Competence by Default: Do listeners assume that speakers are knowledgeable when computing conversational inferences?

Alan Bale,[1] Maho Takahashi,[2] Hisako Noguchi,[1] Marguerite Rolland,[1] and David Barner[3]

[1]Linguistics Program, Concordia University

[2]Department of Linguistics, UC San Diego

[3]Department of Psychology, UC San Diego

Contact:

Alan Bale
Linguistics Program
Concordia University
1455 de Maisonneuve West, FB 1030,
Montreal, Quebec, H2G 1M8
Canada
alan.bale@concordia.ca

Abstract

When engaged in conversation, do listeners make default assumptions about the beliefs of speakers? According to some accounts, when listeners hear a sentence like, "Sarah solved some of the math problems" they infer by default that the stronger statement involving "all" is not true (i.e., that Sarah did not solve all of the problems). However, drawing on tests of reading time, eyetracking, and manipulations of cognitive load, multiple studies have argued that this form of inference - called scalar implicature - is not computed by default. In this study, while acknowledging this claim, we explore whether important sub-processes of implicature might nevertheless involve default inferences. In particular, we tested whether listeners assume by default that speakers are knowledgeable about the utterances that they make - a critical precondition for computing scalar implicature. To do this, we tested 60 English speaking participants, who heard utterances made by either knowledgeable speakers - where scalar implicatures were expected - or ignorant speakers - where implicatures were not expected. Also, half of these participants were placed under cognitive load using a dot-array memory task. We replicated the previous finding that, when under load, participants compute fewer implicatures than expected when a speaker is knowledgeable. However, we also found that participants placed under load over-computed implicatures when speakers were ignorant - as though assuming that they were knowledgeable by default. These results are compatible with Neo-Gricean accounts of implicature, but difficult to explain on accounts which argue that scalar implicatures are computed entirely by default.

Keywords: Scalar Implicature, Ignorance Implicature, Default Reasoning, Cognitive Load, Epistemic reasoning, Contextual processing

1. Introduction

A defining property of natural language is that utterances commonly convey information about the truth of stronger propositions that could have been said but weren't. For example, a sentence like "Sarah solved some of the math problems," in (1), implies that a more informative, proposition such as "Sarah solved all of the math problems," in (2), is false (sometimes called a Strong Scalar Implicature, or SSI).

(1) Sarah solved some of the math problems.

(2) Sarah solved all of the math problems.

Most accounts of SSI propose that listeners reason about alternative statements that a speaker could have said in order to compute implications that go beyond the literal meaning of a given utterance.[1] On such accounts, the computation of implicatures requires cognitive resources to generate alternative propositions, process contextual cues, and reason about a speaker's epistemic state (e.g., what information the speaker has access to). Considerable debate surrounds the nature of the computations involved in SSIs and how cognitive resources are deployed when executing them. Some have argued that SSIs like the one triggered by the utterance in (1), which Grice called *Generalized Implicatures*, occur by default unless additional contextual information inhibits or overrides their derivation (see the discussions in Levinson 2000, Zimmerman, 2000; Sauerland, 2004; van Rooij & Schulz, 2004; Geurts, 2010; among others). On some such accounts, the meanings derived by SSIs are incorporated into the preferred literal interpretation

---

[1] Technically, grammatical theories of SSIs, as discussed in Chierchia et al. 2012, incorporate the SSI into the literal meaning of an utterance. Still, the type of processing that is involved in computing a strengthened meaning is very similar to the processes hypothesized in most pragmatic accounts. We will present and address the grammatical point of view more thoroughly in the General Discussion.

of an utterance (e.g., Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2012; Meyer, 2013). In contrast, others have argued that SSIs are not computed by default, but that the computation of an SSI involves a form of rational choice made by the listener that depends on consulting contextual information, including the speaker's mental state. A version of this hypothesis is endorsed by Relevance Theorists such as Sperber and Wilson (1986/1995, 2012) and is also implicit in many Gricean and Neo-Gricean discussions (e.g., see Grice, 1975; Soames, 1982; Leech, 1983; Horn, 1989; Matsumoto, 1995).

