

Interpreting referential noun phrases in belief reports – the *de re/de dicto* competition

Yuhan Zhang¹ and Kathryn Davidson¹

¹Department of Linguistics, Harvard University

yuz551@g.harvard.edu

Abstract

The *de re/de dicto* ambiguity centers on the referential and/or attributive properties of noun phrases in the scope of intentional operators such as belief reports. For the belief report *Julie believes Elizabeth's poem will win the competition*, a *de re* reading of the embedded referential noun phrase *Elizabeth's poem* entails that the referential association between this noun phrase and the target poem is true from the perspective of the speaker but may not be registered as such in the belief holder's (e.g. Julie's) mind. In contrast, a *de dicto* reading describes Julie's beliefs as she registers them in her mind. While both the *de re* and *de dicto* readings of definite noun phrases like *Elizabeth's poem* are generally available given supporting contexts, we show that the *de re* acceptability is vulnerable to contextual and pragmatic manipulations to the extent of obtaining the opposite truth value, even when the rest of the context permits a *de re* reading. Specifically, when the *de re* permitting context also allowed a *de dicto* reading of a competing noun phrase that refers to the same entity, the speaker-oriented *de re* reading of the target noun phrase achieved bimodal acceptability and participants preferred the *de dicto* noun phrase that oriented toward the belief holder's mind. Our study is the first that systematically lays out the empirical landscape of *de re/de dicto* readings of definite noun phrases, which solidifies the foun-

dation for formal theory development and endorses the practice of collecting reliable empirical judgment data for nuanced semantic phenomena.

Key words: *de re/de dicto*, definite noun phrases, belief reports, linguistic judgment, experimental semantics, context effects

1 Introduction

The *de re/de dicto* distinction refers to an interpretive ambiguity of noun phrases embedded in an intensional domain¹. For example, in (1), under a *de re* reading of the noun phrase *a prince*, Aurora wants to marry a particular individual who the speaker of (1) knows to be a prince. This is in fact the scenario in the first part of the story of *Sleeping Beauty*, in which Aurora falls in love with a man she meets in a forest who the narrator knows to be Prince Phillip, although Aurora herself is not aware of his royal station. In this scenario, someone can truthfully describe Aurora’s desires with (1). However, this sentence also has another interpretation – the *de dicto* interpretation of the noun phrase *a prince*, in which Aurora’s beliefs are characterized as, basically, wanting to marry a prince, whoever one might be. In the context of *Sleeping Beauty*, the same sentence under the *de dicto* interpretation is false because Aurora desires to marry only the man she met in the forest, thus she is not in the state of desiring to marry a prince *de dicto* (perhaps contrary to a common fairytale/royalty expectation for princesses).

- (1) Aurora wants to marry a prince.
 - a. True under the *de re* reading in the context of the *Sleeping Beauty* story
 - b. False under the *de re* reading if Aurora did not want to marry Prince Phillip (and Prince Phillip is the only prince in the context)

¹As early as Aristotle, linguistic phenomena related to *de re/de dicto* have been observed. Yet this pair of Latin terms was not intensively applied until the Medieval period by Thomas Aquinas. The adoption of the terms in philosophy and linguistics was initiated by G. Frege (1892), B. Russell (1905), and W.V.O. Quine (1956) even if none of them explicitly use the term *de re/de dicto* in their writings. The current senses of *de re* and *de dicto* may deviate a little from the literal Latin meaning of the terminology (*de re*: “of the thing”, *de dicto*: “of what is said”) (von Fintel & Heim, 2011), so it may be clearer to introduce the *de re/de dicto* distinction via contextualized examples. For more details of the nomenclature, see Keshet and Schwarz (2019).

- c. True under the *de dicto* reading if Aurora were following expectations for royalty
- d. False under the *de dicto* reading in the context of the *Sleeping Beauty* story

The classic approach in formal semantics to model the *de re/de dicto* distinction is scope ambiguity (Partee, Meulen, & Wall, 1990; Quine, 1956; von Fintel & Heim, 2011, a.o.). As represented in (2a) for the *de re* reading of (1), the property of being a prince that holds of the bound variable x is evaluated to be true in the actual world w_0 ; Aurora marries x takes place in the possible world w' that is compatible with what Aurora wants in the actual world. On the other hand, for the *de dicto* reading in (2b), the “prince” property is evaluated to be true in Aurora’s desire world w' .

- (2) a. $\exists x [\text{prince}_{w_0}(x) \wedge \forall w' [\text{WANT}_{w_0}(A, w') \rightarrow \text{marry}_{w'}(A, x)]]$ (*de re*)
 b. $\forall w' [\text{WANT}_{w_0}(A, w') \rightarrow \exists x [\text{prince}_{w'}(x) \wedge \text{marry}_{w'}(A, x)]]$ (*de dicto*)

The scope approach generates different implications for both the *de re* and *de dicto* interpretations which sharpen their contrast. For the *de re* reading in (2a), since the existential quantifier takes a wider scope than the universal quantifier over possible worlds and *prince* is evaluated outside the universal quantifier, Aurora doesn’t have to realize that x is a prince. The noun phrase *a prince* is just one of many possible referential terms that pick up the specific individual in the real world. For the *de dicto* reading in (2b), on the other hand, since *prince* is evaluated within the scope of *want*, Aurora’s beliefs are that whoever she wants to marry must be a prince. Under the *de dicto* reading, given that Aurora’s desires involve her marrying a prince, then if we know that Aurora wants to marry someone, we also know they must be a prince.

The *de re/de dicto* ambiguity has also been extensively studied in definite noun phrases and other DPs (Nelson, 2019; Percus, 2000; von Fintel & Heim, 2011, a.o.). An example is (3) where the possessive noun phrase *your abstract* could have either a *de re* or a *de dicto* reading given the corresponding supporting context. Specifically, the ambiguity hinges on whether the belief holder is aware of the possessive relation between the abstract and the addressee. The ambiguity could also be predicted by the scope analysis (Fodor, 1970; Romoli & Sudo, 2009, a.o.).

- (3) John believes that your abstract will be accepted.

Evaluating context for De Re: John’s belief may be about an abstract that he reviewed, but since the abstract is anonymous, he doesn’t know who wrote it. He tells me about that abstract and he believes that it is sure to be accepted. I know that it is your abstract and inform you of John’s opinion by saying the sentence above.

Evaluating context for De Dicto: You are a famous linguist that John is acquainted with. John knows your work very well and believes that you submitted a (unique) abstract to a linguistic conference. Given his general knowledge about this specific conference and his high opinion of your work, he believes that your abstract will be accepted, even if he doesn’t know which one is your abstract or has read it. He tells me his belief and I am retelling the belief to you.

(von Fintel & Heim, 2011, p.157).

Following the scope solution for this basic ambiguity under simple contexts, more theoretical semantic tools have been proposed to capture the *de re/de dicto* ambiguity in syntactically different complement clauses, for different types of noun phrases, and under more sophisticated contexts (Charlow & Sharvit, 2014; Deal, 2018; Elliott, 2023; Keshet, 2008; Percus, 2000; Percus & Sauerland, 2003, a.o). Diverging a little from yet ultimately contributing to this path, in this paper, we utilize tools in experimental semantics (e.g. Cummins & Katsos, 2019) and investigate the contextual influence on the acceptability of *de re/de dicto* readings of definite noun phrases. The motivation behind going “experimental” is that despite the undoubted availability of *de re* readings of definite noun phrases in examples like (3), we still found literature that shows the acceptability of *de re* readings could be prioritized or deprioritized from context to context even if those contexts theoretically fulfill the requirements that make a *de re* reading true (J. C. Anderson, 2013; Jaszczolt, 1997; Sudo, 2014; Zhang & Davidson, 2021). Given the field’s increasing interest in providing robust replicable linguistic evidence to support theoretical development (Davidson, 2020; Tonhauser & Matthewson, 2015, a.o.) as well as the limited experimental research on the topic of *de re/de dicto*, we believe it is crucial to understand under what contexts *de re* readings would be more acceptable than others. We

hope this line of research could further solidify the empirical foundation of referential properties of noun phrases in intensional semantics.

The paper is structured as follows. In Section 1.1, we present claims of an asymmetry in the acceptability of the *de re* reading compared to the *de dicto* reading for definite noun phrases, among broader observations of both interpretations generally being available for noun phrases in the scope of intensional operators. In Section 1.2, we reinforce the motivation for experimental investigation and highlight the potential contribution of this study to linguistic theory. In Section 1.3, we lay out a finer-grained categorization of the *de re* permitting contexts and of the *de re/de dicto* ambiguity. We use these new categories to design well-controlled stimuli for empirical testing and causal analysis. In Section 1.4, we introduce the experiment outline. Then in Section 2 to 4, we report designs and results of three experiments. In Section 5, we raise potential explanations to the contextual effect on *de re* acceptability and conclude with suggestions for future work.

1.1 Diverging judgments of *de re* readings

In this section, we present existing literature suggesting that the *de re* reading of noun phrases embedded in intensional domains may be less acceptable or preferred in some contexts than a *de dicto* interpretation.

First, we see claims that argue for a preference toward the *de re* reading out of the *de re/de dicto* ambiguity from works in Default Semantics (Capone, 2011; Jaszczolt, 1999, 2005, 2015, a.o.). In this framework, where a main claim is that utterance meaning is jointly determined by its compositionality, the intention of interlocutors, and their cognitive biases in communication, Jaszczolt (1997) argues that the *de re* reading of definite noun phrases is the default and the most salient one out of the ambiguity because the primary objective of communication is understanding the speaker's intention by securing the referent of the speaker's utterance in the conversational context. Since the *de re* reading highlights the referential property of noun phrases and is able to select objects in the broader conversation context, it should stand out during interpretation. This *default de re* theory not only predicts the availability of *de re* interpretations but also predicts them to be even more accessible

relative to *de dicto* during communication.

Second, in legal studies, [J. C. Anderson \(2013\)](#) reported that the *de re/de dicto* ambiguity has been overlooked in the interpretation of legal statutes – the emphasis on *de re* readings of statutes and the lack of attention on the *de dicto* one have even led to puzzling judicial results². For example, in the famous Enron Scandal in 2001, Enron’s auditor Arthur Anderson anticipated litigation and urged its employees to destroy related financial documents. Their action stopped on the day when the U.S. Securities and Exchange Commission subpoenaed records. At that time, two federal obstruction statutes applied to document destruction. The more general statute makes it a federal offense to “corruptly...endeavor to influence, obstruct, or impede the due administration of justice.”³ The more specific one prohibits “knowingly...corruptly persuad[ing] another person...with the intent to...destroy an object...[or]...impair the object’s...availability for use in an official proceeding.”⁴ The intentional words in these two statutes *endeavor* and *intent* grant each of the statutes a *de re* and a *de dicto* interpretation, as shown in (4) and (5).

(4) *De re* interpretations

(i) For the general statute: There is some X, which is in fact an instance of justice being administered. The defendant endeavors to influence or obstruct X.

(ii) For the specific statute: There is a specific official proceeding Y in which the defendant intends to impair some objects’ availability for use.

(5) *De dicto* interpretations

(i) For the general statute: The defendant has the corrupt intention to influence what we describe as “the administration of justice” ([J. C. Anderson, 2013](#), p.28).

(ii) For the specific statute: The defendant has the intention to initiate some kind of impairment against any possible official proceeding.

Because there was an overwhelming reliance among the judges on the *de re* interpretation of the two statutes and there was no sufficient evidence that suggests Arthur Anderson’s destruction

²For more legal cases where the bias towards *de re* reading influenced judicial results, see [J. C. Anderson \(2013\)](#).

³18 U.S.C. §1503 (2010)

⁴18 U.S.C. §1512(b) (2010)

was under the knowledge of a *specific* pending proceeding, the defendant was not charged under the general clause. What is worth noting is that the ruling could have been different had the *de dicto* reading been picked up. Anderson resorts to the difficulty for children to acquire the Theory of Mind ability (Apperly et al., 2010; Baron-Cohen, Leslie, & Frith, 1985; Wellman, 1992, a.o.) to back up her observation on the preference over *de re*. That is, the observation that children tend to focus on reality and the broader context and find it difficult to reason about others' mind helps explain why (even) adults have a preference toward *de re* that emphasizes what is going on from the actual world and from the global context compared with *de dicto* that emphasizes mental status. This study in law cites different sources from the Default Semantics framework but arrives at similar conclusions on the empirical acceptability of the *de re/de dicto* ambiguity.

