choice of a source of relevance, or at least no extra-linguistic factors in the case of Roberts. These accounts could however be extended to explain how interlocutors make use exhaustivity to individuate the current purpose of the conversation. One possibility could be to posit that the Gricean maxims of conversation, and the expectations they cause in the hearer, are not only instrumental in inferring implicature but also in guiding the hearer to the current purpose of the conversation. So in the case of this experiment, the participants' expectation that the other player would be abiding the maxims of Quantity and therefore be exhaustive influenced their reconstruction of what QUD was being addressed. This solution is not far from Sperber and Wilson's proposal, which is in essence that hearer's choice of a source of relevance is determined by the hearer's expectations of relevance. Their account is in fact compatible with the results of this experiment as it seems reasonable to assume that the source of relevance that will make the utterance more informative will give rise to more cognitive effects and so it will result in a more relevant interpretation. For example, in the scenario with one spider and two empty boxes interpreting the character's utterance as addressing the question of *which boxes contain a spider*

(i.e. the spider-interpretation) may results in more cognitive effects (e.g. due to knowing both which box to choose and what each box contains) compared to the empty-interpretation.

Although this study focused on how exhaustivity affects the process of choosing a QUD and an interpretation, other factors influence this process. Despite the fact that the experimental design and choice of an aversive stimulus like the spider were aimed at making the two QUDs equally salient in the baseline condition, the question of which box was empty was more salient *a priori*. And even in the conditions where the exhaustivity manipulation was pushing participants towards a spider interpretation, preference for the non-pointed box barely reached 50% and preference for the spider-interpretation response only reached 30%. The difficulty of making the question of *which box contains a spider* salient enough may be the reflection of other factors besides exhaustivity that are influencing the behaviour of participants. One such factor could be a low level bias to go where a friendly hand is pointing, which would cause participants to be attracted to the pointed box. This may be related to subpersonal processes by which cues *orient* our attention even before we are able to detect these cues (Posner, 1980). If such a bias was perceived as common ground between the two players it may help them coordinate on the QUD of which boxes are empty.

Furthermore, participants may have also used other strategies for inferring what QUD is being addressed that do not involve exhaustivity. For example, besides the expectation that the utterance would be maximally informative with respect to the QUD, participants may have had an expectation that the character would point to what the participant *needs* (i.e. an empty box) rather than what they do *not need*. This view would also offer an explanation for how the exhaustivity manipulation affected asymmetrically the 8-spider-2-empty scenario, where both strategies pushed participants to choose the pointed box, and 2-spider-8-empty scenario where two strategies pushed in opposite directions.

In conclusion, this experiment provides evidence that exhaustivity of the utterance as an answer to the QUD affects the choice of the source of relevance for the utterance and the interpretation of the utterance itself. Although this study focuses on exhaustivity, the view proposed here is exhaustivity affects the choice of a source of relevance among many other factors. I proposed that in this task these other factors may include a low level bias to "go" where a hand is pointing and the expectations that pointing picks out what the addressee *needs* rather than what they do *not need*.

7.3   Experiment 2

Experiment 1 addressed the question of whether exhaustivity affects hearers' selection of a source of relevance for the utterance of a reliable speaker. However, hearers also select a source

of relevance when the interlocutor is unreliable. Experiment 2 explores this issue by addressing the question of how deception interacts with the effect that exhaustivity has on the selection of a source of relevance and an interpretation for an utterance. In order to address this question Experiment 2 uses exactly the same paradigm as Experiment 1 but the character who gives participants a clue is introduced as *dishonest*.

The fact that this second experiment involves a deceitful communicator makes it relevant to an aspect of the account of Epistemic Vigilance proposed by Sperber et al. (2010). Sperber and colleagues separate the comprehension process of an utterance from the epistemic assessment of its content. They argue that when talking to an untrustworthy interlocutor hearers interpret their utterances *as if they were trustworthy* and then judge whether they are likely to be true. In the setting of the present experiment this account predicts that participants' interpretation of the utterance (i.e. their responses to the meaning question) should not differ between Experiment 1 and Experiment 2, but participants in Experiment 2 should not trust the content of the utterance and therefore their box choices should be different from the choices of participants in Experiment 1.

7.3.1 Methods

290 participants were recruited online (179 females, Mean age= 27.9 years). The design, scenarios and procedure of Experiment 2 were exactly the same as in Experiment 1 with the only exception that the character (David) was presented as a *dishonest guy* rather than *an honest guy*. Participants of Experiment 1 and 2 were recruited from the same crowdsourcing website (prolific.co.uk) and at the same time. They were randomly assigned to participate in one experiment or the other. As in Experiment 1, each participants were randomly assigned to see only one scenario. Each scenario was seen by roughly 58 participants. It was ensured that participants in the two experiments came from the same population and were tested under the same conditions in order to compare their responses.

7.3.2 Results

Each participant's box choice (see **Error! Reference source not found.**) was collected and coded as pointed (i.e. choosing the box that was pointed to) or not non-pointed (i.e. choosing a box that was not pointed to). Participants answers to the three follow up questions were also collected (***Table 14***).
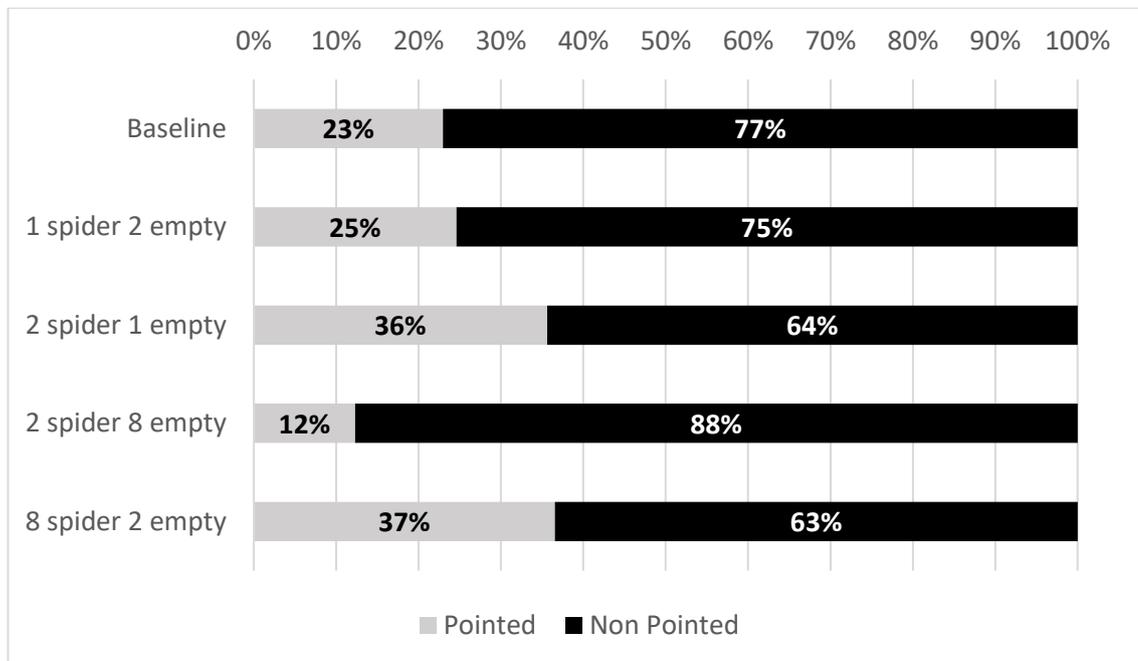
Figure 13 *Box choices in each scenario*

Table 14 *Frequencies of answers to the Meaning, Clue Use and Mislead question in each scenario*

Scenario	Meaning question			Clue use		Mislead	
	Spider	empty	Not sure	Used clue	Chose random	yes	no
Baseline	20.75%	62.26%	16.98%	75.41%	24.59%	85.25%	14.75%
1 spider 2 empty	18.97%	39.66%	41.38%	59.02%	40.98%	73.77%	26.23%
2 spider 1 empty	10.00%	55.00%	35.00%	64.41%	35.59%	79.66%	20.34%
2 spider 8 empty	19.09%	34.55%	46.36%	42.11%	57.89%	73.68%	26.32%
8 spider 2 empty	7.89%	42.11%	50.00%	50.00%	50.00%	84.62%	15.38%
Total	15.19%	46.64%	38.16%	58.62%	41.38%	79.31%	20.69%

In contrast to Experiment 1, the responses to the Mislead question (***Table 14***) indicate that the vast majority of participants believed that the character was trying to mislead them, which was expected given that the character was presented as dishonest. A considerable number of them reported that they ignored the character's clue in choosing a box and chose at random. As for Experiment 1, the distribution of box choices for baseline scenario was significantly different from chance level (X^2(1, n=61)=17.85, P<0.001).

The question of whether participants' box choices differed in the honest and dishonest condition was addressed by comparing each scenario across the two conditions. The distributions of responses was significantly different across the two conditions for all scenarios: there was a significant difference in the baseline scenario (X^2(1, n=122)=29.58, P<0.001), 1-spider-2-empty scenario (X^2(1, n=122)=6.07, P=0.014), 2-spider-1-empty scenario (X^2(1, n=119)=37.79, P<0.001),

2-spider-8-empty scenario (X^2(1, n=116)=18.04, P<0.001) and in the 8-spider-2-empty scenario (X^2(1, n=110)=24.60, P<0.001).

The question of whether participants' interpretation of the utterance differed across the two conditions was addressed by comparing the responses to the meaning question (excluding the *not sure* responses) in each scenario across the two conditions. There were no significant differences in any of the scenarios except for the 2-spider-1-empty scenario (X^2(1, n=96)=7.83, P=0.005).

The effect of exhaustivity in the dishonest condition was investigated by comparing each scenario of the dishonest condition with the baseline scenario in terms of box choices and responses to the meaning question. The responses to the meaning question in each scenario showed a shift relative to the baseline in the same direction as in Experiment 1. The shift however was smaller than in Experiment 1 and there was no significant differences with the baseline for any of the scenarios. In terms of box choices, only the scenarios where the empty-interpretation was the most informative were significantly different from the baseline. Both the 2-spider-1-empty scenario (X^2(1, n=59)=5.33, P=0.021) and the 2-spider-8-empty scenario (X^2(1, n=52)=5.42, P=0.019) showed a significant increase in the number of pointed box choices compared to the baseline scenario.
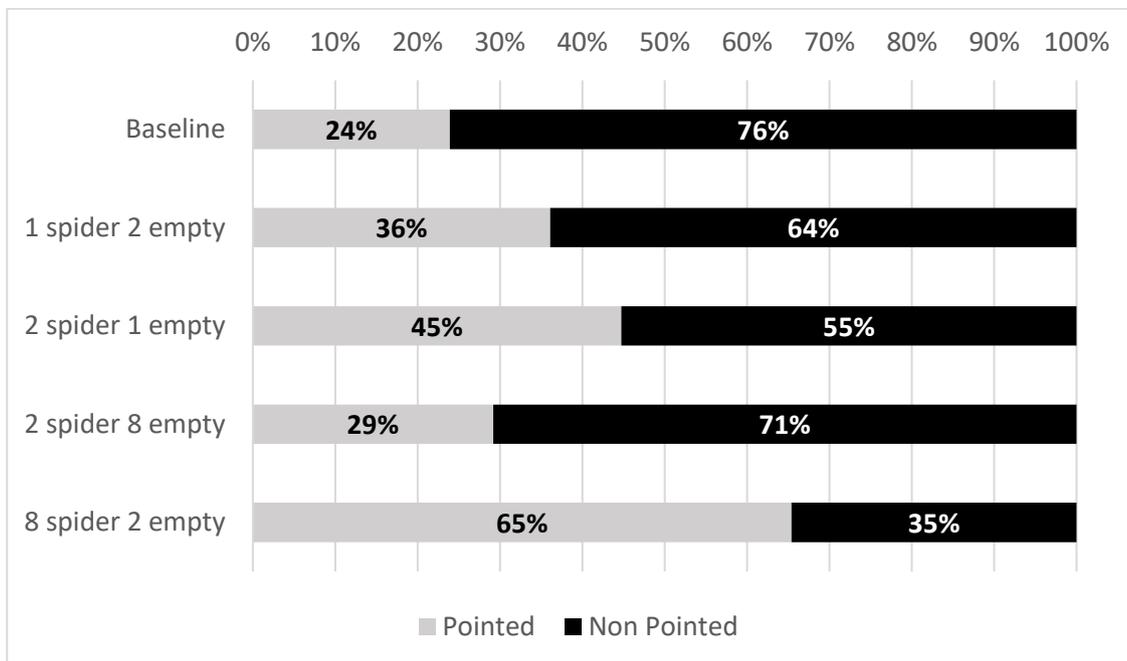
Figure 14 Box choices of participants who indicated having used the clue in the Clue Use question

As in Experiment 1, the box choice data was re-analysed after excluding the participants who indicated that they had ignored the character's clue in choosing a box and they had chosen at random in the Clue Use question (**Figure 14**). The re-analysis confirmed the results of the first

analysis: only the 2-spider-1-empty scenario (X^2(1, n=38)=9.06, P=0.002) and the 2-spider-8-empty scenario (X^2(1, n=26)=24.57, P<0.001) were significantly different from the baseline scenario.