Previous studies have sought to adjudicate between these alternative views using a variety of experimental methods that probe online processing of SSIs via eye-tracking, reading time, ERP, and tasks that require participants to compute inferences under cognitive load. Many of these studies have concluded that SSIs are not computed by default, because they (1) take time relative to the computation of literal meanings, (2) are effortful, and (3) are less likely under conditions of cognitive load. For example, in a study by Bott and Noveck (2004), participants were presented with sentences such as "Some elephants have trunks" and were asked to judge whether they were true or false. Using this design, they found that participants who judged the sentences as "true" were faster to reply than those who judged them to be "false". Given that the sentences were only "true" under a literal interpretation and only "false" under an interpretation that included an SSI, Bott and Noveck concluded that the longer response times reflected the additional processing needed to compute an SSI, an effect that they took to be inconsistent with the view that SSIs are computed by default. Similar results, suggesting that literal meanings of statements are computed faster and require less cognitive processing than SSIs have also been reported by numerous subsequent studies (e.g., Noveck & Posada, 2003; Breheny et al., 2006;

Huang & Snedeker, 2009; Tomlinson et al., 2013; Politzer-Ahles et al., 2013; see also Zhao et al. 2015, 2021).

The idea that implicatures are not computed by default has received further support from experiments that test how participants compute SSIs under cognitive load. For example, in an experiment similar to Bott and Noveck (2004), De Neys and Schaeken (2007) also presented participants with sentences like "Some elephants have trunks". However, before the presentation of these sentences, participants were shown a dot pattern that was either easy to recall (e.g., a small number of dots arranged in a straight line) or difficult to recall (e.g., a greater number of dots arranged randomly). Participants then judged whether the sentences were true or false, followed by a trial in which they were asked to recall the dot pattern. Interestingly, when participants saw the more difficult dot pattern before the test sentence, they computed implicatures less often than when they saw the simpler pattern. Based on this result, De Neys and Schaeken (2007) concluded that SSIs are not computed by default, but instead require the use of additional cognitive resources that are inhibited when listeners are placed under cognitive load. In support of this conclusion, similar results using a cognitive load manipulation have also been reported by multiple subsequent studies, including Dieussaert et al. (2011), Marty and Chemla (2013), Marty et al. (2013), and Cho (2020).

Although previous experimental results have repeatedly shown that SSIs require cognitive effort and are not computed automatically, such results do not completely settle the debate concerning SSIs and default heuristics. This is due to the fact that more than one version of the default hypothesis exists. In particular, as discussed by Geurts (2010), a distinction can be made between between the idea that listeners compute SSIs by default—what we will call the "SSI by Default Assumption"—and the hypothesis that listeners make a more modest default assumption, namely that the speaker of an utterance is knowledgeable regarding the truth or falsity of potential alternatives—what we will call the "Competence by Default Assumption" (borrowing from Geurts, 2010). According to most pragmatic accounts of implicature, if a speaker utters a statement such as "Sarah solved some of the math problems," then the listener should compute an SSI only if they believe that the speaker has sufficient evidence to judge whether or not the alternative with "all" is true or false. However, if the listener believes that the speaker does not have enough evidence to make such a judgment, then they should not compute an SSI, given the evidence that the speaker is ignorant. For example, if a listener knows that the speaker only witnessed Sarah successfully solve the first three math problems but that they were absent when Sarah attempted the final two, then such a listener should not compute an SSI from the statement "Sarah solved some of the math problems." According to the Competence-by-Default assumption, listeners assume that speakers know the status of potential alternatives (i.e., have an adequate amount of evidence to judge alternatives) unless other contextual and grammatical information indicates otherwise (Geurts 2010). Thus, on this view, only one component of the SSI—i.e., speaker knowledge—is computed by default. Ignorance is therefore the exception to the rule (e.g., for related discussion, see Zimmermann, 2000, regarding the "Authority Principle", Sauerland, 2004, on the "Epistemic Step", and van Rooij & Schulz, 2004, on the "Competence Assumption").[2]

The SSI-by-Default and the Competence-by-Default assumptions are similar in that they both make a clear distinction between literal interpretation and pragmatically enriched meanings.

[2] As stated by Geurts (2010: 42), "Very roughly, the idea [of Competence by default] is that, in the absence of evidence to the contrary, hearers will assume that the speaker knows what he is talking about." Critically, contextual factors including the speaker's prior statements count as "evidence to the contrary." Because of this nuance, Geurts (2010: 42) states, "The only way to settle this issue [of whether listeners assume speaker Competence by default] would be by collecting quantitative data, but unfortunately I don't see how this might be done."