Finally, the evidence for a dispreference for *de re* comes from experimental findings reported in Zhang and Davidson (2021). They designed an acceptability task as exemplified in (6). The evaluating context featured a protagonist Julie who falsely associated the authorship of the target poem with Nicole but in reality and from the speaker's perspective, the poem was written by Elizabeth. According to a similar scenario in Romoli and Sudo (2009)⁵, the belief report with the noun phrase *Nicole's poem* should be interpreted *de dicto* and the one with *Elizabeth's poem* should be *de re*. In Zhang and Davidson (2021)'s experiment, 115 native speakers of English were given one of the two sentences (6a, 6b) with the same context, and the instruction was "According to this story, please indicate to what extent you agree or disagree with the following statements". The result showed that 18.3% of participants and a significant proportion of trials highly disagreed with the *de re* reading even if more than 50% of the *de re* trials showed high agreement. In contrast, few trials showed disagreement with the *de dicto* reading. This empirical finding, especially the peculiar bimodal dis-

⁵The example in Romoli and Sudo (2009) is shown in (1). The context explicitly shows that the belief holder does not think the *de re* noun phrase refers to the target individual because the belief holder registers the *de dicto* term as associated with the target individual.

- (1) John thinks that the president of the United States is smart.
Evaluating context: Consider the situation as of today [2009], in which Barack Obama is the president of the United States, and suppose that John wrongly thinks that Al Gore is. In this context, the sentence has two interpretations. It can be read as reporting John's belief about Barack Obama or about Al Gore. The former is called the *de re* reading and the latter the *de dicto* reading.

tribution of *de re* acceptability, is surprising in light of other work such as [Romoli and Sudo \(2009\)](#) where both readings are predicted to be generally acceptable.

- (6) *Evaluating Context*: Julie is one of the judges of an ongoing poetry competition. The best poem that she has read so far is an extremely intriguing poem about the ocean. She believes that this poem will win the competition. Julie remembers being told that Nicole, one of the best-known poets, submitted a poem about the ocean to the competition. Therefore, Julie concludes that this poem must be written by Nicole and the first prize will be going to her. However, this poem was actually written by Elizabeth, a younger and lesser-known poet. It is just a coincidence that the two poets wrote about the same topic.
- a. Julie believes that Nicole's poem will win the competition. (*de dicto*)
 - b. Julie believes that Elizabeth's poem will win the competition. (*de re*)

Thus we see several varied sources in the literature that seem to report more easiness or difficulty accessing *de re* readings in comparison to *de dicto* readings of noun phrases. This seems like an area ripe for more careful experimental consideration, especially given the role that context may play in these judgments.

1.2 More motivations for experimental investigation

Given some reported disparities in existing literature between *de re* and *de dicto* interpretations, more carefully controlled experimental research seems prudent; in this section we further motivate an experimental turn in this domain.

For one thing, from the perspective of research methodology, the experimental investigation on *de re/de dicto* provides an instance that enriches the discussion on “the nature of empirical evidence in research on meaning” ([Tonhauser & Matthewson, 2015](#), p.1). To yield stable replicable and transparent data for theoretical development, [Tonhauser and Matthewson \(2015\)](#) argues one needs to provide not only the linguistic expression but also the context in which the expression is uttered, a response by a native speaker to a task involving the linguistic expression in the context, and information about

the native speakers that provide the response. [Gibson and Fedorenko \(2010\)](#) and [Davidson \(2020\)](#) also argue for using multiple items with controlled experiments to eliminate confounding factors like specific lexical properties in a single sentence and idiosyncrasies from the contexts. Existing research in semantics and pragmatics has found the influence of contexts, experimental paradigm, and/or response options on linguistic judgments and therefore corroborated the necessity to become more “experimental” (see, e.g. [Schwarz, Clifton Jr, and Frazier \(2007\)](#) and [Jasbi, Waldon, and Degen \(2019\)](#) for scalar implicature and the influence of contexts as well as response options; see [Jasbi, Bermudez, and Davidson \(2023\)](#) and [Jasbi, Bermudez, Zhang, Siro, and Davidson \(2023\)](#) for logical connectives and the effect of experimental paradigm on cross-linguistic findings). Our research is another case study that advocates for rigorous experimental practice in the study of meaning.

Second, empirically investigating the *de re/de dicto* reading acceptability provides a case study of how formal theories of meaning that involve logic and mathematical techniques align with psycholinguistic findings that delineate the psychological representation of meaning. While we might be familiar with some well-known phenomena like scalar implicature under this topic ([Chemla & Singh, 2014a, 2014b](#)), this *de re/de dicto* issue could open up more perspectives. For example, studying the empirical acceptability of the *de re/de dicto* ambiguity against a semantic representation utilizing the formal technique of scope interaction could be analogous to research that explores the relation between the formal quantificational scope and related psycholinguistic processing mechanisms in classic scope interactions (see [Brasoveanu & Dotlačil, 2019](#), for a review). While we know that the inverse scope reading is sometimes hard to obtain, as in (7) (e.g. [C. Anderson, 2004](#); [Tunstall, 1998](#)), it would be interesting to compare the scope interactions and the corresponding reading behavior between the case of multiple quantifiers and the *de re/de dicto* ambiguity.

- (7) a. A caregiver comforted Mary every night. (The inverse scope reading is, at every night, there was a different caregiver who comforted Mary.)
- b. A boy climbed every tree. (The inverse scope reading is, for every tree, there is a different boy who climbed it.)

Last but not least, the *de re/de dicto* ambiguity is one of many phenomena that describes how

language expresses thought and, in turn, how thought can be accessed via language (see Nelson, 2023, for a review). It attracts interdisciplinary discussion from philosophy (e.g. Cohen, Tang, & Wang, 2021; Epstein, Naumov, & Tao, 2023; Lederman, 2022; Richard, 1990), cognitive science (e.g. Apperly & Robinson, 2003; Robinson & Apperly, 2001), and even artificial intelligence (e.g. Wiebe & Rapaport, 1986). Understanding what contexts facilitate or impede a certain reading of the belief report could shed light on the research agenda that understands language and mind. More practically, this line of research could also complement existing research about false belief tasks and Theory of Mind (e.g. Baron-Cohen et al., 1985; Wimmer & Perner, 1983) where the focus has been on the predicate of the belief content (e.g. *Sally believes that the marble is in the box.*) rather than the referential properties of noun phrases (e.g. *the marble*) inside belief reports.

1.3 A finer-grained categorization of *de re* contexts and the *de re/de dicto* term

In this section, we set up a finer-grained categorization of (i) the *de re* permitting contexts and (ii) the *de re/de dicto* terminology to better operationalize our experimental investigation. The aim is to collect accurate linguistic judgments for a specific linguistic expression with a clarified meaning in a well-controlled context.

First, we resort to Sudo (2014) and categorize the *de re* permitting contexts into two types: the “ignorance” context and the “misapprehension” context. The “ignorance” context applies to cases where the belief holder is ignorant of the association between the *de re* noun phrase and the referred object. The context in the “abstract reviewing” example in (3) that permits the *de re* reading can be categorized as the “ignorance” context since the belief holder John is not aware that the abstract was written by the addressee. On the other hand, the “misapprehension” context applies to cases where the belief holder is wrong about the identity of the referent and associates the wrong noun phrase to the referent. The *de re* context in the “poetry competition” example in (6) is the “misapprehension” context since the judge Julie believes the poem was written by Nicole but in reality and in the broader story context the poem was written by Elizabeth (the latter noun phrase *Elizabeth’s poem* is read *de re*).

In Sudo (2014), there is not a discussion related to the influence of the two contexts on the acceptability of the *de re* reading. Both contexts are assumed to permit *de re*. In Zhang and Davidson (2021), the judgment data shows the “misapprehension context” disfavors *de re*. One potential hypothesis is that the “misapprehension” context highlights the contrast between the belief holder’s mental state and the story’s broader context and thus prioritizes the *de dicto* reading (e.g. *Julie believes that Nicole’s poem_{de dicto} will win the competition*) where the use of the *de dicto* term is consistent with the matrix clause introducing the belief holder’s mental state. Previous research on children’s interpretation of belief reports shows that highlighting the belief state of the protagonist in the context increased the likelihood that children attend to the belief content (Lewis, Hacquard, & Lidz, 2017), it is interesting to see whether this treatment can be borne out in *de re/de dicto* research.

Secondly, we adopt a three-way distinction of the *de re/de dicto* ambiguity as referential *de re*, referential *de dicto* and attributive *de dicto*. This is motivated not only by the fact that ontologically two types of contexts emerge under the umbrella of *de dicto* for definite noun phrases but also that a stricter control of *de dicto* can provide a clearer reference for the compared *de re* reading. Inspired by Domellan (1966), Fodor (1970), and Jaszczolt (1997) on their discussion around the referential and attributive properties of definite noun phrases, we instantiate this tripartite taxonomy in the “poetry competition” example, repeated here as (8)⁶.

- (8) Julie believes that Elizabeth’s poem will win the competition.

The referential de re context: Julie does have a particular poem in mind that she believes will win the competition. However, Julie doesn’t recognize the description *Elizabeth’s poem* as a description of the poem she has in mind.

The referential de dicto context: Julie does have a particular poem in mind that she believes will win the competition. She has the poem in mind as *Elizabeth’s poem*. However, in reality, the poem is Nicole’s poem. In other words, Julie falsely represents the poem by using a referential term that is false in the actual world.

⁶Note that there is a fourth reading here: Julie believes that poem A will win the competition and knows that poem A was written by Elizabeth. This reading is not controversial and less interesting from a theoretical sense so we didn’t put it in parallel with the other three.

The attributive de dicto context: Julie does not have any particular poem in mind but simply believes that whichever poem written by Elizabeth will win the competition.

By integrating the two context types and the three readings of intensional definite noun phrases, we learn that the “ignorance” context only permits a referential *de re* reading of the corresponding belief report and the “misapprehension” context (theoretically) permits both a referential *de re* and a referential *de dicto* reading of corresponding belief reports. The following experiments aim to test whether the division of context types and terminologies can help address the bimodal judgment of *de re* in Zhang and Davidson (2021). Since this study is one of the starting points for experimental investigations on *de re/de dicto* ambiguity, we mainly focus on addressing the bimodal pattern of *de re*.

1.4 Experiment outline

This section introduces the outline of three experiments. All three experiments adopted the truth-value judgment task with adult native speakers of English.

Experiment 1 replicated Zhang and Davidson (2021) and found that while the context featured misapprehension and allowed a referential *de re* reading for one noun phrase and a referential *de dicto* reading for another, the referential *de re* reading did receive bimodal judgments.

Experiment 2 put the critical sentence from Exp.1 under contexts that supported an attributive *de dicto* reading (as the control context) and contexts that only supported a referential *de re* reading (as the ignorance context). There, we found no statistical difference in acceptability ratings between both contexts – both received high agreement. This shows that (referential) *de re* interpretations of belief reports are acceptable in the ignorance contexts.

Experiment 3 tested the hypothesis that the misapprehension context, i.e., the co-existence of a referential *de re* definite noun phrase and a referential *de dicto* definite noun phrase in the same context, would make the former reading less acceptable. By juxtaposing the misapprehension context from Experiment 1 and the ignorance context from Experiment 2, we found evidence supporting this hypothesis.

In sum, we show that while the *de re* reading of definite noun phrases is undoubtedly allowed by the grammar, its acceptability is vulnerable to contextual factors. When a competing referential *de dicto* appears in the same context, in other words, when the belief holder associates a wrong term with the target object, the *de re* reading with a different term becomes much less acceptable. We discuss the potential mechanisms and implications in Section 5.

2 Experiment One

Experiment 1 aims to replicate [Zhang and Davidson \(2021\)](#). It asked whether the bimodal distribution of *de re* judgments would persist with another round of testing.

2.1 Methods

2.1.1 Participants

60 participants who self-identified as English monolinguals from the United States (aged 39.15 ± 12.05) took this study. They were recruited from the online crowdsourcing platform [Amazon’s Mechanical Turk](#) and were paid \$2.00 for their participation. All experiments were conducted with the approval of our university’s Institutional Review Board.

2.1.2 Materials & Procedures

Since Experiment 1 was a replication of [Zhang and Davidson \(2021\)](#), the experimental materials were the same as theirs. The only difference was that we adopted a discrete fully-labeled Likert scale to collect the judgment rather than a continuous slide bar in the original study. We chose a Likert scale over a binary option or a continuous slider based on the following considerations: First, Likert scales provide finer-grained levels to reveal potential judgments that would otherwise remain concealed on a binary scale (e.g. [Jasbi et al., 2019](#); [Katsos & Bishop, 2011](#); [Zhang, Ricciardi, & Davidson, 2021](#)); second, the labels at the intermediate levels (e.g. “somewhat agree”, “uncertain”, “somewhat disagree”) offer better interpretability than the continuous slider bar when it comes to

mapping participants' intermediate choices with their actual interpretations; third, choosing Likert scale does not lose sensitivity compared with the continuous slider bar (Marty, Chemla, & Sprouse, 2020; Sprouse & Almeida, 2017).