7.3.3 Discussion

Participants' box choices in Experiment 1 and Experiment 2 were significantly different as participants in each scenario were much more likely to choose the non-pointed box in Experiment 2 than in Experiment 1. However there were no corresponding differences in the responses to the meaning question, with the exception of the 2-spider-1-empty scenario. I do not have a principled explanation for the significant difference in the 2-spider-1-empty scenario, however I can speculate that given the extremely low rate of spider-interpretation responses which this scenario elicited in Experiment 1, any noise introduced by the added feature of deception may have made the preference for one response less prominent and caused the observed difference. Overall this pattern of results seems to support the view of Sperber et al. (2010) that the interpretation process is not influenced by whether the speaker is trustworthy, while the epistemic assessment, which in this experiment comes to light in participants box choices, may be influenced. This is also in line with other studies pointing to a dissociation between comprehension and acceptance such as the study by Mazzarella, Trouche, Mercier and Noveck (2016) and the experiments I presented Chapters 4 and Chapter 5.

The exhaustivity manipulation in Experiment 2 had a significant effect on box choices only in the two scenarios where the empty interpretation was the most exhaustive (the 2-spider-1-empty scenario and the 8-spider-2-empty scenario) As for the responses to the meaning question, the scenarios in Experiment 2 followed a similar pattern to the scenarios in Experiment 1, but the differences relative to the baseline scenario were not as large and none of them were significant. It is interesting that the differences in box choices found in the 2-spider-1-empty and the 8-spider-2-empty scenarios are in the same direction as the differences found in Experiment 1: participants tend to choose the box that is pointed to more often than in the baseline scenario. This result does not sit well with the fact that the task involved deception. One possible explanation for this finding is that while participants in Experiment 1 had clear intuitions about the interpretation of the utterance in these two scenarios and almost unanimously chose the pointed box, participants in Experiment 2 might have thought that deceiving the other player by pointing to a spider would be too predictable and they may have inferred that the character was double-bluffing.

As for the responses to the meaning question, which would have provided the more telling evidence for an effect of exhaustivity, there are two factors that may have worked against the

chance of finding an effect of the manipulation. Firstly the very high rate of 'not sure' responses in Experiment 2 compared to Experiment 1 decreased the number of responses in the categories which were analysed and therefore reduced the power of the analysis. Secondly the element of deception may have made the task more complicated for participants, causing more noise in the responses as well as boosting the rate of 'not sure' responses. So, in conclusion, even though the box choices seem to be affected by the manipulation in two scenarios, these results do not provide clear evidence for an effect of exhaustivity.

7.4   Conclusion

The two experiments presented in this chapter looked at the effect that exhaustivity has on the choice of a QUD that an utterance might be addressing and its interpretation. Experiment 1 tested the hypothesis that given two possible QUDs for the same utterance, hearers interpret the utterance as addressing the question to which it can give the most exhaustive answer. The results of this experiment are in support of this hypothesis. At least when faced with an honest communicator, exhaustivity with respect to the QUD is a factor in choosing what QUD an utterance is addressing. The view proposed here is that exhaustivity is a factor among the many constraints which help the hearer determine the source of relevance for the speaker's utterance. This view echoes the so-called *constraint based* accounts which propose that listeners take multiple cues into consideration in deciding whether the interpretation of an utterance is likely (Brown-Schmidt & Hanna, 2011; Degen & Tanenhaus, 2015). This experiment therefore aims to add to previous research on what factors affect the choice of QUD in interpretation (Tian, Breheny & Ferguson, 2010; Kehler & Rohde, 2016).

Experiment 2 explored how the effect of exhaustivity interacted with the knowledge that the speaker is deceptive. This experiment did not find evidence for the hypothesis that exhaustivity affects the choice of QUD when the speaker is unreliable. In the previous section I discussed a number of reasons that may have contributed to this null result. Experiment 2 also addresses a topic that is common to all the experimental studies presented in this thesis: communication in a non-cooperative context. The expression *non-cooperative* applies only loosely to this experiment as the instructions or the task give no information regarding the speaker's motives and whether they may share the goal that the participant chooses a winning box. However, deception can be considered in itself a feature of non-cooperative communication. The comparison of the results of the two experiments presented in this chapter and the fact that the utterance of the dishonest speaker was interpreted in the same way as the utterance of the honest one but elicited different box choices can be seen as another demonstration of the dissociation between utterance interpretation and epistemic evaluation of its content (Sperber et al., 2010; Mazzarella, 2015a). It seems that participants in Experiment 2 interpreted the

character's utterance *as if it were trustworthy* (i.e. in the same way as participants in Experiment 1) and then tended to reject the content that the utterance communicated.

This Chapter presented the last two experimental studies of this thesis. While previous chapters focused on Quantity implicatures, the experiments presented in this chapter focused on how addressees infer an utterance's source of relevance. The next chapter will discuss the findings of all these studies together and it will offer a view of how these findings address the bigger issues concerning the relationship of cooperation and communication.

8 Conclusion

8.1 Summary of findings

This thesis investigated the role of cooperation in communication. It looked at communication in non-cooperative contexts, with a particular focus on implicatures, with different paradigms.

The experiment presented in Chapter 4 addressed the question of how scalar implicatures and Ad hoc implicatures are affected in a non-cooperative situation that was operationalised in a competitive game. Participants played the role of receivers in a signalling game where they had to choose one of two cards in each round and their co-player gave them short descriptions of the winning card. In the cooperative condition the signaller was an ally, in the competitive condition the signaller was an opponent and allowed to lie, in the competitive-truthful condition the signaller was an opponent but they were not allowed to lie. The results indicated that the maxim of quality was suspended in the competitive condition, as participants did not expect the signaller to give truthful descriptions. The distribution of the data in the competitive condition also suggested that participants were drawing implicatures from the signaller's utterance. In the cooperative-truthful condition participants did not think that the signaller would lie and they also calculated and trusted the implicatures arising from the signaller's utterances to some extent. I interpreted this result as indicating that participants considered false implicatures to be lies to some degree and therefore considered the content of implicature trustworthy and protected by the no-lying rule in the competitive-truthful condition. In this respect scalar implicatures and Ad hoc implicatures differed, with scalar implicatures being 'trusted' more than Ad hoc implicatures. The results of this experiment indicate that hearers infer implicatures from the utterance of an uncooperative speaker. The difference in the extent to which the contents of scalar and Ad Hoc implicatures were trusted in the competitive-truthful condition suggest that false implicatures arising from lexicalised scaled are considered to be lies (and therefore part of the truth-evaluable content of the utterance) to a greater extent than false Ad hoc implicatures.

The experiment presented in Chapter 5 addressed the question of whether hearers infer less implicatures from an uncooperative speaker who is not manifestly opting out. Participants were asked to read a context which presented a character as either cooperative or uncooperative. Participants then read an utterance of the character which could give rise to the scalar implicature of *some* and they were asked to indicate: (i) whether they thought that *all* was possible given what the character said, (ii) whether they thought the character meant to communicate the *not-all* implicature and (iii) whether they thought the character was trying to mislead them. Participants who faced an uncooperative character were more likely to indicate

142

that they thought that *all* was possible given what the character has said and more likely to indicate that the character was misleading but they were not less likely to indicate that the character intended the *not-all* implicature compared to participants who faced a cooperative speaker. I interpreted these results to indicate that hearers do draw implicatures from uncooperative speaker even though they may reject the content of the implicatures afterwards. This finding exemplifies the dissociation between the processes of comprehension of acceptance of the implicature (Sperber et al. 2010; Mazzarella, 2015a, 2015b) and reinforces the findings of the experiment presented in Chapter 4.

The experiment presented in Chapter 6 addressed the question of how speakers communicate when they are in a non-cooperative context. As in the experiment presented in Chapter 4, the non-cooperative context was operationalised as a competitive signalling game, except that in this case participants played the role of signallers. Participants saw sets of two cards, a winning card and a losing card. They were asked to complete descriptions of the winning card and they were told that a receiver would use their descriptions to decide which of the two cards was the winning card. The descriptions templates invited two obvious completions that resulted in either true or false assertions in the case of control items, and completions that gave rise to either true or false implicatures in the case of experimental items. Experimental items employed three types of quantity implicatures: scalar implicature (i.e., *exact* reading) of numerals, scalar implicature of the quantifier *most* and Ad hoc implicatures. Participants were told that the receiver would be told that they were opponents and that signallers were allowed to lie. Knowing that the hearer would be distrustful pushed signallers to overall produce true and false descriptions at the same rate, both for control items and for experimental items. This matches the overall expectations of receivers in the competitive condition of the experiment presented in Chapter 4. Participants also gave a relatively high rate of uninformative completions for control items and items with numerals. In this and most other respects, items with numerals patterned with control items rather than *most* and *Ad hoc* items. I interpreted this pattern of results in favour of the view that the *exact* interpretation of numerals is part of their truth-conditional meaning (Carston, 1998; Breheny 2008) and against the view that the *exact* reading arises in context as an implicature (Horn, 1972; Gazdar, 1979; Levinson, 2000). Although participants chose true and false completions at the same rate overall, a cluster analysis suggested that different groups of participants followed different strategy profiles. A large group of participants preferred false completions for both control and experimental items and expected their receivers to perform worse as a result. This suggests that these participants were expecting their hearers to infer and trust implicatures arising from their descriptions.

Chapter 7 focused on the topic of how hearers infer the source of relevance (or QUD) for an utterance. Experiment 1 addressed the question of whether the exhaustivity of an utterance with respect to the possible QUDs in the context influences the choice of the source of relevance used to interpret the utterance. Participants were shown a set of boxes where one (or more) of the boxes contained a spider. They were told that if they chose the box containing the spider they would loose and if they chose an empty box they would win. Before they made their choice an *honest* character gave them a piece of information by pointing to one of the boxes. Participants could interpret the pointing gesture to be addressing the question of *which box contains a spider* or the question of *which box is empty*. The number of spider-boxes and empty boxes varied in different scenario and this affected whether the pointing utterance was more exhaustive as an answer to one question or the other. Each scenario was compared to a baseline with only one spider-box and one empty box, where the utterance was an equally exhaustive answer to both questions. The results showed that the exhaustivity affected the choice of the source of relevance. In scenarios where the utterance could address the question of *which box is empty* more exhaustively participants were more likely to choose the box that the character pointed to (compared to the baseline) and they were more likely to indicate that they interpreted the utterance as pointing to the empty box (compared to the baseline). The reverse happened in scenarios where the utterance could answer more exhaustively the question of *which box contains a spider*. Although the baseline scenario aimed to make the two QUD equally salient participants were heavily biased to interpret the utterance as addressing the question of *which box is empty*. I suggested that this does not affect the interpretation of the exhaustivity effect but it indicates that exhaustivity is only one of the many factors that affect the choice of QUD and other factors may have brought about this bias.

Experiment 2 used the same paradigm as Experiment 1, but the pointing character was presented as *dishonest*. This second experiment aimed to explore how the choice of QUD and the effect of exhaustivity on the choice of QUD were affected by a deceptive speaker. The results of this experiment offered very limited evidence of an effect of exhaustivity on the choice of the source of relevance. The most interesting aspect of the results of Experiment 2 is that the box choices of participants were significantly different from box choices of participants in Experiment 1 but their interpretation of the character's utterance (i.e. what they thought the character meant to communicate) was not different from the interpretation given by participants in Experiment 1. Participants in Experiment 2 mostly indicated that they thought the character was trying to mislead them. This pattern of results is again an example of the dissociation between comprehension and epistemic assessment of the communicated content (Sperber et al. 2010; Mazzarella, 2015a). Participants in both experiments interpreted the

utterance as if the communicator was trustworthy, but while participants in Experiment 1 accepted the content of the utterance, participants in Experiment 2 rejected it.