Also, both hypothesize that deriving a pragmatic interpretation requires cognitive resources to generate alternatives, make inferences, process contextual cues, and evaluate discourse information. However, because they assume different defaults, they also make different experimental predictions. For example, on the SSI-by-Default view, the listener begins with a strong meaning for an utterance which can be modified by various sorts of information, including the discovery that the speaker is ignorant. In contrast, the Competence-by-Default view does not begin with a strong meaning, but merely the assumption that the speaker is knowledgeable—an assumption that is necessary but not sufficient for computing an SSI. Notably, because the Competence-by-Default hypothesis assumes that the listener begins with a literal meaning and requires additional pragmatic reasoning to derive an SSI, the hypothesis is compatible with previous data arguing that SSIs are not computed by default, but require additional steps.

Given the compatibility of the Competence-by-Default Assumption with previous studies arguing against default SSIs, few studies provide data that directly test Competence-by-Default. In one relevant study, building on the reading-time paradigm of Breheny et al. (2006), Bergen and Grodner (2012) tested how listeners computed SSIs under conditions in which the speaker's knowledge states varied. For example, in one condition the speaker began by making a statement that implied full knowledge such as, "At my client's request, I meticulously compiled the investment report." In a second condition, the speaker began by making a statement that implied partial knowledge such as, "At my client's request, I skimmed the investment report." Following this, in both conditions, the speaker then produced an under-informative statement featuring "some" (e.g., "Some of the real estate investments lost money") which was immediately followed by a third sentence that was compatible with only the "not all" implicature (e.g., "The rest were successful despite the recent economic downturn"). Bergen and Grodner found that reading times were longer for this third sentence under conditions of speaker ignorance (when an implicature was not expected), a result that they took as evidence that SSIs are not computed automatically. However, while the study demonstrated the importance of speaker knowledge to SSIs, it did not test whether participants make the Competence-by-Default Assumption, since there was no condition in which the knowledge state of the speaker was unknown. Consequently, the study did not test whether, in absence of direct evidence about speaker states, listeners are more likely to assume ignorance or knowledge, leaving open whether speaker knowledge is assumed by default, or must be computed on the basis of positive evidence.

In a related study, Dieuleveut, Chemla and Spector (2019) presented participants with a set of ten playing cards, all of the same suit, and a character who could either see all of the cards or only half of them. Participants were then asked whether the character could have said the sentence, "Some of the cards are hearts" given what they could see. Dieuleveut et al. found that participants responded "no" significantly more in the knowledgeable speaker condition (~60%) than in the ignorant speaker condition (~30%), which they took as evidence that listeners were sensitive to speaker knowledge when judging the appropriateness of sentences. In particular, participants were aware that an utterance containing "some" is less appropriate if the speaker knows that a stronger statement containing "all" is possible, but not if the speaker is ignorant. Relevant to the present study, Dieuleveut et al. noted that participants responded "no" in the ignorant speaker condition more often than expected, a result which might be taken as evidence that they often compute implicatures without considering speaker knowledge - i.e., as if computing SSIs by default—and possibly compatible with the Competence-By-Default view. However, as they point out, other factors might also explain these judgments. For example, participants in the ignorant speaker condition might reject utterances containing “some” when “all” is true not because they themselves compute SSIs, but because they are aware that the utterance might create such an implicature, and is therefore potentially misleading. Further, if participants computed SSIs by default in the study, then one might expect high rates of implicature in the knowledgeable speaker condition. But here, a surprisingly large number of participants judged that the utterance containing “some” was appropriate even though it was under-informative and a stronger statement containing “all” might be preferable, a result not expected if SSIs are computed by default.

In another relevant study, Hochstein et al. (2018) investigated the role of epistemic reasoning in scalar implicature by testing adolescents with autism spectrum disorders (ASDs). They reasoned that if scalar implicature depends upon Gricean epistemic reasoning, then adolescents with ASDs might struggle to use speaker knowledge as a condition for computing SSIs. In the study, participants witnessed a scene in which a speaker sat in front of three boxes, two of which were open, revealing identical contents (e.g., strawberries), the third of which was closed. The speaker then made a statement like, “Some of the boxes have strawberries” after either peeking into the third box (full knowledge condition), or not peeking inside it (partial knowledge condition). After the speaker made their statement, participants were asked whether the speaker knew what was in the third box in order to confirm that they were aware of the speaker’s knowledge state. Next, they were asked whether the third box contained the same objects as the first two (e.g., whether it contained strawberries) in order to test whether they had computed an SSI. A “no” response was taken as evidence that participants computed an SSI, while an “I don’t know” response indicated that they did not (note that “yes” responses were possible but not expected, since the listener never had direct evidence of what was contained in the third box).