As for the specific experimental design, participants read four stories (113 ± 6.4 words) in a Qualtrics survey and gave their acceptability judgments on four declarative sentences in terms of how accurately each sentence reflected the facts in each story. The story and the sentences appeared on separate pages. There was no time pressure to complete the experiment. Participants could return to any story and change their answers at any time before submitting their answers.

Each story portrayed a protagonist who holds a belief toward a person or object (we simplify this using "entity"). The story fosters a "mistaken identity context", i.e., a "misapprehension context" where the protagonist falsely attributes one definite noun phrase to refer to the target entity but in reality, the correct attribution should come from the other definite noun phrase. One of the trials is shown in Table 1⁷ which is the same as (6). Julie falsely believes the poem was written by Nicole. In reality, it was written by Elizabeth, which Julie is unaware of. Given this context, the target sentence to be judged featured a report describing the protagonist's belief. The experimental manipulation concerned what definite noun phrase to be used in the belief report as the referring expression. Using the definite noun phrase held true in the protagonist's mind would render the belief report an unambiguous referential *de dicto* reading. In contrast, using the definite noun phrase held true in the broader story context and from the speaker's perspective would render the report a referential *de re* reading. In theory (Romoli & Sudo, 2009; Sudo, 2014), both readings are predicted to be true.

Additionally, for each story, there were three sentences accompanying the target sentence as fillers and controls; of these one was true given the context, one was false, and the third one was undecided because of the lack of verifying information. For each sentence to rate, participants were asked to map their judgment onto a five-point fully labeled Likert scale where each point had a label and represented the categorical level of agreement. Participants gave their judgment depending on whether and to what degree they agreed that the sentence correctly reflected the story content.

⁷The critical sentences that determined the condition were italicized in the table for illustration purposes. They were not italicized in the actual experiment.

<p>Context</p> <p>Julie is one of the judges of an ongoing poetry competition. The best poem that she has read so far is an extremely intriguing poem about the ocean. She believes that this poem will win the competition. Julie remembers being told that Nicole, one of the best-known poets, submitted a poem about the ocean to the competition. Therefore, <i>Julie concludes that this poem must be written by Nicole and the first prize will be going to her. However, this poem was actually written by Elizabeth, a younger and lesser-known poet.</i> It is just a coincidence that the two poets wrote about the same topic.</p>										
<p>Instruction</p> <p>According to this story, please indicate to what extent you agree or disagree with the following statements.</p>										
<p>Target sentence 1</p> <p>Julie believes that Nicole’s poem will win the competition. (Referential <i>de dicto</i>)</p> <p>Target sentence 2</p> <p>Julie believes that Elizabeth’s poem will win the competition. (Referential <i>de re</i>)</p> <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr> <td style="width: 20%;">Highly Disagree</td> <td style="width: 20%;">Somewhat Disagree</td> <td style="width: 20%;">Uncertain</td> <td style="width: 20%;">Somewhat Agree</td> <td style="width: 20%;">Highly Agree</td> </tr> <tr> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> </table>	Highly Disagree	Somewhat Disagree	Uncertain	Somewhat Agree	Highly Agree	<input type="radio"/>				
Highly Disagree	Somewhat Disagree	Uncertain	Somewhat Agree	Highly Agree						
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>						

Table 1: Example story from Experiment One

By comparing the proportion of different levels of agreement between conditions, we expected to approximate the representative judgment distribution of the two belief report versions.

Altogether each participant read four stories. We chose four items because this number was chosen in [Zhang and Davidson \(2021\)](#) and it was easier to compare item difference with a smaller set of scenarios at the initial stage of the *de re/de dicto* investigation agenda.

Furthermore, the condition manipulation was within-subjects and participants read stories in both conditions. Across the four stories, two were randomly assigned to be referential *de dicto* and the other two were referential *de re*. The order of stories and the sentences within a story were randomized. To achieve a Latin Square design, we manually created six lists⁸ and the participants were randomly assigned to each list. During the actual experiment, participants started by completing three practice trials (sentences to be judged without contexts) to familiarize themselves with the experiment design. In the end, participants completed a survey and provided their demographic in-

⁸We created six lists because there are six combinations where two stories are randomly interpreted *de re* and the other two are *de dicto*. That is {AB/CD, AC/BD, AD/BC, BC/AD, BD/AC, CD/AB} where the first two stories corresponded to a referential *de dicto* condition and the latter two corresponded to a referential *de re* condition.

formation and their self-reported linguistic profile (e.g. reading and writing proficiency, knowledge of other dialects/languages).

2.2 Results

We only analyzed the judgments from participants who passed the practice trials and whose judgments of the fillers were correct more than 75% of the time, 51 out of 60 participants (85%) contributed their data to our final analysis⁹.

Figure 1 shows that in the referential *de dicto* condition, the majority of judgment goes to the “highly agree” side. The sharp contrast is in the referential *de re* condition where around 25 % of the judgments are “highly disagree” with the majority of choices still going to “highly agree”.

To assess the condition differences statistically, we fit the judgment data into Bayesian multilevel cumulative ordinal models using the *brms* package (Bürkner, 2017, 2018) in R. The five-point judgments were the dependent variable with non-equidistant intervals between levels on the Likert scale. They were coded from 1 to 5 where 1 indicated “highly disagree” and 5 indicated “highly agree”. The two critical condition levels were entered as a dummy-coded fixed effect (reference level = referential *de dicto*). The story was also entered as a fixed effect, and the interaction between the two fixed effects was also included¹⁰. Random intercepts and slopes for the full fixed effect structure for the subjects were included as random effects to obtain the maximal random effect requirements for mixed-effects models (Barr, Levy, Scheepers, & Tily, 2013). The prior distributions for all the intercepts and coefficients of fixed effects were fitted to a normal distribution with the mean as 0 and the standard deviation as 2 (i.e., $Normal(0, 2)$); the prior for the correlation matrices was set to be $LKJ(2)$ (LKJ has been the default weakly informative prior for correlation matrices in *brms* (Lewandowski, Kurowicka, & Joe, 2009; Nalborczyk, Batailler, Lœvenbruck, Vilain, & Bürkner, 2019)); the variances for the correlation matrices were set as the default in R. These setup of priors

⁹The reason why we included the practice trials in the participant screening procedure was that we explicitly asked the participants to choose, e.g. “highly agree”, over other choices. A failure to do so indicated a lack of attention to our materials and instructions.

¹⁰We consulted a statistician from our funding institution and took his suggestion to treat the story (N=4) as a fixed effect. Another motivation was that we could derive more insights into story-specific effect on the acceptability of *de re* reading from the statistical analysis.

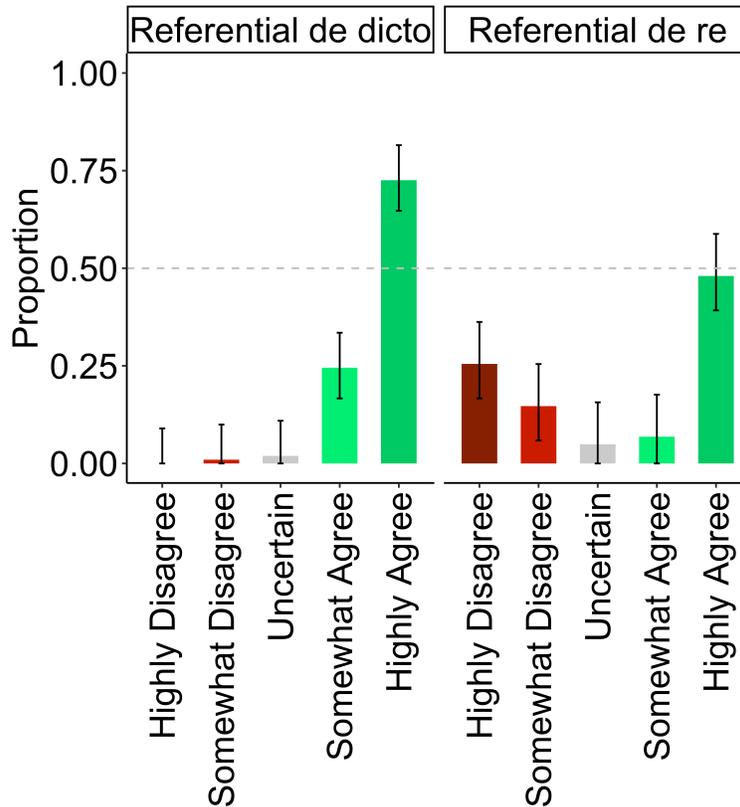


Figure 1: Proportion of different Likert scale choice in the referential *de dicto* and the referential *de re* condition (The 95% CIs were output from the *MultinomialCI* package (Sison & Glaz, 1995)).

mildly restrict the possible coefficient for each parameter but still allow reasonably large variance. The model had four sampling chains each with 4000 iterations. The first 2000 samples were taken as a warmup. An \hat{R} close to 1.0 marks the convergence of the sampling chain to the underlying posterior distribution of the target predictor (Gelman & Rubin, 1992). The parameter setup also followed previous acceptability rating tasks in psycholinguistics (e.g. Paape, Vasishth, & von der Malsburg, 2020; Zhang, Ryskin, & Gibson, 2023).

All \hat{R} s for the sampling chains for all fixed effects were 1.0, indicating successful convergence. We used the package *emmeans* (Lenth, Singmann, Love, Buerkner, & Herve, 2018) to evaluate the main effect exerted by the *de re/de dicto* manipulation and the judgment distinction in each story setting. Here we use β to refer to the coefficient estimate and HPD, i.e., highest posterior density, to refer to the shortest interval with the highest density in the posterior distribution of target coefficient

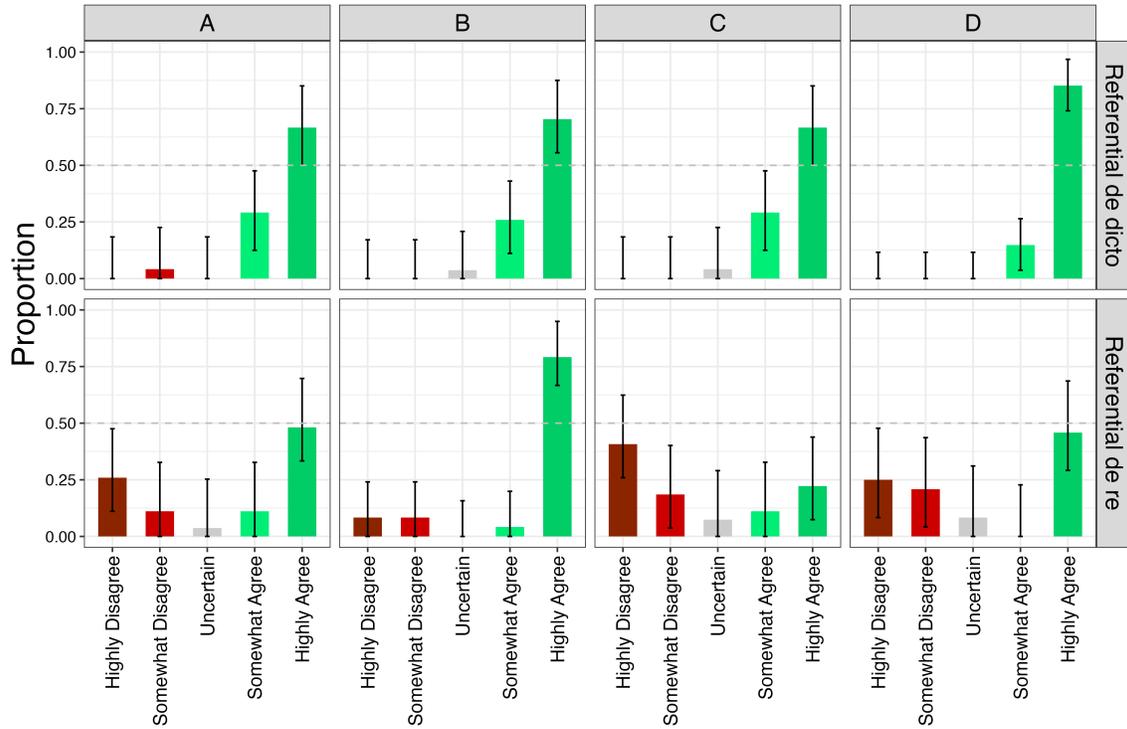


Figure 2: Judgement distribution across the four stories in Exp.1

(Box & Tiao, 2011).