8.2 Theoretical implications and outlook

8.2.1 Implicatures in non-cooperative contexts

The studies presented in this thesis addressed the question of the role of cooperation in the derivation of implicatures. The findings of these studies support the proposition that hearers do infer implicatures from speakers that they believe to be uncooperative. The results of the experiments presented in Chapter 4 and Chapter 5 speak to this point. Furthermore, the results of the production study presented in Chapter 5 indicate that uncooperative speakers expect hearers to infer implicatures arising from their utterances even though they have reason to believe that they will be seen as uncooperative from their audience. This proposition is in contrast with the prediction I derived from Grice's (1989) account that hearers should not infer implicatures from the utterances of an uncooperative speaker. In section 3.2 I explained how this prediction applies to contexts where the fact that the speaker is uncooperative is known to both hearer and speaker but it is not mutually known or not part of the common ground. It could be argued that some of the participants in the experiments presented in this thesis may have constructed the speaker's unwillingness to cooperate to be mutually known between themselves and their interlocutor. This however does not change the prediction derived from Grice's account. In a situation where the speaker's unwillingness is mutually known to interlocutors, either because this is apparent from the context or because the speaker makes it apparent, the speaker should be seen a *opting out* and therefore the hearer should not infer implicatures from the speaker's utterances.

The proposition that hearers infer implicatures from uncooperative speakers is not incompatible with Relevance Theory (Sperber & Wilson, 1986/1995). According to this account, hearers are guided by an expectation that the speaker's utterances will be *optimally* relevant and not by an assumption that the speaker is cooperative with regards to the hearer's goals. This means that hearers will infer a relevant interpretation of the speaker's utterance, including any implicatures that this may involve, regardless of whether the speaker is cooperative or uncooperative. This however does not entail that hearers always infer relevant implications even in cases where the speaker may have *preferred* not to communicate them. Since the resulting interpretation has to be *compatible with the speaker's abilities and preferences* (Sperber & Wilson, 1995) hearers do not infer implicatures that are incompatible with the speaker preferences, such as implicatures related to a topic about which the speaker manifestly does not wish to give information.

Another proposition that is supported by the results of the studies presented is that there is a dissociation between the processes of comprehension and acceptance of communicated content (Sperber et al. 2010; Mazzarella, 2015a). This also applies to implicatures arising from the utterances of uncooperative / untrustworthy speakers and it touches on an important methodological point raised by Mazzarella (2015b). The measures of experimental studies of implicatures may reflect the epistemic acceptance of the implicature content but not whether the implicature was inferred (e.g., Bonnefon, Feeney & Villejoubert, 2009). With these measures it is not possible to distinguish between cases where the hearers infer and then reject an implicature because they consider the speaker untrustworthy, and case where hearers do not infer implicatures at all. Experiments that investigate implicatures in non-cooperative context and employ measures based on epistemic acceptance (e.g., Chapter 4; Pryslopska, 2013) may show that hearers accept the content of implicatures of uncooperative speakers less compared to implicatures of cooperative speakers. However, one should be careful not to conclude from this that they are inferring less implicatures from the utterances of the uncooperative speakers. In fact, the experiment presented in Chapter 5 demonstrates that hearers do not seem to infer less implicatures from cooperative speakers compared to cooperative ones even though they are less likely to accept the content of the uncooperative speaker's implicatures. The difference in the results of the two experiments presented in Chapter 6 provides an example of the dissociation between comprehension and acceptance with a different kind of pragmatic inference.

One aspect that the experiments presented in Chapter 4 and Chapter 5 have in common is the fact that the design of the experimental paradigms aimed to recreate a situation where the hearer believed the speaker to be uncooperative but they did not believe them to be manifestly opting out in Grice's terms. This was pursued by making sure that the 'uncooperative' goals of the speaker were clear to hearers and at the same time by not having the speaker signal that they were opting out. However, as I point out in section 3.2.1, there is a provision in Grice's account for a situational opting out, in which the speaker does not signal that they are opting out but it is clear from the situation and therefore mutually manifest to all interlocutors that they are opting out. It could be argued that at least some of the participants in the experiments presented in Chapters 4 and 5 if not all of them may have interpreted the speaker as opting out because the way in which the competitive game was presented supported a situational opting out. If this was the case, the predictions for both Grice and Relevance Theory would be that participants should not have drawn any implicatures as the speaker's unwillingness to be informative was part of the common ground between speaker and hearer. This would make the results of the two experiments surprising and possibly incompatible with both theories. I say

'possibly' incompatible because in fact Relevance Theory has a second route to the derivation of the scalar inferences which may allow their derivation even in cases where the speaker's unwillingness to be informative is mutually manifest among interlocutors.

Relevance Theory hypothesises two routes for scalar inferences. One route is the global inference route, in which implicatures that are calculated on the basis of the proposition explicitly expressed. This global derivation is triggered when there is some "mutually manifest, actively represented reason to wonder" (Noveck & Sperber, 2007, p. 10) whether the stronger alternative is true, or when the stronger alternative "was being entertained as a relevant possibility" (Noveck & Sperber, 2007, p.11). The second route is by local enrichment, whereby the relevance expectations of the hearer in a specific context push the hearer to narrow the meaning of the scalar expression (e.g. some narrowed to mean some and not all). This happens when the explicature appears to be insufficiently relevant for the hearer's expectations of relevance. For example, if the sentence "Some of the open questions must be answered" (see the experiment materials in Chapter 5) is used in a context where it is already part of the common ground that at least some of the open questions must be answered, the meaning of some can be locally enriched to mean some and no all. With the local route, Relevance Theory could account for hearers drawing scalar inferences (not scalar *implicatures*) as part of the explicature in situations where the speaker's lack of cooperation and unwillingness to be informative is part of the common ground. Therefore, it could account for the results concerning scalar inferences emerging from Chapter 4 and 5. It is not clear where this local route is also supposed to be available for the derivation of Ad hoc inferences. If scalar and Ad hoc inferences differed in terms of the availability of a local route, this may explain to some extent the difference between scalar and Ad hoc inferences discussed in Chapter 4. In any case it is difficult to determine a priori whether the inferences in these experiments should be considered global implicatures or local enrichments. This is because whether a scalar inference is a local enrichment or a global implicature depends on two factors that are hard to measure: (i) whether the explicature without the unenriched scalar term is sufficiently relevant for the hearer's expectations of relevance and (ii) whether among interlocutors there is a mutually manifest and actively represented reason to wonder about the stronger alternative as a relevant possibility.

In sum, the results of the experiments presented in this thesis give some indication regarding the role of cooperation in the derivation of implicatures. As discussed in section 3.2, here I mean *perlocutionary* cooperation as without *locutionary* cooperation there is simply no communication. The answer that seems to emerge from the results presented here is that cooperation is not a necessary assumption for deriving implicatures or at least the kind of pragmatic inferences used in the experiments presented in this thesis. However, cooperation is

tightly connected with the assumptions of trustworthiness that hearers make about the speaker. Therefore, cooperation (or the lack thereof) can deeply affect the hearer's epistemic assessment of what is communicated to them. This conclusion highlights once again the importance of not confounding comprehension and epistemic assessment (Mazzarella, 2015b) – and how this is particularly true when studying communication in uncooperative contexts.

8.2.2 False implicatures

When hearers infer an implicature from the utterance of an uncooperative speaker and reject its content because they believe it is false they are inferring that the speaker is communicating a false implicature (Meibauer, 2014). The results of the experiment presented in Chapter 5 indicate that hearers expect uncooperative speakers to communicate false implicatures and the results of the production study presented in Chapter 6 suggest that these expectations are correct as uncooperative speakers do communicate false implicatures. False implicatures are in themselves compatible with Grice's account, but the fact that a hearer would infer an implicature from an uncooperative speaker and regard it as false because they see the uncooperative speaker as untrustworthy is not compatible with the prediction derived from his account. This scenario can instead be explained by the view that comprehension and acceptance are distinct processes (Sperber et al. 2010; Mazzarella, 2015a). The results of the experiment presented in Chapter 4 indicate that hearers consider false implicatures to be part of the truth-conditional content of an utterance to some extent. This means that a speaker could *lie* by uttering a true statement that gives rise to a false implicature in context. This suggests that drawing a sharp distinction between lies and false implicatures based on Grice's distinction between *what is said* and what is implicated may not reflect lay judgements of what counts as a lie (Meibauer, 2005, 2014; Dynel, 2011, 2015).

Although I presented some of the results in this thesis as illustrating the dissociation between the processes of comprehension and acceptance I do not take these results to also support the view that the comprehension process is completely independent from the mechanisms of Epistemic Vigilance. In fact, in the Relevance Theoretic framework a hearer deriving a false implicature from a speaker that they know to be deceitful may have to entertain interpretive hypotheses that they already know to be false because these hypotheses are compatible with the speaker's preferences (i.e. their intention to deceive). Therefore, as Mazzarella (2015c, 2016) points out, the same Epistemic Vigilance mechanisms for the evaluation of the speaker's competence and benevolence seem to have an early effect on the interpretation process as they allow the hearer to modulate their expectations of relevance according to the speaker's *abilities and preferences*. Indeed it seems that in cases where the hearer needs to adjust their expectations of relevance according to the preferences of a deceitful speaker, the Epistemic

Vigilance mechanisms that afford this adjustment are necessary for the hearer to derive false implicatures.

The results of the experiment presented in Chapter 4 also show a difference between false scalar implicatures arising from lexicalised scales and false Ad hoc implicatures in the extent to which they are considered to be lies. False scalar implicatures were more likely to be considered lies than false Ad hoc implicatures. This result also ties in with the findings of a study presented by Katsos (2009) where participants found under-informative statements which could give rise to false scalar implicatures to be a worse "offence" than under-informative statements which could give rise to false Ad hoc implicatures. If the distinction between falsely implicating and lying is taken to reflect the distinction between what is implicated and *what is said*, this result are also suggests that scalar implicatures are more likely to be incorporated into '*what is said'* in Gricean terms than Ad hoc implicatures. In this perspective, this result adds to existing evidence that implicatures can be incorporated into the truth-conditional content of an utterance (Doran, Baker, McNabb, Larson & Ward, 2009; Doran, Ward, Larson, McNabb & Baker, 2012) and it can be taken to support the view that implicatures can be interpreted as being part of the explicit content of the utterance (Noveck & Sperber, 2007). Although this result may reveal something about how these two types of inferences are perceived it does not in itself support an argument for or against a unified view of implicatures arising from lexicalised and non-lexicalised scales because it does not give any information about how these two types of implicatures are processed.

8.2.3   Source of relevance

This thesis also touched on the topic of how hearers infer the source of relevance of an utterance. Although the source of relevance of an utterance is often formalised as a Question Under Discussion (van Kuppervelt, 1996; Roberts, 1996; Ginzburg, 2012) I proposed that the notion of source of relevance can be treated as theory neutral. In this perspective I am connecting this notion and the notion of QUD to the notion of shared goal that appears in theories that cast conversation as a cooperative activity (Grice, 1989; Clark, 1996). The first experiment presented in Chapter 7 indicates that the exhaustivity of an utterance as an answer to the possible QUDs in the context guides the hearer's choice of QUD. The view proposed here is that exhaustivity is only one among the many factors that guide hearers in inferring what source of relevance the speaker is addressing. The experiment presented aimed to add to the existing literature on what affects the process of inferring the source of relevance of an utterance (Tian, Breheny & Ferguson, 2010; Kehler & Rohde, 2016). Given the fundamental role that inferring the source of relevance plays in interpretation, under all accounts, it would be desirable that this topic gained a stronger footing as an avenue of experimental research.

8.3 Outlook

Uncooperative situations such as competitive games are the setting where interlocutors may communicate strategically. The studies presented in Chapter 4 and Chapter 6 explore strategic communication in competitive contexts where deception is expected. Although the results of both studies on the surface suggest that interlocutors produce and expect truths and lies in equal amounts in these scenarios, closer analysis of the data suggest that individual communicators may choose very different strategies. This result may be expected in a strategic setting, however it highlights the importance of taking individual differences of communicators into consideration in the study of uncooperative situations. This strategic aspect of communication is relevant in many settings, from negotiations to legal scenarios, but it is rarely studied from a linguistic perspective. Research on deception, which is a feature of strategic communication at least in the competitive scenarios considered in this thesis, has focused on the detection of detection by studying cues to deception (DePaulo et al., 2003 for a review) and accuracy in detection (Bond & DePaulo, 2006 for a review). Further research into this strategic aspect of communication would certainly shed light on how communication is used in non-cooperative situations.

The research presented in this thesis focused on a topic that has received very little attention so far: communication in non-cooperative contexts. Although an important part of the thesis focuses on testing the hypothesis that hearers do not infer implicatures from uncooperative speakers, which is of theoretical relevance, the spirit of most of the studies presented in this thesis is exploratory. Out of four studies presented, three are based on novel experimental paradigms and sets of stimuli that had not been used before and were developed especially for these experiments. Hopefully, the work that has gone into creating these paradigms and items will benefit future research on communication in non-cooperative contexts. This line of research may inform theories of communication by providing insight on the relationship between cooperation and communication and many questions revolving around this issue are still open. A general issue that this thesis does not address is that of the variety of non-cooperative situations (e.g. competitive situations, negotiations, etc.). It would be interesting to investigate whether there are valid and useful distinctions among *types* of non-cooperative situations and whether communication is affected in different ways depending on the type of non-cooperative situations. One possibility is that differences between *types* of non-cooperative situations are determined by which and how many *perlocutionary* goals are shared or conflicting. Furthermore, the experiments in this thesis focus on situations where the speaker's willingness or unwillingness to cooperate was known to interlocutors. It would be interesting to investigate situations where interlocutors are unsure about each other's goal and observe the effect of

modulating the hearer's epistemic trust both on comprehension and production. In addition, it would be interesting to explore the role of *locutionary cooperation* from an experimental point of view in order to find empirical correlate of the *locutionary-perlocutionary* distinction and test its validity. It would also be worth investigating further and with more sophisticated measures whether the comprehension process is in any way affected in a non-cooperative contexts suggesting an interplay with the epistemic assessment (contra Sperber et al., 2010). On the topic of the distinction between lying and falsely implicating, it would be interesting to explore further how different types of pragmatic inferences are perceived by hearers and speakers and how this is affected by the perceived commitment of the speaker (Morency, Oswald, & de Saussure, 2008).

9 References

Antoniou, K., Cummins, C., & Katsos, N. (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics*, *99*, 78-95.

Attardo, S. (1997). Locutionary and perlocutionary cooperation: The perlocutionary cooperative principle. *Journal of Pragmatics, 27(6)*, 753-779.

Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Bach, K. (1994). Semantic slack: What is said and more. In S. L. Tsohatzidis (Ed.), *Foundations of speech act theory: Philosophical and linguistic perspectives* (pp. 267-291.). Routledge

Bach, K. (2001). You don't say?. *Synthese*, *128*(1-2), 15-44.

Bale, A. C., Brooks, N., & Barner, D. (2010). Quantity implicature and access to scalar alternatives in language acquisition. *Semantics and Linguistic Theory, 20*, 525-543.

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1450–1460.

Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and social psychology Review*, *10*(3), 214-234.

Bonnefon, J-F., Feeney, A., & Villejoubert, G. (2009). When some is actually all: scalar inferences in face-threatening contexts. *Cognition, 112*, 249-258.

Bonnefon, J. F., De Neys, W., & Feeney, A. (2011). Processing Scalar Inferences in Face-Threatening Contexts. In L. Carlson, C. Hölscher, and T. Schipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX.

Bott, L., Bailey, T. M., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language*, *66*(1), 123-142.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*, 437–457.

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25(2*), 93-139.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalized scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100(3),* 434–463.

Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition*, *126*(3), 423-440.

Breheny, R. (to appear). Language Processing, Relevance and Questions. In in K. Scott, B. Clark & R. Carston (Eds.), *Relevance: Pragmatics and Interpretation*. Cambridge University Press

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review, 101(2),* 327-341.

Bratman, M. E. (1993). Shared intention. *Ethics*, *104(1),* 97-113.

Brennan, S. E., & Hanna, J. E. (2009). Partner-Specific Adaptation in Dialog. *Topics in Cognitive Science, 1(2)*, 274-291.

British National Corpus, (2007). BNC XML Edition, version 3. Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialogue & Discourse*, *2*(1), 11-33.

Bull, P. (2016, October 24). Theresa May has a very special technique for avoiding questions. *The Conversation*. Retrieved from: theconversation.com

Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the 'negativity bias', studied through event-related potentials. *International journal of psychophysiology*, *41*(1), 75-85.

Carston, R. (1995). Quantity maxims and generalised implicature. *Lingua*, *96*(4), 213-244.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In: Carston, R. & Uchida, S. (eds.), *Relevance Theory: Applications and Implications* (pp. 179-236). Amsterdam: John Benjamins.

Carston, R. (1999). Review of the book *Using language,* by H. H. Clark. *Journal of Linguistics, 35*, 167-171

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Carston, R. (2009). The explicit/implicit distinction in pragmatics and the limits of explicit communication. *International Review of Pragmatics, 1(1)*, 35-62.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Belletti, A. (Ed.), *Structures and beyond* (pp. 39-103). Oxford: Oxford University Press.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L.

Domingues, & A. Johansen, (Eds.), *Proceedings of the 25th Boston University Conference on Language Development* (pp. 157–168). Somerville, MA: Cascadilla Press.

Coleman, L., & Kay, P. (1981). Prototype semantics: The English word lie. *Language, 57(1)*, 26-44.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception--behavior link and social interaction. *Journal of personality and social psychology, 76(6)*, 893-910.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clifton, C., & Frazier, L. (2012). Discourse integration guided by the "question under discussion." *Cognitive Psychology, 65*, 352-379.

Coleman, L., & Kay, P. (1981). Prototype semantics: The English word lie. *Language*, *57(1),* 26-44.

Cooperation (2017). In Cambridge Dictionary. Retrieved from: http://dictionary.cambridge.org/dictionary/english/cooperation

Cooperation (2017). In OxfordDictionaries.com. Retrieved from: https://en.oxforddictionaries.com/definition/cooperation

Cummins, C. (2017). Contextual Causes of Implicature Failure. *Discourse Processes*, *54*(3), 207-218.

Cummins, C., & Rohde, H. (2015). Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in psychology*, *6*.

Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Oxford University Press.

Davies, B. L. (2007). Grice's cooperative principle: meaning and rationality. *Journal of Pragmatics, 39(12)*, 2308--2331.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations?. *Lingua, 120(8)*, 1956-1972.

Degen, J., Franke, M., & Jäger, G. (2013). Cost-based pragmatic inference about referential expressions. *Proceedings of the 35th annual meeting of the Cognitive Science Society*, 376–381.

Degen, J., & Goodman, N. (2014). Lost your marbles? The puzzle of dependent measures in experimental pragmatics. In *Proceedings of the Cognitive Science Society, 36*, 397-402.

Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive science*, *39*(4), 667-710.

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental psychology*, *54*(2), 128-133.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin*, *129*(1), 74.

Dieussaert, K., Verkerk, S., Gillard, E., & Schaeken, W. (2011). Some effort for some: Further evidence that scalar implicatures are effortful. *The Quarterly Journal of Experimental Psychology*, *64*(12), 2352-2367.

Dijksterhuis, A., & Aarts, H. (2003). On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science*, *14*(1), 14-18.

Doran, R., Baker, R. E., McNabb, Y., Larson, M., & Ward, G. (2009). On the non-unified nature of scalar implicature: an empirical investigation. *International Review of Pragmatics, 1(2)*, 211-248.

Doran, R., Ward, G., Larson, M., McNabb, Y., & Baker, R. E. (2012). A novel experimental paradigm for distinguishing between what is said and what is implicated. *Language, 88(1)*, 124-154.

Dynel, M. (2011). A web of deceit: A neo-Gricean view on types of verbal deception. *International Review of Pragmatics, 3(2)*, 139-167.

Dynel, M. (2015). Intention to deceive, bald-faced lies, and deceptive implicature: Insights into Lying at the semantics-pragmatics interface. *Intercultural Pragmatics, 12(3)*, 309-332.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Attention, Perception, & Psychophysics, 63(6)*, 1004-1013.

Fallis, D. (2009). What is lying?. *The Journal of Philosophy*, *106*(1), 29-56.

Feeney, A., & Bonnefon, J. F. (2013). Politeness and honesty contribute additively to the interpretation of scalar expressions. *Journal of Language and Social Psychology*, *32*(2), 181-190.

Feeney, A., Scrafton, S., Duckworth, A., & Handley, S. J. (2004). The story of some: everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *58*(2), 121-132.

Fox, D. (2014). Cancelling the Maxim of Quantity: Another challenge for a Gricean theory of scalar implicatures. *Semantics and Pragmatics*, *7*, 5-1.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition and logical form.* New York, NY: Academic Press.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Gilbert, D. T. (1993). The assent of man: Mental representation and the control of belief. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control Century psychology series* (pp. 57–87). Englewood Cliffs: Prentice-Hall.

Gilbert, M. (1989). *On Social Facts*. Princeton, N.J.: Princeton University Press

Gilbert, M. (2006). *A theory of political obligation: Membership, commitment, and the bonds of society*. Oxford University Press on Demand.

Ginzburg, J. (2012). *The interactive stance*. Oxford University Press.

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization, 93*, 293-300.

Grice, H. P. (1957). Meaning. *Philosophical Review, 66*, 377-388.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (eds.), Syntax and semantics 3: Speech acts, pp. 41–58. New York: Academic Press.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Grodner, D., & Sedivy, J. (2011). The effect of speaker-specific information on pragmatic inferences. *The Processing and Acquisition of Reference*, *2327*, 239–72.

Groenendijk, J. & Stokhof, M. (1984). *Studies on the semantics of questions and the pragmatics of answers* (doctoral dissertation). Institute for Logic, Language & Computation (ILLC), University of Amsterdam.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M. T. (2001). At the semantics/pragmatics interface in child language. *Semantics and Linguistic Theory, 11*, 231-247.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and cognitive processes*, *20*(5), 667-696.

Hardin, K. J. (2010). The Spanish notion of lie: revisiting Coleman and Kay. *Journal of Pragmatics, 42(12),* 3199-3213.

Hespanha, J. P., Ateskan, Y. S., & Kizilocak, H. (2000, July). Deception in non-cooperative games with partial information. *Proceedings of the 2nd DARPA-JFACC Symposium on Advances in Enterprise Control*, 1-9.

Hirschberg, J. (1985). *A Theory of Scalar Implicature* (doctoral dissertation). University of Pennsylvania, PA.

Horn, L. (1972). *On the Semantic Properties of Logical Operators in English* (doctoral dissertation). University of California, LA.

Horn, L. (1989). *A natural history of negation*. University of Chicago Press: Chicago, IL

Horn, L. (2006). Implicature. In L. Horn, G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 2–28). Blackwell: Oxford

Horn, L. (2005). The border wars: A neo-Gricean perspective. In K. Turner & K. von Heusinger (Eds.), *Where semantics meets pragmatics* (pp. 21-48). Oxford: Elsevier

Huang, Y., & Snedeker, J. (2009). On-line interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology, 58*, 376–415.

Huang, Y. T., & Snedeker, J. (2011). Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes*, *26*(8), 1161-1172.

Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What exactly do numbers mean?. *Language Learning and Development, 9(2)*), 105-129.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development, 2*, 77-96.

Isenberg, A. (1973). Deontology and the Ethics of Lying. In A. Isenberg (Ed.), *Aesthetics and Theory of Criticism: Selected Essays of Arnold Isenberg* (pp. 245–264). Chicago: University of Chicago Press

Ishihara S. (1917). Tests for Color-Blindness. Handaya, Tokyo: Hongo Harukicho

Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press

Katsos, N. (2009). Evaluating under-informative utterances with context-dependent and contextindependent scales: experimental and theoretical implications. In U. Sauerland & K. Yatsushiro (eds.), *Experimental Semantics and Pragmatics*, (pp. 51-73). Basingstoke: Palgrave Studies in Pragmatics, Language & Cognition.

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67-81.

Katsos, N., & Cummins, C. (2010). Pragmatics: from theory to experiment and back again. *Language and Linguistics Compass, 4(5)*, 282-295.

Kehler, A. & Rohde, H. (2016). Evaluating an Expectation-Driven QUD Model of Discourse Interpretation. *Discourse Processes 54(3)*), 219-238.

Kennedy, C. (2015). A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics & Pragmatics, 8(10)*), 1-44.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action: Theory and Data. In B. Ross (Ed.), *The Psychology of Learning and Motivation Vol. 54* (pp. 59-101). Burlington: Academic Press.

Krebs, J. R. & Dawkins, R. (1984) Animal signals: mind-reading and manipulation? In J. R. Krebs and N. B. Davies (eds.), *Behavioural Ecology: An Evolutionary Approach, 2nd edn*. (pp. 380-402) Oxford: Basil Blackwell Scientific Publications.

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological review 117(3)*, 785-807.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Lundquist, T., Ellingsen, T., Gribbe, E., & Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior & Organization, 70(1)*, 81-92.

Mahon, J. E. (2007). A definition of deceiving. *International Journal of Applied Philosophy, 21(2)*, 181-194.

Marty, P. P., & Chemla, E. (2013). Scalar implicatures: working memory and a comparison with only. *Frontiers in psychology*, *4*.