Overall, the referential *de dicto* condition receives more agreement than the referential *de re* ($\beta = 1.66$, HPD = [0.16, 3.12]). Interestingly, the agreement distribution of the two conditions varied by the story: in story A, there doesn't seem to be a difference ($\beta = 1.40$, HPD = [-0.16, 2.99]); neither does story B ($\beta = -1.30$, HPD = [-3.84, 0.95]); in both story C and D, the referential *de dicto* reading is rated as more agreeable than the referential *de re* reading (C: $\beta = 3.48$, HPD = [1.33, 5.98]; D: $\beta = 3.10$, HPD = [0.65, 5.68]). The by-item distribution can also be seen in Figure 2.

When it comes to individual differences, we see that for the referential *de dicto* condition, more than 50% participants chose “Highly Agree” for both trials; nobody chose “Highly Disagree” and only one chose “Somewhat Disagree” once. In comparison, only 15 out of 51 participants (29.4%) chose “Highly agree” twice for the referential *de re* condition; 5 (9.8%) chose “Highly disagree” twice. It is clear that participants showed more disagreement and chose more intermediate options for the referential *de re* readings.

2.3 Discussion

In this experiment, we incorporated the contexts that should in theory support the referential *de dicto* interpretation and the referential *de re* interpretation of belief report into one single story and varied the target sentence to gauge the judgment distribution of the two readings. We found that while the referential *de dicto* reading was overwhelmingly agreed upon, the referential *de re* led to a bimodal pattern of judgments. This successfully replicated [Zhang and Davidson \(2021\)](#) but raised judgment patterns that are not entirely the same as judgments in [Sudo \(2014\)](#).

The by-item investigation shows that while all four stories received almost identical rating patterns for the referential *de dicto* reading, they exhibited different patterns for the referential *de re* reading. Specifically, in story B, the proportions for “Highly Disagree” and “Somewhat Disagree” was the lowest by comparison, replicating the finding in [Zhang and Davidson \(2021\)](#). We speculate that the unique judgment pattern for the *de re* sentence in story B, i.e., *Mrs. Jackson believes that Grace’s gift was sent by Mike*, might be related to the information structure of the passive complement clause (see [Appendix A](#) for the complete experimental details). Since passives could (i) highlight the relative newness of the information in the *by* phrase – it was Mike, not someone else, who sent the gift, and/or (ii) emphasize that the subject is affected by the action denoted by the verb – the gift was sent not received ([Ambridge, Bidgood, Pine, Rowland, & Freudenthal, 2016](#); [Pullum, 2014](#)), this passive construction could potentially modulate the information flow so that readers might focus on verifying information in the predicate and ignore the subject part in the belief report. Nevertheless, story D also featured a passive structure in the complement clause, i.e., *Tracy believes that Alice’s spare apron needs to be washed*, but there is still a significant proportion of disagreement on *de re*. We speculate that the *by* phrase in story B could play a role here and we leave to future work for more investigation on the effect of passives as well as the information structure of the complement clause on the judgment of the referential *de re* reading.

The investigation of individual differences shows that while more than half of the participants had no problem with accessing the referential *de dicto* condition, only 30% did so with the referential *de re* condition plus a 10% going for the opposite truth-value judgment.

In sum, the bimodal distribution of *de re* was replicated and observed to be systematic. Future studies are needed to disentangle the effects of linguistic information structure and individual differences on *de re* judgments.

3 Experiment Two

It is clear that the contexts theoretically permitting the referential *de re* in Experiment 1 featured misapprehension of the belief holder and therefore was more complicated than common *de re* contexts that usually feature an ignorance context (e.g. (1) and (3)). Besides, to our best knowledge, there has been no experimental research that lays out the judgments of the canonical *de re/de dicto* paradigm in a systematic way. Therefore, Experiment 2 simplified the context for judgment collection: by juxtaposing a context that only permits the referential *de re* reading of a definite noun phrase and a context that only permits the attributive *de dicto* reading, we aim to bring more relevant and clear-cut evidence for comparison between the two experiments¹¹.

3.1 Methods

3.1.1 Participants

66 participants took this study who self-identified as English monolinguals from the United States (aged 32.94 ± 10.25). They were recruited from the online crowdsourcing platform [Prolific](#) and were paid \$2.00 (\$12-15/hr) for their participation.

3.1.2 Materials & Procedures

The experimental design and materials were very similar to Experiment 1 except that the manipulation took place in the context, not at the sentence level and each context only theoretically permitted one reading. Additionally, the readings to be tested were attributive *de dicto* and referential *de re*.

¹¹The reason we tested the *attributive de dicto* instead of the *referential de dicto* reading of definite noun phrases in Experiment 2 was that the attributive one resembles the *de dicto* reading of an indefinite noun phrase, which is closer to the canonical interpretation of *de dicto* (please compare (1) and (8)).

Each story (81.14 ± 13.75 words) portrayed a protagonist who held a belief. The target sentence to be judged featured a belief report. We created two conditions, the attributive *de dicto* condition and the referential *de re* condition, by varying the context of the story in which the same target sentence is to be evaluated. Table 2 exhibits an example and the full list of the materials is in Appendix B. In the attributive *de dicto* condition, the protagonist Julie believes that whichever poem written by Elizabeth will win – the noun phrase *Elizabeth’s poem* does not refer to any specific individual entity in the mind of the speaker, only the contents of Julie’s mind. In the referential *de re* context, Julie believes of a particular poem (that exists, according to the speaker) that it will win the competition but does not know that the authorship of this poem belongs to Elizabeth – the noun phrase *Elizabeth’s poem* possesses the referential function. The *de re* context in Experiment 2 was also the “ignorance context”. Additionally, for each story, there were three sentences accompanying the target sentence as fillers and controls; of these one was true given the context, one was false, and the third one was undecided because of the lack of verifying information.

Condition 1: Attributive De Dicto				
Julie is a judge of an ongoing poetry competition. She is told that Elizabeth Johnson, one of the best-known poets in the US, submitted a poem to the competition. Julie is a huge fan of Elizabeth. Even though Julie is blind to the authors and does not know which poem is written by Elizabeth, <i>she believes that no matter which poem Elizabeth submitted, it will win the competition.</i>				
Condition 2: Referential De Re				
Julie is a judge of an ongoing poetry competition. <i>She encounters an extremely well-written poem and believes that this poem will be the winner of the competition.</i> This poem happens to be written by Elizabeth Johnson, a well-known poet in the US. But unfortunately, as a judge, <i>Julie is blind to the authors and therefore does not know it is Elizabeth Johnson who wrote this excellent poem.</i>				
Instruction				
According to this story, please indicate to what extent you agree or disagree with the following statements.				
Target sentence				
Julie believes that Elizabeth’s poem will win the competition.				
Highly Disagree	Somewhat Disagree	Uncertain	Somewhat Agree	Highly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Table 2: Example story from Exp.2

For each sentence to rate, participants were asked to map their judgment onto a five-point fully labeled Likert scale where each point had a label and represented the categorical level of agreement. Crucially, the condition manipulation was within-subjects and participants read stories in both conditions. Across the four stories, two were randomly assigned to be attributive *de dicto* and the other two were referential *de re*. The randomization, the counterbalance treatment, and the experimental procedure were the same as in Experiment 1.

3.2 Results

We only analyzed the judgments from participants who passed the practice trials and whose judgments of the fillers were correct more than 75% of the time. As a result, 60 (90.9%) participants contributed to the analysis.

Figure 3 shows that in the attributive *de dicto* condition, more than 75% of the judgments fall within the “highly agree” category and the distribution is strongly skewed towards the agreement edge. In the referential *de re* condition, more than half of the judgments aggregate to the “highly agree” category with a similar skewness pattern. By visual comparison, the proportion of judgments from “highly disagree” to “somewhat agree” in the *de re* condition is slightly larger, indicating that the *de re* interpretation might be less acceptable, but undoubtedly in general both readings are highly acceptable.

We fit the judgment data into Bayesian multilevel cumulative ordinal models. The condition, the story, and their interaction were entered as fixed effects; random intercepts and random slopes for the full fixed effect structure for the subjects were entered as random effects. The prior setting as well as all the other parameters were the same as in Experiment 1.

The Bayesian model shows that all \hat{R} s for the sampling chains for all fixed effects were 1.0, indicating successful convergence. There was no difference between the attributive *de dicto* condition and the referential *de re* in their agreement distribution ($\beta = 0.825$, HPD = [-0.18, 1.76]) and only in story A was there a difference in judgment of the two conditions ($\beta = 1.31$, HPD = [0.069, 2.63]).

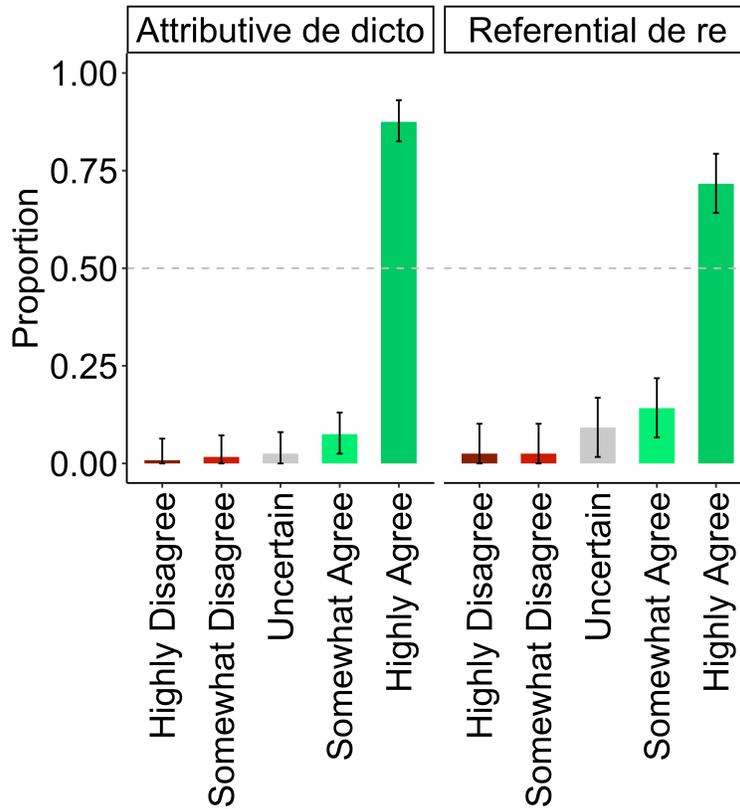


Figure 3: Proportion of different Likert scale choice in the attribute *de dicto* and the referential *de re* condition. (Error bars indicate 95% multinomial confidence intervals calculated by the R package *multinomialCI*)

3.3 Discussion

Experiment 2 found no statistical difference between the referential *de re* reading and the attributive *de dicto* reading of the definite noun phrase in the belief reports. This is consistent with common knowledge about the availability as well as the acceptability of *de re/de dicto* ambiguity of definite noun phrases in belief reports. The comparison between Experiment 1 and 2 reflects that the unacceptability of *de re* reading of definite noun phrases is temporary and conditional on the context. More specially, we know that the *de re* context in Experiment 1 featured the misapprehension of the belief holder while the *de re* context in Experiment 2 featured the ignorance of the belief holder with the *de re* expression.

4 Experiment Three

Experiment 3 was set up to test whether the *de re* reading of definite noun phrases was degraded in the misapprehension context but received full agreement in the ignorance context. To do so, we juxtaposed contexts that only supported the referential *de re* reading as in Experiment 2 (the ignorance context) and contexts that were designed to support both the referential *de dicto* and *de re* reading as in Experiment 1 (the misapprehension context) in a within-subjects design. In both conditions, Experiment 3 asked participants to just rate the sentence with the *de re* interpretation. If this specific context setup plays a role in affecting the acceptability of *de re*, in other words, whether the belief holder has a competing noun phrase in mind associated with the target object, we would expect that in contexts that mirrored Experiment 1, the bimodal distribution would still persist; in contexts that mirrored the *de re* condition in Experiment 2, there would be no to few disagreements for the *de re* readings.

4.1 Methods

4.1.1 Participants

60 participants who self-identified as native speakers of English from the United States (aged 33.02 ± 8.35) were recruited from [Prolific](#). They were paid \$2.00 for their participation.

4.1.2 Materials & Procedures

Table 3 shows an example story. In the referential *de re* condition (the ignorance context), there is only one valid nominal expression referring to the target object but the belief holder is unaware of such relation; the scenario only supports a referential *de re* interpretation of the belief report. In the referential (*de dicto* + *de re*) condition (the misapprehension context), there are two valid nominal expressions: one is interpreted *de re* that the protagonist is unaware of and the other is referential *de dicto* which the protagonist associates with the object in her mind but is wrong in the broader story context. Please see the full list of materials in Appendix C.