Marty, P., Chemla, E., & Spector, B. (2013). Interpreting numerals and scalar items under memory load. *Lingua*, *133*, 152-163.

Mazzarella, D. (2015a) *Inferential pragmatics and epistemic vigilance* (doctoral dissertation). UCL, London, UK.

Mazzarella, D. (2015b). Politeness, relevance and scalar inferences. *Journal of Pragmatics*, *79*, 93-106.

Mazzarella, D. (2015c). Pragmatics and epistemic vigilance: The deployment of sophisticated interpretative strategies. *Croatian Journal of Philosophy, 15(44)*, 183-199.

Mazzarella, D. (2016). Pragmatics, modularity and epistemic vigilance. *Argumenta 1(2)*, 181-193.

Mazzarella, D., Trouche, E., Mercier, H., & Noveck, I.A. (2016). *Believing what you are told: Politeness and scalar inferences.* Poster presented at AMLaP, Bilbao, Spain

Meibauer, J. (2005). Lying and falsely implicating. *Journal of Pragmatics, 37(9)*, 1373-1399.

Meibauer, J. (2014). *Lying at the semantics-pragmatics interface*. Berlin: Mouton de Gruyter.

Mercier, H. (2009). *La théorie argumentative du raisonnement* (doctoral dissertation). E.H.E.S.S Paris.

Millikan, R. G. (2004). *Varieties of Meaning: The 2002 Jean Nicod Lectures*. MIT Press.

Millikan, R. G. (2005). *Language: A Biological Model*. New York: Oxford University Press.

Montague, R., Navarro, D., Perfors, A., Warner, R., & Shafto, P. (2011). To catch a liar: The effects of truthful and deceptive testimony on inferential learning. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 1312-1317.

Morency, P., Oswald, S., & de Saussure, L. (2008). Explicitness, implicitness and commitment attribution: A cognitive pragmatic approach. *Belgian journal of linguistics, 22(1)*, 197-219.

Newstead, S. E., & Collis, J. M. (1987). Context and the interpretation of quantifiers of frequency. *Ergonomics*, *30*(10), 1447-1462.

Nicolle, S., & Clark, B. (1999). Experimental pragmatics and what is said: A response to Gibbs and Moise. *Cognition, 69(3)*, 337-354.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, *63*(3), 324-346.

Noveck, I. A. (2001). When children are more logical than adults. *Cognition, 86*, 253-282

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and language*, *85*(2), 203-210.

Noveck, I., & Sperber, D. (2007). The why and how of experimental pragmatics: The case of scalar inferences. In N. Roberts (Ed.), *Advances in Pragmatics*. Palgrave.

Pacherie, E. (2012). The phenomenology of joint action: self-agency vs. joint-agency. In A. Seemann (Ed.), *Joint attention: New developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 343-389). Cambridge, MA: MIT Press.

Papafragou, A. & Musolino, J. (2003). Scalar implicatures: experiments at the semantics/pragmatics interface. *Cognition, 86*, 253-282.

Pettit, P. (2003). Groups with Minds of their Own. In Schmitt, F. F. (Ed.), *Socializing metaphysics*, (pp. 167–93). Rowman and Littlefield: Lanham, MD

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of sciences, 105(3)*, 833-838.

Politzer-Ahles, S., & Fiorentino, R. (2013). The Realization of Scalar Inferences: Context Sensitivity without Processing Cost. *PLOS ONE*, *8*(5), e63943.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Pouscoulous, N., Noveck, I., Politzer, G., Bastide, A. (2007) Processing costs and implicature development, Language Acquisition, 14.4, 347-376.

Primoratz, I. (1984). Lying and the "Methods of Ethics", *International Studies in Philosophy, 16*, 35–57.

Pryslopska, A. (2013). *Implicatures in uncooperative contexts: Evidence from a visual world paradigm*. Poster presented at XPRAG, Utrecht, Netherlands.

Rakoczy, H. (2008). Pretence as individual and collective intentionality. *Mind & Language*, *23*(5), 499-517.

Rakoczy, H., & Tomasello, M. (2006). Two-year-olds grasp the intentional structure of pretense acts. *Developmental Science*, *9*(6), 557-564.

Ransom, K., Voorspoels, W., Perfors, A., & Navarro, D. (2017). A cognitive analysis of deception without lying. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, 992-997

Reboul, A. (2017). Is implicit communication a way to escape epistemic vigilance?. In Assimakopoulos S. (Ed.), *Pragmatics at its Interfaces*, (pp. 91-112). De Gruyter.

Récanati, F. (1979). *La trasparence et l'énunciation*. Seuil: Paris

Récanati, F. (2001). What is said. *Synthese*, *128*(1), 75-91.

Récanati, F. (2004). 'What is said' and the semantics/pragmatics distinction. In C. Bianchi (Ed.), *The Semantics/Pragmatics Distinction.* Stanford, CA: CSLI Publications.

Roberts, C. (1996/2012). Information structure in discourse: towards an integrated formal theory of pragmatics. In J. Yoon & A. Kathol (Eds.), *Papers in semantics (Working Papers in Linguistics 49)*. The Ohio State University. {Reproduced with minor amendments in. Semantics and Pragmatics, 5(6): 1–69.}

Roberts, C. (2004). Discourse context in dynamic interpretation. In L. Horn & G. Ward (eds.), *Handbook of contemporary pragmatic theory*. (pp. 197–220). Oxford: Blackwell

Roth, A. S. (2004). Shared agency and contralateral commitments. *The Philosophical Review*, *113*(3), 359-410.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics, 20*, 53-65.

Russell, B. (2012). *Probabilistic reasoning and the computation of scalar implicatures* (doctoral dissertation). Brown University.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and philosophy, 27(3),* 367-391.

Schwarz, F., Clifton, C., & Frazier, L. (to appear) Strengthening 'or': Effects of Focus and Downward Entailing Contexts on Scalar Implicatures. In Anderssen, Moulton, Schwarz & Ussery (Eds.), *UMOP 37: Semantic Processing*. Amherst, MA: GLSA.

Scott-Phillips, T. (2014). *Speaking Our Minds: Why human communication is different, and how language evolved to make it special*. Palgrave MacMillan.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Searle, J. (1990). Collective intentions and actions. In P. Cohen, J. Morgan, & M. M. Pollack (Eds.), *Intentions in communication*. Cambridge, MA: MIT Press.

Searle, J. R. (1995). *The construction of social reality*. New York: The Free Press

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1234-1246.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. Topics in Cognitive Science, 1(2), 305-319.

Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia*, *41*(2), 171-183.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language 25(4)*, 359-393.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2004). Relevance theory. In L. Horn, G. Ward (Eds.), *Handbook of Pragmatics.* (pp. 607-632) Oxford: Blackwell.

Sperber, D., & Wilson, D. (2015). Beyond speaker's meaning. *Croatian Journal of Philosophy*, *15*(2), 117-149.

Stiller, A., Goodman, N., & Frank, M. (2011, January). Ad-hoc scalar implicature in adults and children. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 2134-2139.

Tian, Y., Breheny, R., & Ferguson, H. (2010). Why we simulate negated information: A dynamic pragmatic account. *The Quarterly Journal of Experimental Psychology, 63(12)*, 2305–2312.

Tian, Y. & Breheny, R. (2016). *Representing Polar Questions and Inferring States of Inquiry*. XPrag Workshop: Questions Answers and Negation. SAS, Berlin.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences 35(1)*, 75-97.

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science, 10(1)*, 121-125.

Tuomela, R. (1993). What is cooperation? *Erkenntnis, 83(1*), 87-101.

Tuomela, R. (1995). *The Importance of Us – A Study of Basic Social Notions*. Stanford: Stanford University Press.

Tuomela, R. (2011). Cooperation as joint action. *Analyse & Kritik*, *33*(1), 65-86.

Tuomela, R., & Miller, K. (1988). We-intentions. *Philosophical Studies, 53*, 367-389.

Van Kuppevelt, J. (1996). Inferring from Topics: Scalar Implicatures as Topic-Dependent Inferences. *Linguistics and Philosophy*, *19*(4), 393–443.

van Tiel, B. (2014). *Quantity Matters: Implicatures, Typicality and Truth* (doctoral dissertation). University of Nijmegen.

van Tiel, B., & Geurts, B. (2013). Embedded scalars. *Semantics and Pragmatics, 6,* 1–37.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, *20*(2), 415-432.

Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind*, *111*(443), 583-632.

Zondervan, A. (2009). Experiments on QUD and focus as a contextual constraint on scalar

      implicature calculation. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and*

      pragmatics: From experiment to theory (pp. 94–112). Basingstoke: Palgrave Macmillan.

Zondervan, A. (2010a). *Scalar implicatures or focus: An experimental approach* (doctoral

      dissertation). Netherlands Graduate School of Linguistics: Amsterdam

Zondervan, A. (2010b). The role of QUD and focus on the scalar implicature of most. In J.

      Meibauer &M. Steinbach (Eds.), *Experimental semantics/pragmatics* (Vol. 175, pp.

      221–238). Amsterdam: John Benjamins.

Zondervan, A., Meroni, L., & Gualmini, A. (2008). Experiments on the role of the question

      under discussion for ambiguity resolution and implicature computation in adults. In

      Proceedings of SALT (Vol. 18, pp. 765–777).

10 APPENDICES

10.1 APPENDIX A1

Instructions shown to participants for each condition.

Instructions for the Competitive condition:

HOW THE GAME WORKS (please read carefully)

- This is a competitive game with two players: a describer and a guesser. In this game you are the guesser.
- Each round of the game has a winning card and a losing card.
- You score points when you click on winning cards.
- The describer instead scores points when you click on losing cards.
- The describer sees both cards (with the winning card highlighted in green) and they have to complete a description of the winning card as in the example below

[Chapter 3 *Figure 1*]

- In each round you'll see both cards but you won't know which one is the winning card.
- Please read what the describer wrote in each round and click on what you think is the winning card.
- The describer is allowed to lie and they are playing against you so their goal is to make you lose.

Instructions for the Cooperative condition:
HOW THE GAME WORKS (please read carefully)

- This is a cooperative game with two players: a describer and a guesser.
- In this game you are the guesser.
- Each round of the game has a winning card and a losing card.
- You score points when you click on winning cards.
- The describer also scores points when you click on winning cards.
- The describer sees both cards (with the winning card highlighted in green) and they have to complete a description of the winning card as in the example below

[Chapter 3 *Figure 1*]

- In each round you'll see both cards but you won't know which one is the winning card.
- Please read what the describer wrote in each round and click on what you think is the winning card.
- The describer is allowed to lie but they are helping you so their goal is to make you win.

Instructions for the Competitive-truthful condition:
HOW THE GAME WORKS (please read carefully)

- This is a competitive game with two players: a describer and a guesser. In this game you are the guesser.
- Each round of the game has a winning card and a losing card.
- You score points when you click on winning cards.
- The describer instead scores points when you click on losing cards.

- The describer sees both cards (with the winning card highlighted in green) and they have to complete a description of the winning card as in the example below

[Chapter 3 *Figure 1*]

- In each round you'll see both cards but you won't know which one is the winning card.
- Please read what the describer wrote in each round and click on what you think is the winning card.
- The describer is not allowed to lie but they are playing against you so their goal is to make you lose.

10.2 APPENDIX A2

List of all items used in the experiment.

Item	Description	Matching card	Other card
SCALE01	On the winning card most of the objects are ribbons		
SCALE02	On the winning card most of the objects are lamps		
SCALE03	On the winning card most of the objects are flowers		
SCALE04	On the winning card most of the objects are boats		
SCALE05	On the winning card some of the objects are brooms		
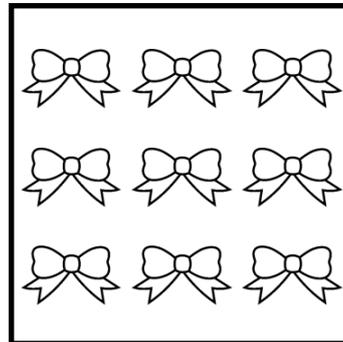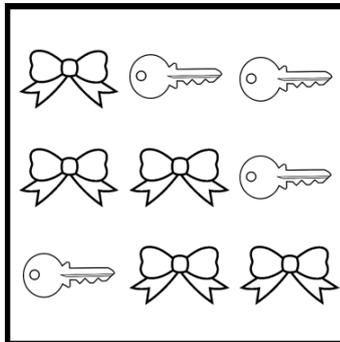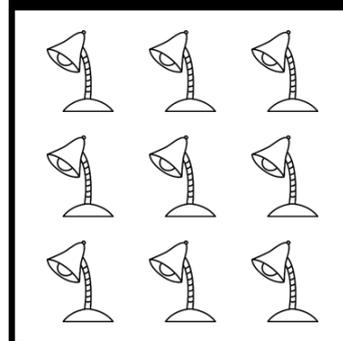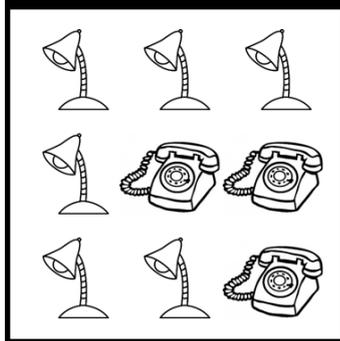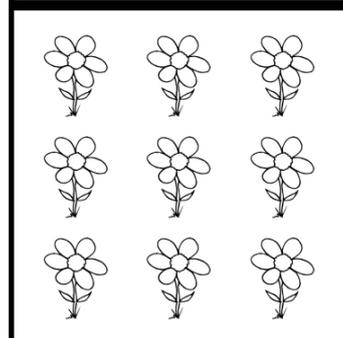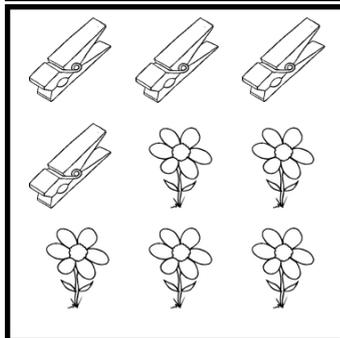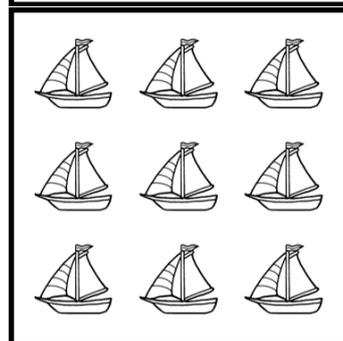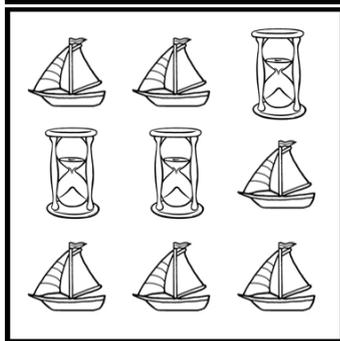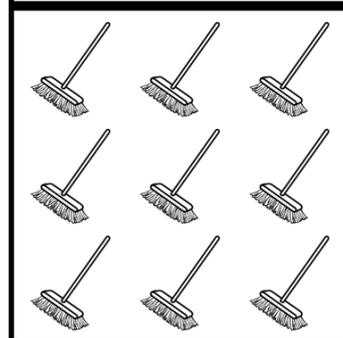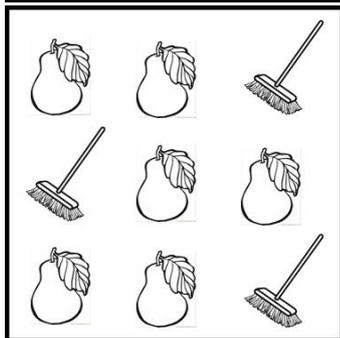
SCALE06	On the winning card some of the objects are tents		
SCALE07	On the winning card some of the objects are flags		
SCALE08	On the winning card some of the objects are cakes		
HOC09	On the winning card the top object is a fork		
HOC10	On the winning card the top object is a bike		
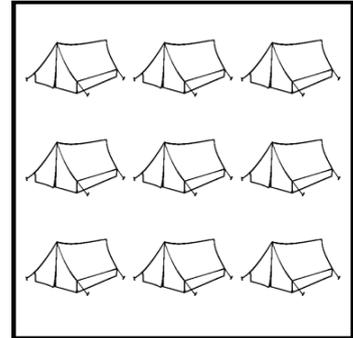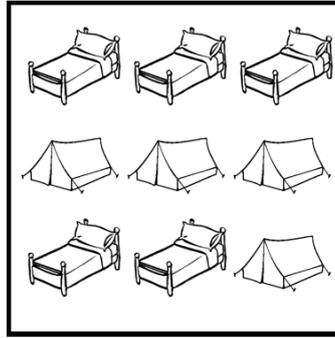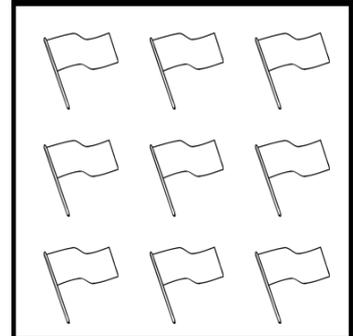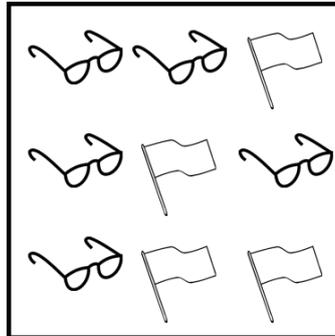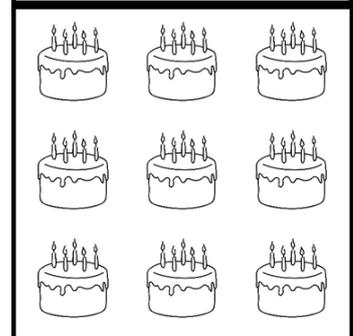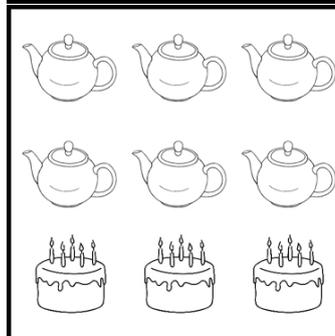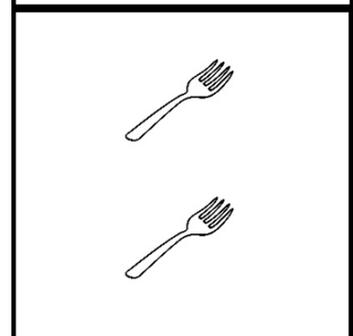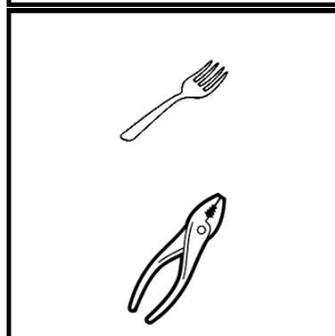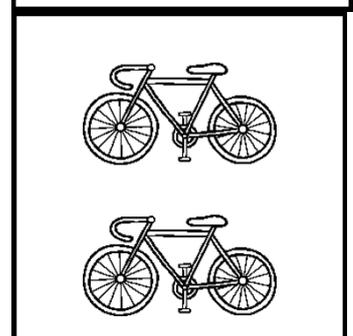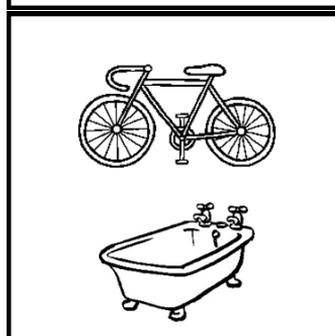
HOC11	On the winning card the bottom object is a candle		
HOC12	On the winning card the bottom object is a shoe		
HOC13	On the winning card the small object is a bucket		
HOC14	On the winning card the small object is an apple		
HOC15	On the winning card the big object is a kite		
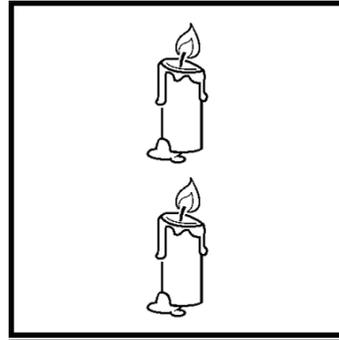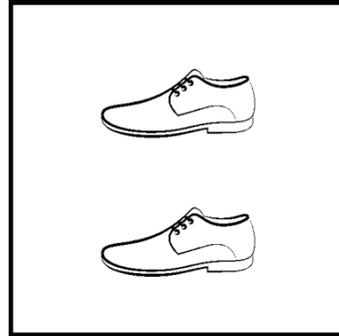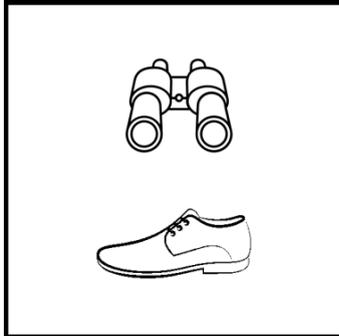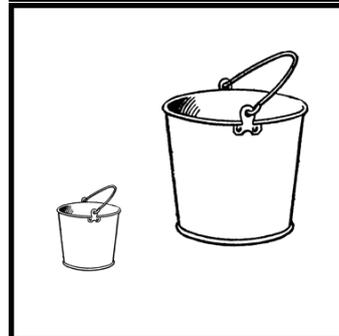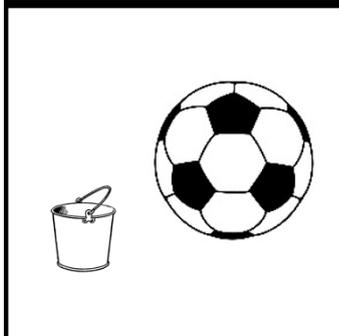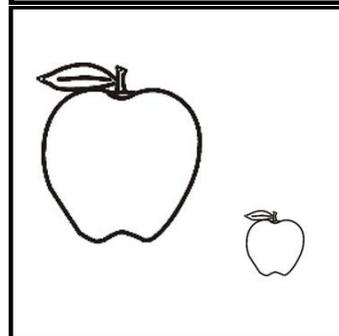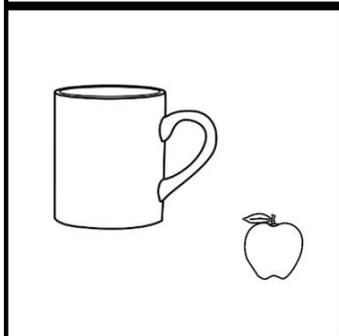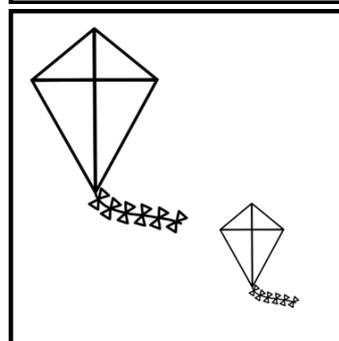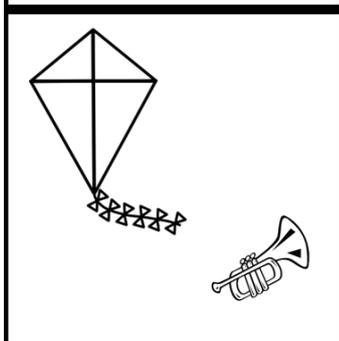
HOC16	On the winning card the big object is a sofa		
FIL17	On the winning card neither of the objects is a carrot		
FIL18	On the winning card neither of the objects is a rocket		
FIL19	On the winning card neither of the objects is a ladder		
FIL20	On the winning card neither of the objects is a sword		
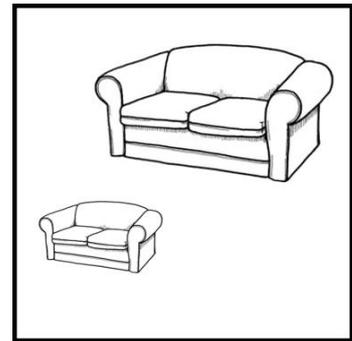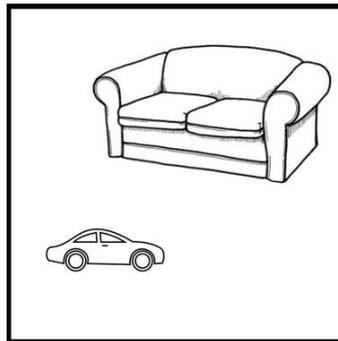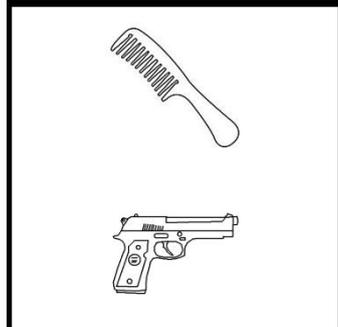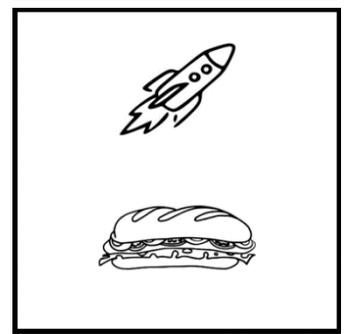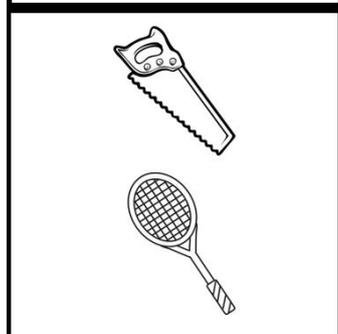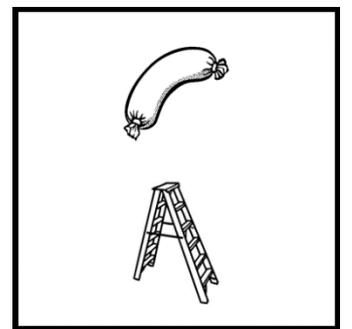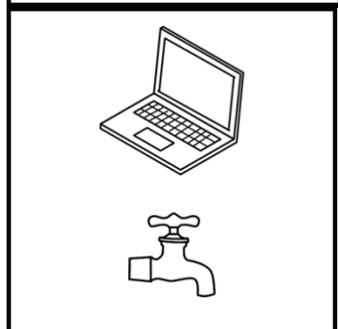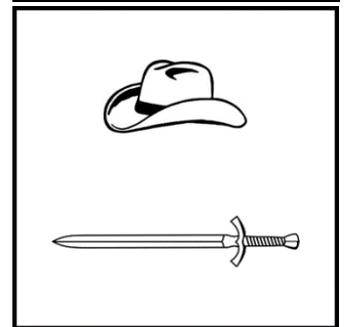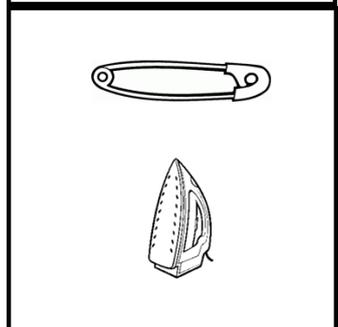
| FIL21	On the winning card neither of the objects is a camera		
FIL22	On the winning card neither of the objects is a guitar		
FIL23	On the winning card neither of the objects is a bell		
FIL24	On the winning card neither of the objects is a tower		
FIL25	On the winning card none of the objects are dice		
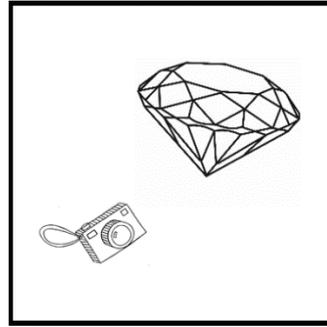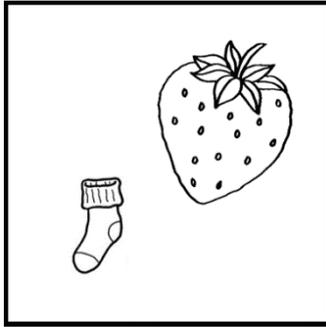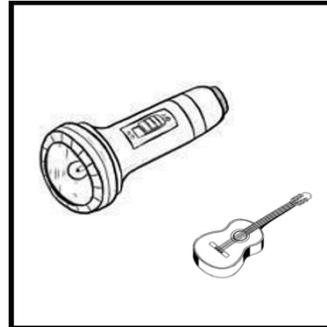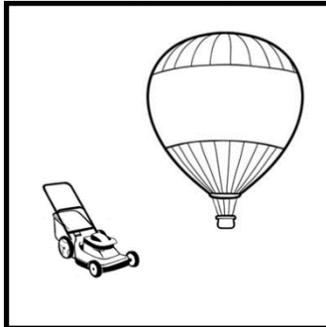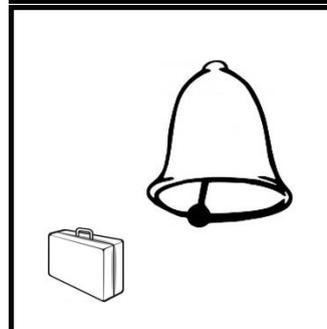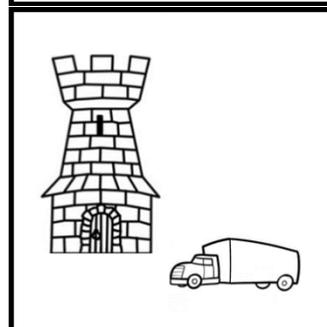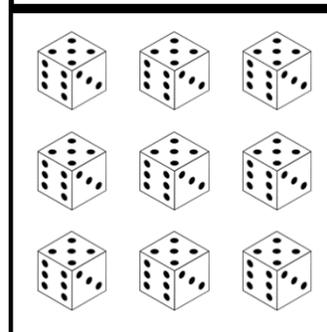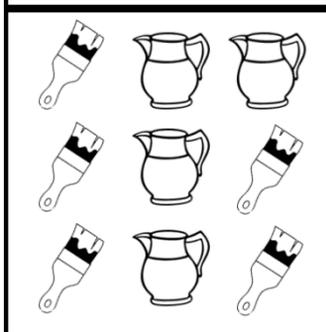
| FIL26 | On the winning card none of the objects are vases | | |
FIL27	On the winning card none of the objects are bananas		
FIL28	On the winning card none of the objects are pills		
FIL29	On the winning card none of the objects are drums		
FIL30	On the winning card none of the objects are benches		
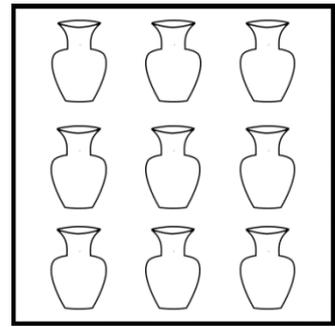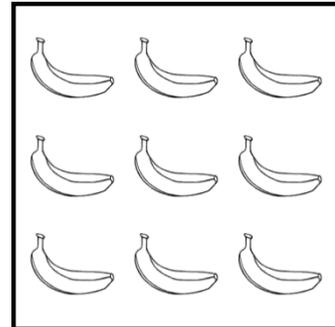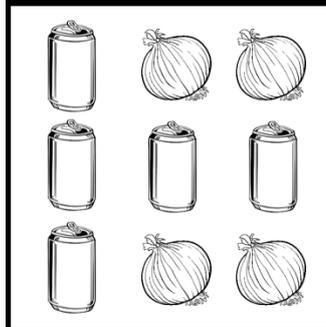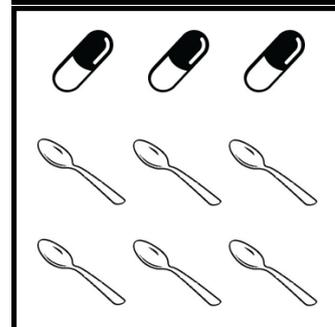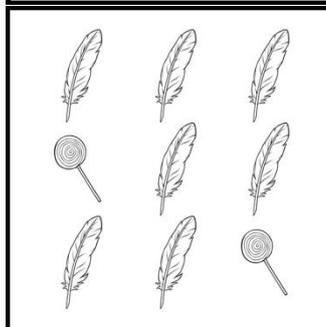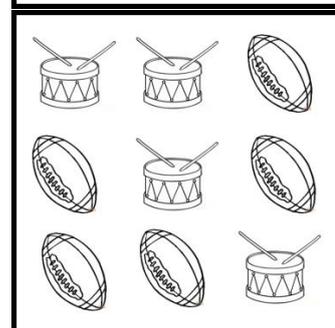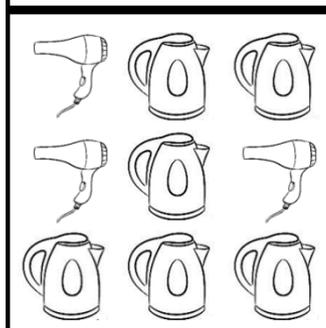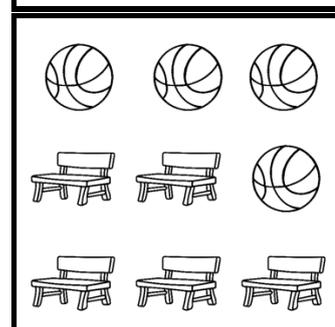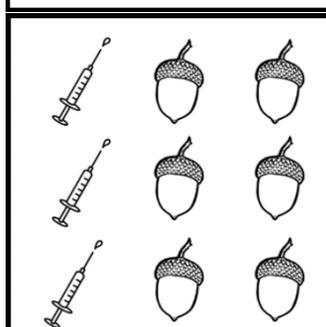
FIL31   On the winning card none of the objects are bottles
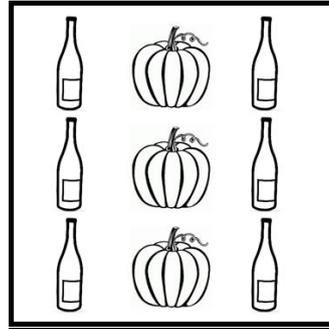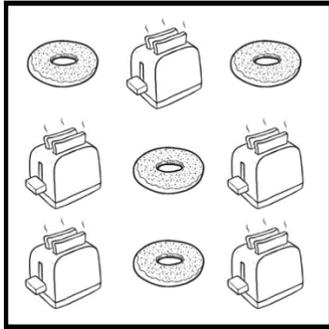
FIL32   On the winning card none of the objects are rings
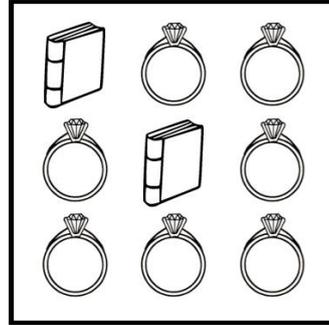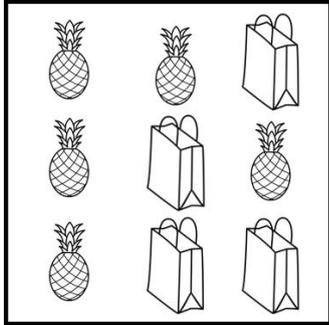
10.3 APPENDIX B1