<p>Condition 1: Referential <i>de re</i></p> <p>Julie is a judge in an ongoing poetry competition. She encounters an extremely well-written poem and believes that this poem will be the winner of the competition. This poem happens to be written by Elizabeth Johnson, a well-known poet in the US. But unfortunately, as a judge, <i>Julie is blind to the authors and therefore does not know it is Elizabeth Johnson who wrote this excellent poem.</i></p>										
<p>Condition 2: Referential (<i>de dicto</i> + <i>de re</i>)</p> <p>Julie is a judge in an ongoing poetry competition. She encounters an extremely well-written poem about the ocean and believes that this poem will be the winner of the competition. Julie remembers being told that Nicole, one of the best-known poets, submitted a poem about the ocean to the competition. <i>Therefore, Julie concludes that this poem must be written by Nicole and the first prize will be going to her. However, this poem was actually written by Elizabeth, a younger and lesser-known poet.</i> It is just a coincidence that the two poets wrote about the same topic.</p>										
<p>Instruction</p> <p>According to this story, please indicate to what extent you agree or disagree with the following statements.</p>										
<p>Target sentence</p> <p>Julie believes that Elizabeth’s poem will win the competition. (Referential <i>de re</i>)</p> <table style="width: 100%; text-align: center;"> <tr> <td style="width: 20%;">Highly Disagree</td> <td style="width: 20%;">Somewhat Disagree</td> <td style="width: 20%;">Uncertain</td> <td style="width: 20%;">Somewhat Agree</td> <td style="width: 20%;">Highly Agree</td> </tr> <tr> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> <td><input type="radio"/></td> </tr> </table>	Highly Disagree	Somewhat Disagree	Uncertain	Somewhat Agree	Highly Agree	<input type="radio"/>				
Highly Disagree	Somewhat Disagree	Uncertain	Somewhat Agree	Highly Agree						
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>						

Table 3: Example story from Experiment Three

With this context manipulation, participants read the story and judged the following belief report on a five-point Likert scale where the nominal expression inside the belief report was *de re*. There were four scenarios to be read. The counterbalance and the randomization design were kept the same as those in Experiments 1 and 2.

4.2 Results

54 out of 60 participants (90%) contributed to the crucial analysis after the same screening procedure. Aligning with the prediction, Figure 4 shows that the bimodal distribution only appeared in the referential (*de dicto* + *de re*) condition; the statement was highly agreed with in the *de re* only condition. This finding was also supported by Bayesian multilevel cumulative ordinal models. Here both the dum-encoded condition (reference level = *de re* only) and the story (reference level = a) as

well as their interaction were entered as the fixed effects; random intercepts and random slopes for the full fixed effects structure for the subject were entered as the random effect. The priors and all the meta parameters were set the same as in previous experiments. All \hat{R} for the sampling chains for all fixed effects were 1.0, indicating successful convergence. The result shows that overall the *de re* only condition elicited more agreement than the (*de dicto* + *de re*) condition ($\beta = 1.21$, HPD = [0.15, 2.71]).

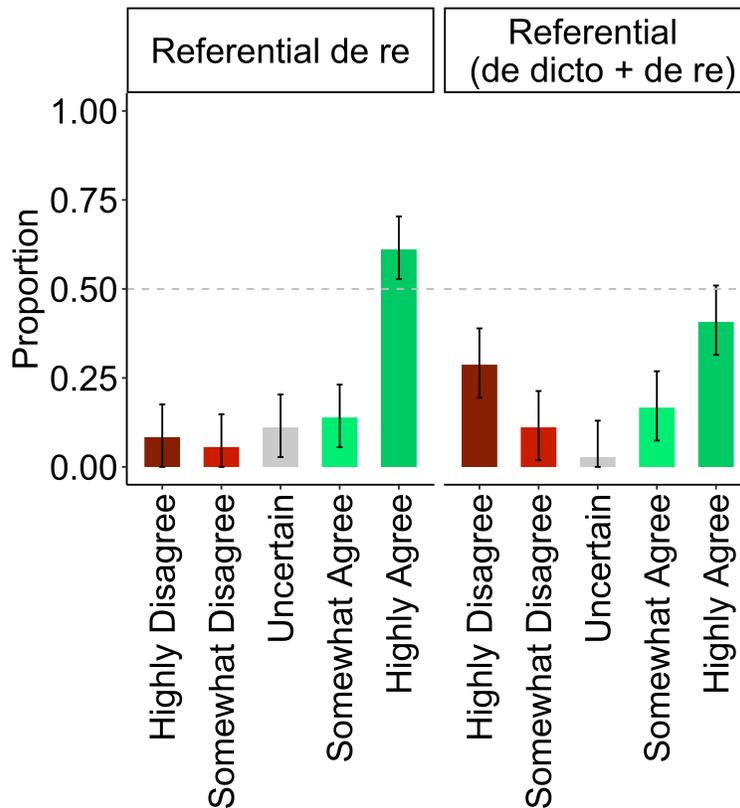


Figure 4: Proportion of different Likert scale choice of the referential *de re* reading in two contexts (The 95% CIs were output from the *MultinomialCI* package). The Referential *de re* context refers to the “ignorance context”. The Referential (*de dicto* + *de re*) context refers to the “misapprehension context”.

Figure 5 exhibits the by-story judgment pattern between the two conditions. We see clearly that within the context that allows both the referential *de dicto* and referential *de re* reading, there are larger proportions of disagreement on the target sentence compared with the *de re* only condition. The statistical analysis shows that in story A and story C, under the (*de dicto*+*de re*) condition, there

was marginally more disagreement than the *de re* only condition (story A: $\beta = 1.06$, HPD = [-0.11, 2.54]; story C: $\beta = 1.29$, HPD = [-0.19, 3.32]). In story B, there wasn't a significant difference between the two conditions ($\beta = 0.22$, HPD = [-1.79, 2.60]). In story D, there was a statistical significance between the *de re* only ratings and the (*de dicto* + *de re*) ratings ($\beta = 2.32$, HPD = [0.58, 4.70]). The peculiarity of story B persisted here.

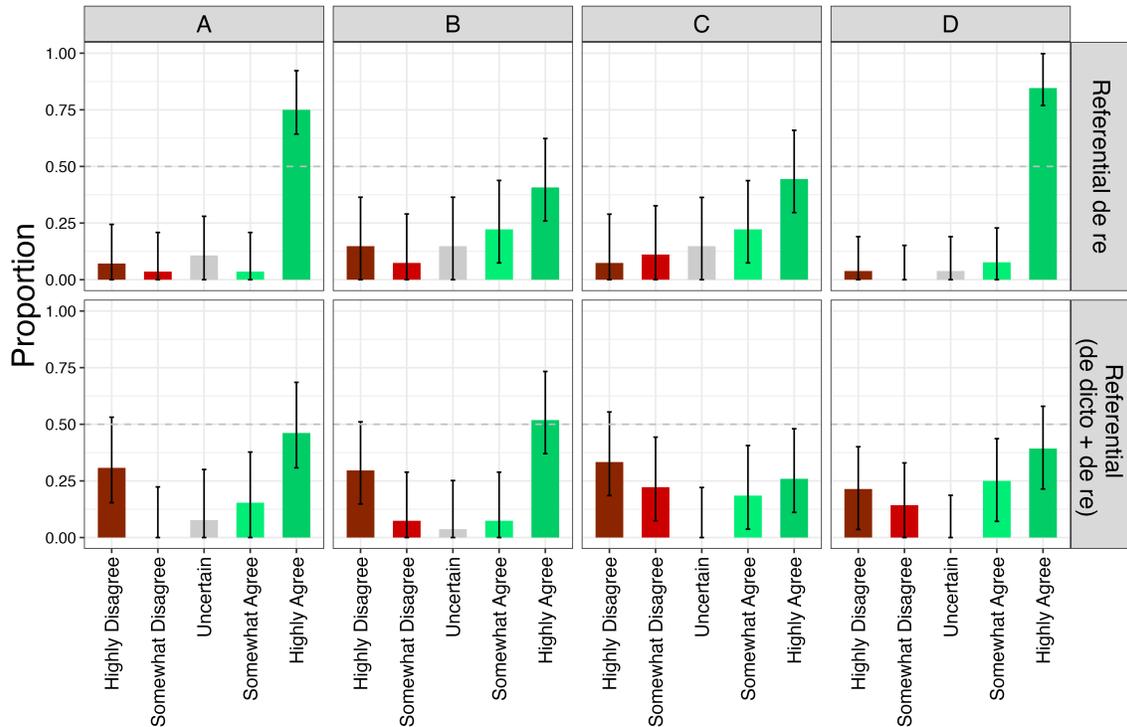


Figure 5: Acceptability ratings on the two conditions across four stories in Exp.3

Similar to Experiment 1, we also observe a difference in the percentage of participants choosing edge judgment labels between the two conditions. In the referential *de re* condition, 18 out of 54 participants (33%) chose “Highly Agree” twice and only 1 participant chose “Highly Disagree” twice. In the referential (*de re* + *de dicto*) condition, 9 out of 54 participants (16.7%) went to “Highly Disagree” for both trials while only 14 (26%) stayed with the “Highly Agree” option twice. This confirms that participants’ judgment behavior changes along with the experimental design.

4.3 Discussion

Experiment 3 further supports our hypothesis: the *de re* readings of definite noun phrases are acceptable in contexts that uniquely support a *de re* reading and that feature an ignorance scenario, and that the acceptability of *de re* is affected in contexts which also permits a referential *de dicto* reading of a competing noun phrase with a misapprehension scenario. We also observe item-wise differences that further support that the *de re* bimodal distribution was caused by contextual factors more than underlying unavailability deriving from the syntactic/semantic structure.

5 General Discussion

One of the biggest challenges in modeling the meaning of a linguistic expression is to understand what aspects of meaning (and more broadly, meaning contributing to acceptability) are stable across contexts and what aspects are context dependent. In this study, we applied an offline truth-value judgment task to investigate contextual effects on the acceptability of *de re* reading of definite noun phrases, with the goal to understand semantic vs. pragmatic factors in *de re* interpretations. In Experiment 1, we replicated the bimodal distribution of the availability of *de re* readings in previous research. In Experiment 2, we found that by changing the context into the canonical *de re* permitting context, the *de re* readings did not suffer from degradation. Inspired by Sudo (2014) we labeled the contexts as “the ignorance context” and “the misapprehension context”. We found that when the *de re* reading of a definite noun phrase was evaluated under the ignorance context – where the context only allowed one noun phrase to refer to the target object and the belief holder did not have that noun phrase in mind – the *de re* reading was almost always accepted; when the context featured a misapprehension scenario – where the belief holder wrongly associated a competing noun phrase with the object – the *de re* reading of the actual noun phrase in the belief report would receive bimodal judgments. Altogether, these results suggest that the *de re* reading of definite noun phrases is generally available in English but that its acceptability is vulnerable to contextual effects, and in particular the effect of competition with the *de dicto* interpretation.

Why might the misapprehension context elicit judgments in favor of the referential *de dicto* reading of the definite noun phrases and against the *de re* reading? We provide our tentative explanation by integrating theories of alternatives with the incremental processing mechanism in psycholinguistics.

We know that alternatives play a critical role in both assessing truth conditional semantics and in tightly related pragmatic processing, where the specific language constrains alternatives that are considered, as does the larger pragmatic context. For example, if we consider the use of alternatives for definite descriptions in an anaphoric environment, we see that the language constrains possibilities for reference: some languages allow entirely covert noun phrases while others do not, some mark definite determiners while others have covert definite determiners, etc. For instance, to express the meaning “I bought a book_{*i*}. BOOK_{*i*} was expensive.”, the specific linguistic construct (e.g. definite noun phrases, pronouns) for BOOK that can be used anaphorically to refer to the book bought by me is determined by available alternative referential nominal expressions in a language and specific anaphoric constraints in that language’s grammar (Ahn, 2020). In Mandarin, the position “BOOK_{*i*}” could be covert or bare noun which influences the relative prominence of other anaphoric expressions. But the lack of these two alternatives in English would render other anaphoric expressions like the definite noun phrase with an explicit definite article (*the book*) more prominent. Another example of linguistic constraints interacting with contextual factors is the well-studied case of processing scalar implicature. As shown by Degen and Tanenhaus (2016), the computation of scalar implicature (i.e., *some* is inferred as *not all*) is affected by the availability of context-specific alternatives. If numerical quantifiers *two* or *three* appeared as alternatives for *some* and *all* to describe the quantity of potential objects, the computation for scalar implicature would be prolonged.