1 COMPANY EXAM	
COMPETITIVE	COOPERATIVE
Imagine that you are about to sit a competitive exam for a particular position in your company. The exam has multiple choice and open answer questions. You don't remember whether you need to answer all the open answer questions in order to pass. You ask the person who has just sat the exam before you what she knows about the open questions. There is only one position opening and it's very well paid. Therefore you know that she probably hopes that you fail the exam so that she has a better chance of getting the job.	Imagine that you are about to sit an exam for a particular qualification in your company. The exam has multiple choice and open answer questions. You don't remember whether you need to answer all the open answer questions in order to pass. You ask the person who has just sat the exam before you what she knows about the open questions. The company has promised bonuses to all the employees if enough people pass the exam. Therefore she probably hopes that you both pass the exam.
• She says that 'Some of the open answer questions must be answered'. • Given what she told you, do you think it's possible that all of the open questions must be answered? • Do you think she meant that you don't need to answer all the open question? • Do you think she was trying to mislead you?	

2 TEAM RETREAT	
COMPETITIVE	COOPERATIVE
Imagine that you are the school's rugby team captain. You are applying to go on a team retreat funded by your school. You don't remember whether all of the players' parents need to sign the application for it to be considered. You ask the football team captain what he knows about it as he has already sent the application. The school will fund only one team to go on the retreat. Therefore you know that he probably wants your application to fail so	Imagine that you are the school's rugby team captain. You are applying to go on a team retreat funded by your school. You don't remember whether all of the players' parents need to sign the application for it to be considered. You ask the football team captain what he knows about it as he has already sent the application. The school is more likely to fund the retreat if many sports teams apply and go together. Therefore you know that he probably wants your application to

the football team have a better chance of going on the retreat.	succeed so both teams have a better chance of going on the retreat.

- He says that 'Some of the parents need to sign the application.'
- Given what he told you, do you think it's possible that all of the parents need to sign the application?
- Do you think he meant that you don't need to get all the parents to sign the application?
- Do you think he was trying to mislead you?

3 GAME SHOW	
COMPETITIVE	COOPERATIVE
Imagine that you are the captain of your team in a gameshow and in one of the challenges your team must sing a famous song. You don't remember whether all the team members need to sing the song for the team to complete the challenge. You ask someone in a different team what she knows about it. The gameshow has a big jackpot. Therefore you know that she probably wants your team to fail so her team has a better chance of winning.	Imagine that you are the captain of your team in a gameshow and in one of the challenges your team must sing a famous song. You don't remember whether all the team members need to sing the song for the team to complete the challenge. You ask a member of your team what she knows about it. The gameshow has a big jackpot. Therefore you know that she probably hopes that your team wins so you can all get the money.

- She says that 'Some of the members need to sing the song.'
- Given what she told you, do you think it's possible that all of the members need to sing the song?
- Do you think she meant that you don't need to get all the members to sing the song?
- Do you think she was trying to mislead you?

4 SCHOOL EDITOR	
COMPETITIVE	COOPERATIVE
Imagine that you are applying to become the editor of your school newspaper. You don't remember whether all your teachers need to approve your application for you to be considered. You ask another student who has already sent the application what he knows about it.	Imagine that you are applying to become an editor of your school newspaper. You don't remember whether all your teachers need to approve your application for you to be considered. You ask another student who has already sent the application what he knows about it.

There is only one editor position. Therefore you know that he probably wants you not to be considered so he has a better chance of becoming editor.	The more editors work at the newspaper the easier the job is. Therefore you know that he probably wants both of you to become editors.

- He says that 'Some of the teachers have to approve your application.'
- Given what he told you, do you think it's possible that all of the teachers have to approve your application?
- Do you think he meant that you don't need to get all the teachers to approve your application?
- Do you think he was trying to mislead you?

5 MUSIC PERFORMANCE	
COMPETITIVE	COOPERATIVE
Imagine that you are studying five piano pieces to perform for a music prize at your school. You don't remember whether at the performance you need to play all of the pieces from memory. You ask a student who has already performed about it. Only one student in the school can win the prize. Therefore you know that she probably hopes your performance will be bad so she has a better chance of winning the prize.	Imagine that you are studying five piano pieces to perform for a music qualification at your school. You don't remember whether at the performance you need to play all of the pieces from memory. You ask a student who has already performed about it. The school has promised to organise a school trip if enough students get the qualification. Therefore you know that she probably hopes that both your performances are successful.

- She says that 'Some of the pieces must be played from memory.'
- Given what she told you, do you think it's possible that all of the pieces must be played from memory?
- Do you think she meant that you don't need to play all the pieces from memory?
- Do you think she was trying to mislead you?

10.4 APPENDIX B2

Raw frequencies of responses segmented by answers to each question

		Meaning - yes			Meaning - no			
		Mislead yes	Mislead no	Tot.	Mislead yes	Mislead no	Tot.	Grand Total
COOP	Epistemic yes	19[a]	56[b]	75	13[c]	35[d]	48	123
	Epistemic no	9	75[e]	84	0	5	5	89
	tot coop	28	131	159	13	40	53	212
COMP	Epistemic yes	76[a]	50[b]	126	40[c]	14[d]	54	180
	Epistemic no	7	20[e]	27	1	5	6	33
	tot comp	83	70	153	41	19	60	213
	Grand Total	111	201	312	54	59	113	425

By crossing the answers to the different question participants loosely be classified as having interpreted the character in the story as being (a) deceptive through scalar implicature, (b) mistaken, (c) stonewalling/deceptive through ignorance implicature, (d) uninformative, (e) honest through scalar implicature.
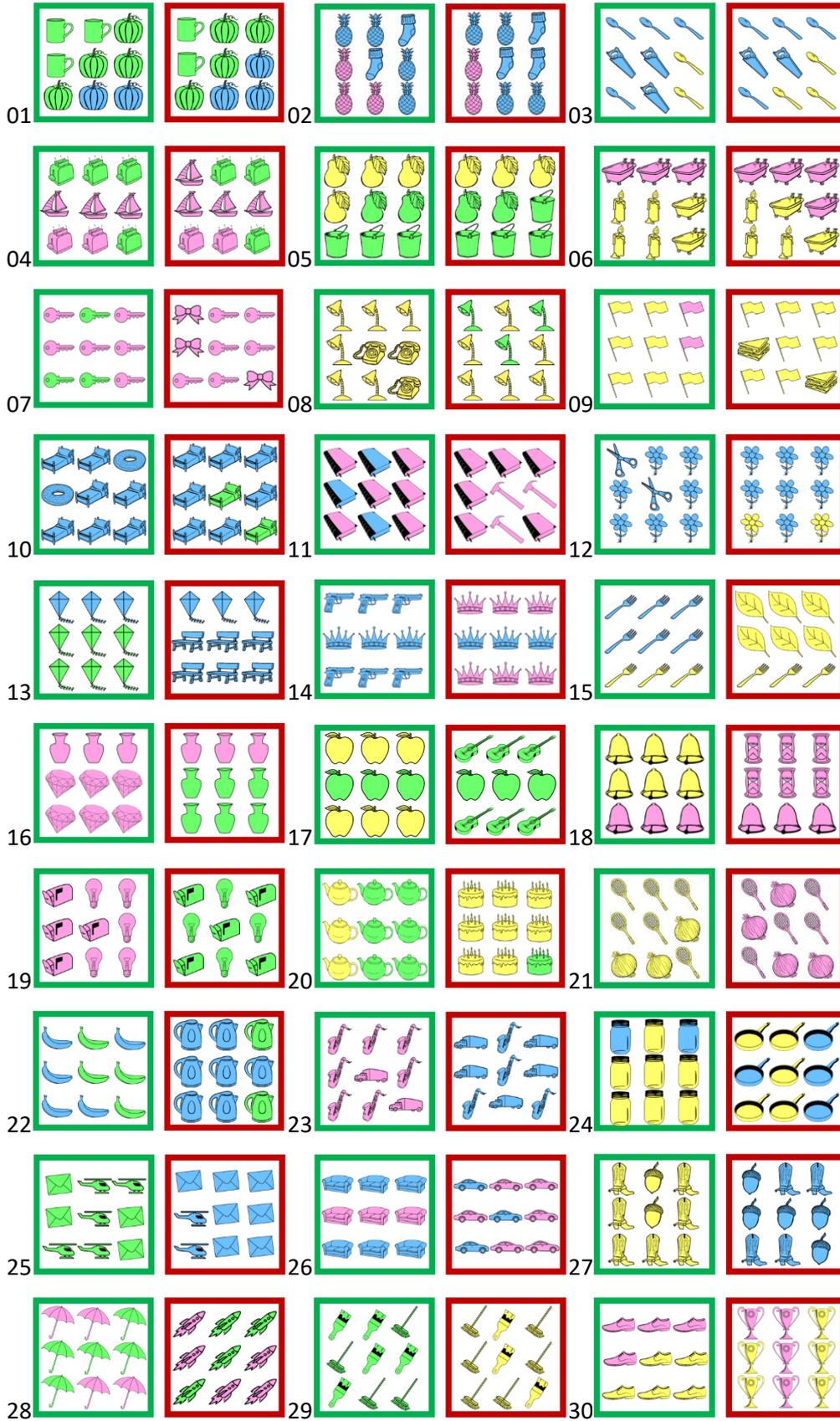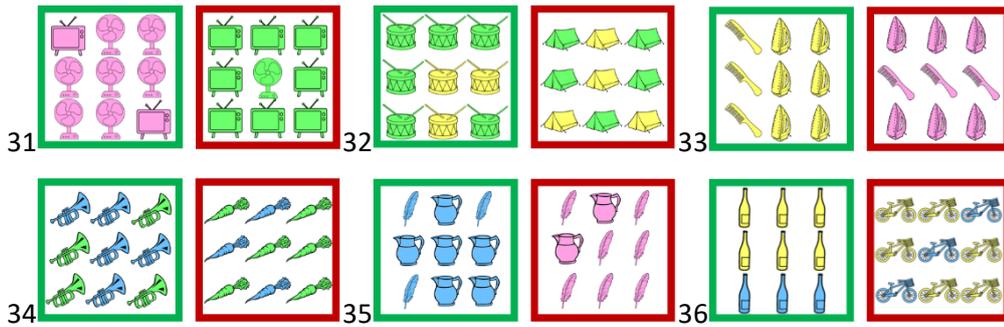
Item Type	Description: "On the winning card…"	True	False	Cards
Num	two of the objects are	blue	Mugs	01
Num	two of the objects are	socks	Pink	02
Num	two of the objects are	yellow	saws	03
Num	three of the objects are	Boats	green	04
Num	three of the objects are	Buckets	yellow	05
Num	three of the objects are	Pink	candles	06
Most	most of the objects are	Pink	key	07
Most	most of the objects are	lamps	yellow	08
Most	most of the objects are	yellow	flags	09
Most	most of the objects are	beds	blue	10
Most	most of the objects are	pink	books	11
Most	most of the objects are	flowers	blue	12
Ad hoc	the objects in the top row are	blue	kites	13
Ad hoc	the objects in the middle row are	crowns	blue	14
Ad hoc	the objects in the bottom row are	yellow	forks	15
Ad hoc	the objects in the top row are	vases	pink	16
Ad hoc	the objects in the middle row are	green	apple	17
Ad hoc	the objects in the bottom row are	pink	bells	18
None	none of the objects are	green	pink	19
None	none of the objects are	cakes	teapots	20
None	none of the objects are	pink	yellow	21
None	none of the objects are	kettles	bananas	22
None	none of the objects are	Blue	pink	23
None	none of the objects are	Pans	jars	24
None	none of the objects are	Blue	green	25
None	none of the objects are	cars	sofas	26
None	none of the objects are	Blue	yellow	27

All	all of the objects are	umbrellas	Rockets	28
All	all of the objects are	green	Yellow	29
All	all of the objects are	Shoes	trophies	30
All	all of the objects are	Pink	green	31
All	of the objects are	Drums	tents	32
All	all of the objects are	yellow	Pink	33
All	all of the objects are	Trumpets	Carrot	34
All	all of the objects are	blue	Pink	35
All	all of the objects are	bottles	Bikes	36

10.6 APPENDIX C2

Instructions for the competitive condition:

HOW THE GAME WORKS (please read carefully)
This is a competitive game with two players: a describer and a guesser.
In this game you are the describer. (the guesser will play in a second phase)
In each round of the game you'll see a winning card (with a green border) and a losing card (with a red border) and you'll have to complete a description of the winning card.
The guesser will read your description and they'll see both cards but they won't know which one is the winning card.
In this game the guesser scores points when they click on winning cards whereas you score points when the guesser clicks on losing cards.
The guesser knows that this is a competitive game but they don't know that you are completing the descriptions instead of writing them freely.
Please complete the instruction with ONLY ONE WORD. You can talk about a colour or a type of object.
You can write false descriptions and remember that you are playing against the guesser so your goal is to make them lose.

Instructions for the cooperative condition:

HOW THE GAME WORKS (please read carefully)
This is a cooperative game with two players: a describer and a guesser.
In this game you are the describer (the guesser will play in a second phase).
In each round of the game you'll see a winning card (with a green border) and a losing card (with a red border) and you'll have to complete a description of the winning card.
The guesser will read your description and they'll see both cards but they won't know which one is the winning card.
In this game both you and the guesser score points when the guesser clicks on winning cards.
The guesser knows that this is a cooperative game but they don't know that you are completing the descriptions instead of writing them freely.
Please complete the instruction with ONLY ONE WORD. You can talk about a colour or a type of object.
You can write false descriptions but remember that you're helping the guesser so your goal is to make them win.

Instructions comprehension questions [and answers] for both conditions:

Before you play, let's check that you know the rules:
(You won't be able to play if you get these wrong)
The winning card is...   [The red one / The green one]
Your descriptions can be... [Only true / Either true or false]
In this game you are... [Helping the guesser click on winning cards / Playing against the guesser]
You can write... [Maximum 1 word / Maximum 3 words]

10.7 APPENDIX D

Full set of instructions for Experiment 1 (1 spider 1 empty scenario):

You will see the image of a fake spider. If you are afraid of spiders and you think you will find it distressing do not take part in this study.

This experiment will take approximately 1 minute.

\

David is an honest guy.

When it's time for you to choose a box, David will give you a piece of information.

David can't talk but he can point.

[**Figure 10** The image of David shown to participants]

\

INSTRUCTIONS (please read carefully):

In this task you will see 2 boxes and you will have to choose one.

As you can see below:

1 box is empty

1 box contains a spider

If you choose the empty box you win

If you choose the box containing a spider you lose

[open box view in ***Table 12***]

\

Now we shuffled the boxes.

David knows what's in each box and on the next page he'll give you a piece of information.

[picture of the boxes with the lid closed]

\

Do you remember how many of these boxes contain a spider?

(multiple choice: 1/2/3/4/5/6/7/8/9/10)

\

Please choose a box (A/B)

[Clue view in ***Table 12***]