Along the same lines, our misapprehension context simultaneously provides two linguistic means of referring, one the *de dicto* noun phrase and the other the *de re* noun phrase, which automatically sets up these as competing alternatives. During the course of sentence comprehension, readers in the misapprehension context need to critically analyze their differences and carefully resort to their linguistic knowledge to give their judgment. The referential *de dicto* noun phrase is an appropriate

referential expression only in the context of the belief holder's mind, not in the context of the broader story. The referential *de re* noun phrase is appropriate for reference in the broader context of the story, not in the context of the belief holder's mind.

When the participant reads the preamble of the critical statement from left to right, e.g. *Julie believes that ...*, in the poetry competition story, an incremental parser might anchor the mental representation of the sentence to the belief holder's mind and build a discourse model structure that expects further discussion relevant to what the belief holder believes to be true. The upcoming *Elizabeth's poem*, with a referential *de re* reading and evaluated to be true only in the broader context but not in the belief holder's mind, would require a revision in the reader's mental model and thus lead to degraded linguistic judgment. This might be analogous to discussions on the difficulty of getting the inverse scope reading for classic quantification scope sentences like *A boy climbed every mountain* (Altmann & Steedman, 1988; Brasoveanu & Dotlačil, 2019; Fodor, 1982): when hearing the preamble *a boy climbed...*, listeners add a boy in the discourse model and that boy stands in the climbing relation with whatever would come up as the direct object. The inverse scope reading requires the revision of the discourse model to multiple boys, which could be difficult to process. On the other hand, the referential *de dicto* term *Nicole's poem* is a natural follow-up. Along the line, we hypothesize that it might take longer to process the *de re* sentence under this specific context and future online measurements such as self-paced reading or eye-tracking could provide more evidence. To better understand how the word-by-word incremental processing mechanism plays a role in the interpretation of *de re*, it would be interesting to test variations of belief reports such as *Elizabeth's poem will win the competition, Julie believes* and cross-linguistic variations where the complement clause linearly precedes the matrix clause or the matrix verb. The incremental processing hypothesis would predict that the *de re* reading would receive more disagreement as long as the linguistic constituents in that sentence are evaluated with respect to different contexts, regardless of whether content anchored to the context of the broader story or content anchored to the belief holder's mind comes first.

A related, perhaps even more speculative, perspective to understand the interpretive process of

this phenomenon comes from the false belief tasks in the Theory of Mind (ToM) literature (e.g. Apperly, 2010; Apperly & Robinson, 2003; Onishi & Baillargeon, 2005; Wimmer & Perner, 1983). We might expect ToM plays some role in resolving this ambiguity, and yet it makes essentially the opposite prediction to our findings. ToM generally suggests that accessing others' mental status is harder and costs extra cognitive resources (e.g. Gopnik, 1993; Wimmer & Perner, 1983), while here it is the referential *de dicto* (belief holder oriented) reading that is always available and it is the *de re* (speaker oriented) reading that had bimodal acceptability. It remains for future work to understand the role, if any, for ToM in this kind of task.

This work is not without limitations. First, more work needs to be done to explain the item variance in Experiments 1 and 3 where some stories appear to receive more disagreement on the referential *de re* reading than others. Currently, one hypothesis concerns the information structure of the embedded clause: a passive structure with the *by* phrase could drive the interpretive focus to the predicate or the *by* phrase and cause the ignorance of the subtlety in the critical embedded subject (Ambridge et al., 2016; Pullum, 2014), although this hypothesis requires a more well-controlled investigation before reaching a conclusion. Similarly, individual differences between participants in *de re* disagreement are worth exploring. A valuable research question here is whether there is a natural way to demarcate groups of semantic comprehenders (i.e. those who find both *de re* and *de dicto* equally available) versus pragmatic comprehenders who are more sensitive to contextual factors for interpreting sentences and may preference one reading. We might also find a fruitful division of individuals who retain openness to ambiguity versus individuals who find it difficult to switch interpretation once one is found, and further testing might be able to relate performance on this task to aspects of more general cognitive flexibility, or specifically flexibility in retaining ambiguity. It could also be helpful to see whether other contextual factors (e.g. something other than a competing *de dicto* term which also highlights the belief holder's mental status) can also affect the acceptability of *de re* readings. One potential direction is the Question under Discussion (e.g. Roberts, 2012; Ronai & Xiang, 2021): if the context makes the belief holder's mental activities the main topic of discussion, could the *de dicto* reading be even more prioritized than *de re*? Lastly,

we have not explained why *de re* readings are argued to be the more accessible in some legal cases (J. C. Anderson, 2013) and even the default reading (Jaszczolt, 1997), which awaits future study.

Overall, this study provides the first comprehensive experimental investigation into the acceptability of *de re/de dicto* readings of definite noun phrases and explores the effect of context on linguistic judgment. We hope this piece of work lays out the empirical foundation to study the referential properties of noun phrases in the intensional domain and enriches the set of linguistic phenomena that have increasingly attracted experimental methodological inspection (together with Jasbi, Bermudez, Zhang, et al., 2023; Jasbi et al., 2019; Tonhauser, Beaver, & Degen, 2018, a.o.). Our findings also extend the processing of scopes from the classic quantificational scope (C. Anderson, 2004; Brasoveanu & Dotlacil, 2015; Brasoveanu & Dotlačil, 2019; Tunstall, 1998) to the intensional domain. We hope to see more work along the line that discusses the relation between formal semantic representations of a language and the mental processes of the speaker/listener. Furthermore, this study also sheds insight into the interdisciplinary interest of language and mind. Going beyond the developmental trajectory of ToM which has shown biases toward a speaker-oriented perspective (J. C. Anderson, 2013; Jaszczolt, 1997, 2010; Lewis et al., 2017; Wang, Ciranova, Woods, & Apperly, 2020), we show that contextual manipulations can guide readers toward a preference for making reference with respect to others' mental states over one's own.

Data Accessibility Statement

The experimental materials are in the Appendix. The raw data, graphs, and processing codes can be viewed and downloaded from the OSF platform [here](#).

Competing Interests

The authors declare no competing interests.

Authors' Contributions

Both authors participated in the conceptualization, investigation, visualization, formal analysis, validation, and funding acquisition of the study; Author 1 also covered data curation, wrote the original draft, and carried out the revision of the draft; Author 2 covered supervision, the provision of resources, and the editing and revision of the draft.

Acknowledgment

We thank Shannon Bryant, Gennaro Chierchia, Judith Degen, Masoud Jasbi, Joshua Martin, Giuseppe Riccardi, Jack Robinovitch, Uli Sauerland, Jesse Snedeker, and Julia Sturm for their helpful feedback. We appreciate the discussion with the audience at Harvard LangCog Workshop, Experiments in Linguistic Meaning 2020, and Chicago Linguistics Society 2022.

References

- Ahn, D. (2020, April). *THAT thesis: A competition mechanism for anaphoric expressions [PhD thesis]*. LingBuzz. Retrieved 2023-08-16, from <https://ling.auf.net/lingbuzz/004742> (LingBuzz Published In: PhD thesis, Harvard University)
- Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30(3), 191–238. (Publisher: Elsevier)
- Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2016). Is Passive Syntax Semantically Constrained? Evidence From Adult Grammaticality Judgment and Comprehension Studies. *Cognitive Science*, 40(6), 1435–1459. Retrieved 2023-07-30, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12277> (_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12277>) doi: 10.1111/cogs.12277
- Anderson, C. (2004). *The structure and real-time comprehension of quantifier scope ambiguity*. Northwestern University.

- Anderson, J. C. (2013). Misreading like a lawyer: Cognitive bias in statutory interpretation. *Harvard Law Review*, 127(6), 1–74. Retrieved from <https://harvardlawreview.org/2014/04/misreading-like-a-lawyer/>
- Apperly, I. A. (2010). *Mindreaders: The cognitive basis of “theory of mind”*. Psychology Press.
- Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010, June). Why are there limits on theory of mind use? Evidence from adults’ ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology*, 63(6), 1201–1217. Retrieved 2021-04-19, from <http://journals.sagepub.com/doi/10.1080/17470210903281582> doi: 10.1080/17470210903281582
- Apperly, I. A., & Robinson, E. J. (2003, August). When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children’s reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, 85(4), 297–311. Retrieved 2020-02-07, from <https://linkinghub.elsevier.com/retrieve/pii/S0022096503000997> doi: 10.1016/S0022-0965(03)00099-7
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985, October). Does the autistic child have a “theory of mind” ? *Cognition*, 21(1), 37–46. Retrieved 2020-01-28, from <https://linkinghub.elsevier.com/retrieve/pii/0010027785900228> doi: 10.1016/0010-0277(85)90022-8
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013, April). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. Retrieved 2020-08-05, from <https://linkinghub.elsevier.com/retrieve/pii/S0749596X12001180> doi: 10.1016/j.jml.2012.11.001
- Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.
- Brasoveanu, A., & Dotlacil, J. (2015). Sentence-internal same and its quantificational licensors: A new window into the processing of inverse scope. *Semantics and Pragmatics*, 8(1).
- Brasoveanu, A., & Dotlačil, J. (2019, March). Quantification. In C. Cummins & N. Katsos (Eds.), *The Oxford Handbook of Experimental Semantics and Pragmatics* (p. 0). Ox-

- ford University Press. Retrieved 2023-08-10, from <https://doi.org/10.1093/oxfordhb/9780198791768.013.3> doi: 10.1093/oxfordhb/9780198791768.013.3
- Bürkner, P.-C. (2017). **brms** : An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1). Retrieved 2020-07-19, from <http://www.jstatsoft.org/v80/i01/> doi: 10.18637/jss.v080.i01
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395. Retrieved 2020-07-19, from <https://journal.r-project.org/archive/2018/RJ-2018-017/index.html> doi: 10.32614/RJ-2018-017
- Capone, A. (2011). Default semantics and the architecture of the mind. *Journal of Pragmatics*, 43(6), 1741–1754. (Publisher: Elsevier)
- Charlow, S., & Sharvit, Y. (2014, March). Bound ‘de re’ pronouns and the LFs of attitude reports. *Semantics and Pragmatics*, 7(3), 1–43. Retrieved 2020-04-02, from <http://semprag.org/article/view/sp.7.3> doi: 10.3765/sp.7.3
- Chemla, E., & Singh, R. (2014a). Remarks on the experimental turn in the study of scalar implicature, part i. *Language and Linguistics Compass*, 8(9), 373–386.
- Chemla, E., & Singh, R. (2014b). Remarks on the experimental turn in the study of scalar implicature, part ii. *Language and Linguistics Compass*, 8(9), 387–399.
- Cohen, M., Tang, W., & Wang, Y. (2021). De re updates. *arXiv preprint arXiv:2106.11497*.
- Cummins, C., & Katsos, N. (Eds.). (2019). *The Oxford Handbook of Experimental Semantics and Pragmatics* (1st ed.). UK: Oxford University Press.
- Davidson, K. (2020). Is “experimental” a gradable predicate? *Proceedings of NELS 50*.
- Deal, A. R. (2018, December). Compositional paths to de re. *Semantics and Linguistic Theory*, 28, 622. Retrieved 2020-06-13, from <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/28.622> doi: 10.3765/salt.v28i0.4443
- Degen, J., & Tanenhaus, M. K. (2016). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive science*, 40(1), 172–201. (Publisher: Wiley Online Library)

- Donnellan, K. S. (1966). Reference and definite descriptions. *The Philosophical Review*, 75(3), 281–304.
- Elliott, P. D. (2023, April). A flexible scope theory of intensionality. *Linguistics and Philosophy*, 46(2), 333–378. Retrieved 2023-08-09, from <https://doi.org/10.1007/s10988-022-09367-w> doi: 10.1007/s10988-022-09367-w
- Epstein, S., Naumov, P., & Tao, J. (2023). An egocentric logic of de dicto and de re knowing who. *Journal of Logic and Computation*, exad053.
- Fodor, J. D. (1970). *The Linguistic Description of Opaque Contexts* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Fodor, J. D. (1982). The mental representation of quantifiers. In *Processes, beliefs, and questions: Essays on formal semantics of natural language and natural language processing* (pp. 129–164). Springer.
- Frege, G. (1892). Sense and Reference. *The Philosophical Review*, 209–230.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457–472.
- Gibson, E., & Fedorenko, E. (2010, June). Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences*, 14(6), 233–234. Retrieved 2019-09-20, from <https://linkinghub.elsevier.com/retrieve/pii/S1364661310000525> doi: 10.1016/j.tics.2010.03.005
- Gopnik, A. (1993, March). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, 16(1), 1–14. Retrieved 2020-01-28, from https://www.cambridge.org/core/product/identifier/S0140525X00028636/type/journal_article doi: 10.1017/S0140525X00028636
- Jasbi, M., Bermudez, N., & Davidson, K. (2023, January). Default biases in the interpretation of English negation, conjunction, and disjunction. *Experiments in Linguistic Meaning*, 2(0), 129–141. Retrieved 2023-08-06, from <http://journals.linguisticsociety.org/proceedings/index.php/ELM/article/view/5382> (Number: 0) doi: 10.3765/elm.2.5382

- Jasbi, M., Bermudez, N., Zhang, Y., Siro, R., & Davidson, K. (2023). Crosslinguistic Consistency in the Interpretation of Logical Connectives: The case of English, Hungarian, Spanish, and Mandarin Chinese. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45(45). Retrieved 2023-08-07, from <https://escholarship.org/uc/item/9tw5k7ff>
- Jasbi, M., Waldon, B., & Degen, J. (2019, February). Linking hypothesis and number of response options modulate inferred scalar implicature rate. *Frontiers in Psychology*, 10, 189. Retrieved 2020-12-01, from <https://www.frontiersin.org/article/10.3389/fpsyg.2019.00189/full> doi: 10.3389/fpsyg.2019.00189
- Jaszczolt, K. (1997, September). The ‘default de re’ principle for the interpretation of belief utterances. *Journal of Pragmatics*, 28(3), 315–336. Retrieved 2020-09-20, from <https://linkinghub.elsevier.com/retrieve/pii/S0378216697000064> doi: 10.1016/S0378-2166(97)00006-4
- Jaszczolt, K. (1999). *Discourse, Beliefs and Intentions: Semantic Defaults and Propositional Attitude Ascription*. Brill. Retrieved 2023-01-21, from <https://brill.com/display/title/23328>
- Jaszczolt, K. (2005). *Default semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford: Oxford University Press. (OCLC: ocm60559910)
- Jaszczolt, K. (2010). Default semantics. In E. K. Brown, A. Barber, & R. J. Stainton (Eds.), *Concise Encyclopedia of Philosophy of Language and Linguistics* (2nd ed., pp. 128–131). Oxford, UK: Elsevier.
- Jaszczolt, K. (2015). Default semantics. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (2nd ed.). UK: Oxford University Press.
- Katsos, N., & Bishop, D. V. (2011, July). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120(1), 67–81. Retrieved 2020-12-04, from <https://linkinghub.elsevier.com/retrieve/pii/S0010027711000758> doi: 10.1016/j.cognition.2011.02.015
- Keshet, E. (2008). *Good Intentions : Paving Two Roads to a Theory of the De Re / De Dicto*

- Distinction* (Thesis, Massachusetts Institute of Technology). Retrieved 2019-11-17, from <https://dspace.mit.edu/handle/1721.1/45622>
- Keshet, E., & Schwarz, F. (2019). De re/de dicto. *The Oxford handbook of reference*, 167–202.
- Lederman, H. (2022). Fregeanism, sententialism, and scope. *Linguistics and Philosophy*, 45(6), 1235–1275.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version*, 1(1), 3.
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9), 1989–2001.
- Lewis, S., Hacquard, V., & Lidz, J. (2017, October). “Think” Pragmatically: Children’s Interpretation of Belief Reports. *Language Learning and Development*, 13(4), 395–417. Retrieved 2019-09-30, from <https://doi.org/10.1080/15475441.2017.1296768> doi: 10.1080/15475441.2017.1296768
- Marty, P., Chemla, E., & Sprouse, J. (2020, July). The effect of three basic task features on the sensitivity of acceptability judgment tasks. *Glossa: A Journal of General Linguistics*, 5(1), 72. Retrieved 2020-08-05, from <https://www.glossa-journal.org/article/10.5334/gjgl.980> doi: 10.5334/gjgl.980
- Nalborczyk, L., Batailler, C., Løevenbrück, H., Vilain, A., & Bürkner, P.-C. (2019). An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard indonesian. *Journal of Speech, Language, and Hearing Research*, 62(5), 1225–1242.
- Nelson, M. (2019). The de re/de dicto distinction (Supplement to propositional attitude reports). In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.). Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/spr2019/entries/prop-attitude-reports/dere.html>
- Nelson, M. (2023). Propositional Attitude Reports. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Spring 2023 ed.). Metaphysics Research Lab, Stanford Uni-

- versity. <https://plato.stanford.edu/archives/spr2023/entries/prop-attitude-reports/>.
- Onishi, K. H., & Baillargeon, R. (2005, April). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. Retrieved 2020-01-26, from <http://www.sciencemag.org/cgi/doi/10.1126/science.1107621> doi: 10.1126/science.1107621
- Paape, D., Vasishth, S., & von der Malsburg, T. (2020). Quadruplex negatio invertit? The on-line processing of depth charge sentences. *Journal of Semantics*, 37(4), 509–555. Retrieved 2020-04-09, from <https://academic.oup.com/jos/article/37/4/509/5924260?login=true> doi: <https://doi.org/10.1093/jos/ffaa009>
- Partee, B. H., Meulen, A. t., & Wall, R. E. (1990). *Mathematical Methods in Linguistics* (No. 30). Dordrecht ; Boston: Kluwer Academic.
- Percus, O. (2000, September). Constraints on some other variables in syntax. *Natural Language Semantics*, 8(3), 173–229. Retrieved 2019-09-26, from <https://doi.org/10.1023/A:1011298526791> doi: 10.1023/A:1011298526791
- Percus, O., & Sauerland, U. (2003). On the LFs of attitude reports. In *Proceedings of Sinn und Bedeutung* (p. 15). Universität Konstanz, Germany. Retrieved from <http://ling.uni-konstanz.de/pages/conferences/sub7/>
- Pullum, G. K. (2014, July). Fear and loathing of the English passive. *Language & Communication*, 37, 60–74. Retrieved 2023-07-30, from <https://linkinghub.elsevier.com/retrieve/pii/S0271530913000980> doi: 10.1016/j.langcom.2013.08.009
- Quine, W. V. (1956, March). Quantifiers and propositional attitudes. *The Journal of Philosophy*, 53(5), 177. Retrieved 2019-08-27, from http://www.pdcnet.org/oom/service?url_ver=Z39.88-2004&rft_val_fmt=&rft.imuse_id=jphil_1956_0053_0005_0177_0187&svc_id=info:www.pdcnet.org/collection doi: 10.2307/2022451
- Richard, M. (1990). *Propositional attitudes: An essay on thoughts and how we ascribe them*. Cambridge University Press.
- Roberts, C. (2012). *Information structure: Towards an integrated formal theory of pragmatics*.

Semantics and pragmatics, 5, 6–1.

- Robinson, E. J., & Apperly, I. A. (2001, January). Children's difficulties with partial representations in ambiguous messages and referentially opaque contexts. *Cognitive Development*, 16(1), 595–615. Retrieved 2020-02-07, from <http://www.sciencedirect.com/science/article/pii/S0885201400000356> doi: 10.1016/S0885-2014(00)00035-6
- Romoli, J., & Sudo, Y. (2009). De re/de dicto ambiguity and presupposition projection. *Proceedings of Sinn und Bedeutung*, 13(2), 14.
- Ronai, E., & Xiang, M. (2021, November). Pragmatic inferences are QUD-sensitive: an experimental study. *Journal of Linguistics*, 57(4), 841–870. Retrieved 2022-03-01, from https://www.cambridge.org/core/product/identifier/S0022226720000389/type/journal_article doi: 10.1017/S0022226720000389
- Russell, B. (1905). On denoting. *Mind*, 14(56), 479–493. Retrieved from <http://www.jstor.org/stable/2248381>
- Schwarz, F., Clifton Jr, C., & Frazier, L. (2007). Strengthening 'or': Effects of focus and downward entailing contexts on scalar implicatures. *University of Massachusetts Occasional Papers in Linguistics*, 33(1), 9.
- Sison, C. P., & Glaz, J. (1995). Simultaneous confidence intervals and sample size determination for multinomial proportions. *Journal of the American Statistical Association*, 90(429), 366–369.
- Sprouse, J., & Almeida, D. (2017, February). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa: A Journal of General Linguistics*, 2(1), 14. Retrieved 2021-05-22, from <http://www.glossa-journal.org/article/10.5334/gjgl.236/> doi: 10.5334/gjgl.236
- Sudo, Y. (2014). On de re predicates. In *Proceedings of WCCFL* (Vol. 31, pp. 447–456).
- Tonhauser, J., Beaver, D. I., & Degen, J. (2018). How projective is projective content? gradience in projectivity and at-issueness. *Journal of Semantics*, 35(3), 495–542.
- Tonhauser, J., & Matthewson, L. (2015). *Empirical evidence in research on meaning*. The Ohio State University and University of British Columbia.

- Tunstall, S. L. (1998). *The interpretation of quantifiers: Semantics and processing*. University of Massachusetts Amherst.
- von Stechow, P., & Heim, I. (2011). *Intensional Semantics*. MIT. Retrieved from <http://lingphil.mit.edu/papers/heim/fintel-heim-intensional.pdf>
- Wang, J. J., Ciranova, N., Woods, B., & Apperly, I. A. (2020, October). Why are listeners sometimes (but not always) egocentric? Making inferences about using others' perspective in referential communication. *PLOS ONE*, 15(10), e0240521. Retrieved 2021-04-19, from <https://dx.plos.org/10.1371/journal.pone.0240521> doi: 10.1371/journal.pone.0240521
- Wellman, H. M. (1992). *The child's theory of mind*. The MIT Press.
- Wiebe, J., & Rapaport, W. (1986). Representing de re and de dicto belief reports in discourse and narrative. *Proceedings of the IEEE*, 74(10), 1405–1413. Retrieved 2019-07-07, from <http://ieeexplore.ieee.org/document/1457909/> doi: 10.1109/PROC.1986.13641
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*(13), 103–128.
- Zhang, Y., & Davidson, K. (2021, July). De re interpretation in belief reports: An experimental investigation. *Experiments in Linguistic Meaning*, 1, 310. Retrieved 2022-10-11, from <https://journals.linguisticsociety.org/proceedings/index.php/ELM/article/view/4874> doi: 10.3765/elm.1.4874
- Zhang, Y., Ricciardi, G., & Davidson, K. (2021). How many responses in a TVJT? It depends. University of Pennsylvania, PA.
- Zhang, Y., Ryskin, R., & Gibson, E. (2023). A noisy-channel approach to depth-charge illusions. *Cognition*, 232, 105346.

Appendix A Materials in Experiment One

Table 4: Complete material for Exp.1

Story A
<p>Referential <i>De Dicto</i> + Referential <i>De Re</i></p> <p>Julie is one of the judges of an ongoing poetry competition. The best poem that she has read so far is an extremely intriguing poem about the ocean. She believes that this poem will win the competition. Julie remembers being told that Nicole, one of the best-known poets, submitted a poem about the ocean to the competition. Therefore, Julie concludes that this poem must be written by Nicole and the first prize will be going to her. However, this poem was actually written by Elizabeth, a younger and lesser-known poet. It is just a coincidence that the two poets wrote about the same topic.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>
<p>[S1 Target] Julie believes that Nicole’s poem will win the competition. (Referential <i>de dicto</i> reading)</p> <p>[S1 Target] Julie believes that Elizabeth’s poem will win the competition. (Referential <i>de re</i> reading)</p> <p>[S2 True] Elizabeth is a young poet.</p> <p>[S3 False] Elizabeth and Nicole met each other and decided that they will both write poems about the ocean.</p> <p>[S4 Unsure] Julie will also be the judge for the poetry competition next year.</p>
Story B
<p>Referential <i>De Dicto</i> + Referential <i>De Re</i></p> <p>Mr. and Mrs. Johnson have two high school girls, Annie and Grace. One day, Mrs. Johnson finds a wrapped present lying on the front porch of their house. A note on the box says: “From your secret admirer”. Mrs. Johnson remembers that one day she saw Annie’s classmate Mike standing in front of their house for a long time without knocking at the door. She also remembers being told that Annie is very popular in her class, so she concludes that Mike sent the gift to Annie. It turns out that Mike did send the gift, but to Grace. Grace and Mike met each other in a book club, and Mike has admired Grace since then.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>

[S1 Target] Mrs. Johnson believes that Annie's gift was sent by Mike. (Referential *de dicto* condition)

[S1 Target] Mrs. Johnson believes that Grace's gift was sent by Mike. (Referential *de re* condition)

[S2 True] Annie is the youngest daughter of Mrs. Jackson.

[S3 False] Grace and Mike knew each other from jazz band.

[S4 Unsure] The gift was wrapped in pink paper.

Story C

Referential *De Dicto* + Referential *De Re*

Susan works at a hospital. She is responsible for checking in visitors whose relatives and friends are in the maternity ward. One day, a man comes to Susan and asks to visit Haley. His surname is the same as Haley's and they both have beautiful blond hair. Susan remembers Haley saying that she has a brother, so Susan concludes that this man is Haley's brother. Since Haley will deliver the little baby soon, Susan also thinks that the man will accompany Haley for a while. Yet, it turns out that this man is not Haley's brother but instead, Haley's husband. Haley took her husband's surname, and they both have blond hair.

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Susan believes that Haley's brother will accompany her for a while. (Referential *de dicto* condition)

[S1 Target] Susan believes that Haley's husband will accompany her for a while. (Referential *de re* condition)

[S2 True] Haley is receiving medical care in the maternity ward.

[S3 False] Susan thinks the man is related to Haley because of his brown hair.

[S4 Unsure] The man is bringing a bouquet of daisies to Haley.

Story D

Referential *De Dicto* + Referential *De Re*

Alice and Tracy live in the same apartment and always help each other with daily errands. One day, Tracy is gathering up their laundry and she finds an apron with a large coffee stain lying on the sofa. Tracy remembers Alice saying that she usually wears her favorite apron when she cooks and the other day she spilled a cup of coffee while cooking. Tracy thus concludes that what she found is Alice’s favorite apron and it needs to be washed. As a matter of fact, however, what Tracy found is Alice’s spare apron, not her favorite one. Alice’s favorite apron was already in the laundry at the time when she spilled the coffee onto her spare apron.

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Tracy believes that Alice’s favorite apron needs to be washed. (Referential *de dicto* condition)

[S1 Target] Tracy believes that Alice’s spare apron needs to be washed. (Referential *de re* condition)

[S2 True] Alice usually wears an apron when she cooks.

[S3 False] The apron with a large coffee stain was lying on the table when Tracy discovered it.

[S4 Unsure] Tracy altogether gathered three pounds of laundry.

Appendix B Materials in Experiment Two

Table 5: Complete material for Exp.2

Story A	
<p>Attributive <i>De Dicto</i></p> <p>Julie is a judge of an ongoing poetry competition. She is told that Elizabeth Johnson, one of the best-known poets in the US, also submitted a poem to the competition. Julie is a huge fan of Elizabeth. Even though Julie is blind to the authors and does not know which poem is written by Elizabeth, she believes that no matter which poem Elizabeth submitted, it will win the competition.</p>	<p>Referential <i>De Re</i></p> <p>Julie is a judge of an ongoing poetry competition. She encounters an extremely well-written poem and believes that this poem will be the winner of the competition. This poem happens to be written by Elizabeth Johnson, a well-known poet in the US. But unfortunately, as a judge, Julie is blind to the authors and therefore does not know it is Elizabeth Johnson who wrote this excellent poem.</p>

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Julie believes that Elizabeth’s poem will win the competition.

[S2 True] Elizabeth Johnson is a well-known poet in the US.

[S3 False] Julie knows exactly who submitted which poem for the competition.

[S4 Unsure] Julie will also be the judge for the poetry competition next year.

Story B

Attributive *De Dicto*

Mr. and Mrs. Jackson always play the gift exchange game on Christmas eve. In the game, the two of them and each of their five kids prepare a gift. Not knowing who prepares which gift, they take turns to select an anonymous gift and, in this way, get their Christmas gift from their family member. Despite this rule, Mrs. Jackson knows that the youngest daughter Annie always secretly asks the eldest brother David for his gift and then selects his. Mrs. Jackson believes that this year is no exception, even though she does not know which gift is prepared by David.

Referential *De Re*

Mr. and Mrs. Jackson always play the gift exchange game on Christmas eve. In the game, the two of them and each of their five kids prepare a gift. Not knowing who prepares which gift, they take turns selecting an anonymous gift and, in this way, get their Christmas gift from their family. Mrs. Jackson encounters a very beautifully wrapped gift and believes that it will be selected by Annie who always chooses the gift based on how beautiful the wrapping is. Unbeknown to Mrs. Jackson, this gift was prepared by David.

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Mrs. Jackson believes that David’s gift will be selected by Annie this year.

[S2 True] Annie is the youngest daughter of Mrs. Jackson. (Attributive *de dicto* condition)

[S2 True] Annie is the youngest daughter of Mrs. Jackson. (Referential *de re* condition)

[S3 False] Mr. and Mrs. Jackson only have two kids.

[S4 Unsure] All of the gifts will be wrapped in pink paper.(Attributive *de dicto* condition)

[S4 Unsure] The gift encountered by Mrs. Jackson is wrapped in pink paper. (Referential *de re* condition)

Story C

Attributive *De Dicto*

Referential *De Re*

<p>Susan works at a hospital. She is taking care of Haley who will deliver a little baby soon. Susan has not seen Haley’s husband and does not know who he is yet. But she constantly hears Haley talking to her baby “Daddy will be here and will be with mummy for a few days”, which makes Susan believe that her husband will keep her company for a while.</p>	<p>Susan works at a hospital. She is responsible for checking in visitors whose relatives and friends are in the maternity ward. One day, a man comes to Susan and asks to visit Haley. Since Haley will deliver a little baby soon, Susan thinks that the man will keep Haley’s company for a while. Yet Susan fails to ask who the man is because he immediately goes to Haley’s room upon knowing her room number. Luckily, the man is just Haley’s husband, not someone irrelevant or someone who bears ill will.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>	
<p>[S1 Target] Susan believes that Haley’s husband will keep her company for a while. [S2 True] Haley is receiving medical care in the maternity ward. [S3 False] Haley has just delivered a baby. (Attributive <i>de dicto</i> condition) [S3 False] Susan is a surgeon. (Referential <i>de re</i> condition) [S4 Unsure] Haley’s mom will also visit Haley. (Attributive <i>de dicto</i> condition) [S4 Unsure] The man is bringing a bouquet of daisies to Haley. (Referential <i>de re</i> condition)</p>	

<p style="text-align: center;">Story D</p>	
<p>Attributive <i>De Dicto</i></p> <p>Alice and Tracy live in the same apartment and often help each other with daily errands. Alice is a dancer and always wears different beautiful dresses for work. One day, Tracy is gathering their laundry. Before she throws the dirty clothes to the washing machine, she realizes Alice will be home soon and will want to wash what she is wearing after sweating a lot while dancing. Therefore, she pauses the laundry work and waits for Alice to return home.</p>	<p>Referential <i>De Re</i></p> <p>Alice, Lily, and Tracy live in the same apartment and often help each other with their daily errands. One day, Tracy is gathering up their laundry and finds a dirty dress on the sofa. She thinks that the dress needs to be washed and throws it into the washing machine, even though she doesn’t know whose dress it is. It turns out that, unbeknown to Tracy, this dress belongs to Alice and Alice requires it to be washed by hand.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>	
<p>[S1 Target] Tracy believes that Alice’s dress needs to be washed.</p>	

[S2 True] Alice works as a dancer. (Attributive *de dicto* condition)

[S2 True] The dirty dress is lying on the sofa. (Referential *de re* condition)

[S3 False] Alice, Lily and Tracy never help each other with their daily errands.

[S4 Unsure] Alice’s favorite dress is violet. (Attributive *de dicto* condition)

[S4 Unsure] Tracy altogether gathered three pounds of laundry. (Referential *de re* condition)

Appendix C Materials in Experiment Three

Table 6: Complete material for Exp.3

Story A	
<p><i>De Re + De Dicto</i></p> <p>Julie is a judge of an ongoing poetry competition. She encounters an extremely well-written poem about the ocean and believes that this poem will be the winner of the competition. Julie remembers being told that Nicole, one of the best-known poets, submitted a poem about the ocean to the competition. Therefore, Julie concludes that this poem must be written by Nicole and the first prize will be going to her. However, this poem was actually written by Elizabeth, a younger and lesser-known poet. It is just a coincidence that the two poets wrote about the same topic.</p>	<p><i>De Re only</i></p> <p>Julie is a judge of an ongoing poetry competition. She encounters an extremely well-written poem and believes that this poem will be the winner of the competition. This poem happens to be written by Elizabeth Johnson, a well-known poet in the US. But unfortunately, as a judge, Julie is blind to the authors and therefore does not know it is Elizabeth Johnson who wrote this excellent poem.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>	
<p>[S1 Target] Julie believes that Elizabeth’s poem will win the competition.</p> <p>[S2 True] Julie encounters a very good poem submitted to the competition.</p> <p>[S3 False] Julie knows exactly who submitted which poem for the competition.</p>	

[S4 Unsure] Julie will also be the judge for the poetry competition next year.

Story B

De Re + De Dicto

Mr. and Mrs. Jackson always play the gift exchange game on Christmas eve. In the game, the two of them and each of their five kids prepare a gift. Not knowing who prepares which gift, they take turns selecting an anonymous gift and, in this way, get their Christmas gift from their family.

Mrs. Jackson encounters a very beautifully wrapped gift and believes that it will be selected by Annie who always chooses the gift based on how beautiful the wrapping is. She also believes that this gift was prepared by Jane because Jane's gift is usually beautifully wrapped. However, what Mrs. Jackson doesn't know is that this gift was in fact prepared by David, not Jane. It is just a coincidence that this year Jane's and David's gifts were both beautifully wrapped.

De Re only

Mr. and Mrs. Jackson always play the gift exchange game on Christmas eve. In the game, the two of them and each of their five kids prepare a gift. Not knowing who prepares which gift, they take turns selecting an anonymous gift and, in this way, get their Christmas gift from their family.

Mrs. Jackson encounters a very beautifully wrapped gift and believes that it will be selected by Annie who always chooses the gift based on how beautiful the wrapping is. Unbeknownst to Mrs. Jackson, this gift was prepared by David.

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Mrs. Jackson believes that David's gift will be selected by Annie this year.

[S2 True] The gift prepared by David is very beautiful.

[S3 False] Mr. and Mrs. Jackson only have two kids.

[S4 Unsure] The gift encountered by Mrs. Jackson is wrapped in pink paper.

Story C

De Re + De Dicto

De Re only

<p>Susan works at a hospital. She is responsible for checking in visitors whose relatives and friends are in the maternity ward. One day, a man comes to Susan and asks to visit Haley. His surname is the same as Haley's and they both have beautiful blond hair. Susan remembers Haley saying that she has a brother, so Susan concludes that this man is Haley's brother. Since Haley will deliver the little baby soon, Susan also thinks that the man will keep Haley's company for a while. Yet, it turns out that this man is not Haley's brother but instead, Haley's husband. Haley took her husband's surname, and they both have blond hair.</p>	<p>Susan works at a hospital. She is responsible for checking in visitors whose relatives and friends are in the maternity ward. One day, a man comes to Susan and asks to visit Haley. Since Haley will deliver a little baby soon, Susan thinks that the man will keep Haley's company for a while. Yet Susan fails to ask and thus doesn't know who the man is. He immediately goes to Haley's room upon knowing her room number. Luckily, the man is just Haley's husband, not someone irrelevant or someone who bears ill will.</p>
<p><i>According to this story, please indicate to what extent you agree or disagree with the following four statements.</i></p>	
<p>[S1 Target] Susan believes that Haley's husband will keep her company for a while. [S2 True] Haley is receiving medical care in the maternity ward. [S3 False] Susan works as a surgeon at the hospital. [S4 Unsure] The man is bringing a bouquet of daisies to Haley.</p>	

<p style="text-align: center;">Story D</p>	
<p><i>De Re + De Dicto</i></p> <p>Alice, Lily, and Tracy live in the same apartment and often help each other with their daily errands. One day, Tracy is gathering up their laundry and finds a dirty dress on the sofa. Based on the observation that Lily always throws dirty clothes on the sofa, Tracey thinks that the dress belongs to Lily. Tracey also believes that it needs to be washed and thus throws it into the washing machine. It turns out that this dress actually belongs to Alice, not Lily. Furthermore, Alice always requires her dress to be washed by hand. She will be mad after knowing what happens to her dress.</p>	<p><i>De Re only</i></p> <p>Alice, Lily, and Tracy live in the same apartment and often help each other with their daily errands. One day, Tracy is gathering up their laundry and finds a dirty dress on the sofa. She thinks that the dress needs to be washed and throws it into the washing machine, even though she doesn't know whose dress it is. It turns out that, unbeknownst to Tracy, this dress belongs to Alice and Alice requires it to be washed by hand.</p>

According to this story, please indicate to what extent you agree or disagree with the following four statements.

[S1 Target] Tracy believes that Alice's dress needs to be washed.

[S2 True] The dirty dress is lying on the sofa.

[S3 False] Alice, Lily and Tracy never help each other with their daily errands.

[S4 Unsure] Tracy altogether gathered three pounds of laundry.