As expected, neurotypical adult controls in the Hochstein et al. study generated SSIs in the knowledgeable-speaker condition but not the ignorant speaker condition. In contrast, adolescents with ASDs exhibited a different pattern: they generated SSIs to the same degree in both cases. Of significance to the issue at hand, when they were explicitly asked at the beginning of each trial if the speaker knew what was inside the third box, all participants, including the ASD teens, answered correctly, indicating an ability to differentiate knowledge vs. ignorance. Consequently, in cases where speakers were ignorant but participants with ASDs nevertheless computed SSIs, the conclusion resulting from this inference—e.g., that the speaker knows that there are no strawberries in the third box—contradicted their explicitly attested belief that the speaker did not know the contents of the box. This is notable because it indicates a distinction between their explicit knowledge of the speaker's beliefs vs. the epistemic consequences of the SSI that they computed, as well as a failure to integrate the two. In their discussion, Hochstein et al. noted that this pattern is difficult to explain if we assume that Gricean reasoning about epistemic states is a precondition for SSI, since participants clearly computed SSIs despite having explicit knowledge that speakers were ignorant. Consequently, they argued that the data might be explained if adolescents with ASDs computed SSIs by default, and these default inferences were insensitive to speaker knowledge.

Another possible explanation of the Hochstein et al. findings, not considered in their study, is that adolescents with ASDs don't compute the entire SSI by default, but instead make the weaker Competence-by-Default Assumption, and struggle to override this default assumption using contextually available information about the speaker's knowledge states. On this analysis, a listener might hear an utterance ("some of the boxes contain strawberries"), assume by default that it is spoken from a position of competence, and only after computing the SSI ("not all of the boxes contain strawberries") override the inference to avoid contradiction by considering extra-linguistic contextual information (e.g., evidence that the speaker is in fact ignorant about the boxes they didn't inspect).[3] Critically, if this process of integrating contextual information and the epistemic implications of an utterance were disrupted - as might occur in the case of adolescents with ASDs - then the default assumption of competence would remain uncontested, and an SSI would follow.

As already noted, some studies have investigated the SSI-by-Default hypothesis by placing participants under cognitive load, on the assumption that such tasks will interfere with processes involved in implicature. However, no previous study has tested whether it is possible to interfere with a participant's ability to draw on contextually available speaker knowledge information when computing implicatures. One possibility, for example, is that this integration problem, seen in adolescents with ASD, can also be elicited in neurotypical adults when they are placed under load, allowing for a test of whether neurotypical listeners adopt the Competence-by-Default Assumption. If cognitive load impairs the ability of neurotypical participants to consider speaker knowledge in the service of implicature, then on the Competence-by-Default hypothesis we might expect them to compute SSIs more often than expected under conditions of speaker ignorance.

[3] This view makes the assumption that contextual evidence of a speaker's knowledge or ignorance is only considered after their utterance is completed. This is because a listener cannot know which knowledge states are relevant to an utterance before hearing it (i.e., they cannot know that a speaker is ignorant about utterances that are yet to come). For details, see the General Discussion.

In order to address this question, and to contrast the SSI-by-Default and Competence-by-Default hypotheses, we conducted a study in which we placed neurotypical adults under cognitive load, and tested them using methods similar to those of Hochstein et al. (2018). First, to assess the SSI-by-Default hypothesis, we placed participants under load in conditions where speakers were knowledgeable. Based on previous studies, we predicted that, contrary to this hypothesis, cognitive load should result in fewer SSIs, since it should inhibit processes involved in the strengthening of literal statements. Second, to assess the Competence-by-Default hypothesis, we asked how cognitive load impacted participants in conditions where speakers were ignorant and no SSI was expected. Here, matters are more complex, since past studies predict that load should make it harder to compute SSIs, but the Competence-by-Default hypothesis predicts higher levels of SSI than expected under conditions of speaker ignorance. Specifically, if a listener assumes that a speaker is knowledgeable, then they should believe that an SSI is warranted unless this assumption is overturned by contextual information indicating that the speaker is ignorant. Therefore, under conditions of speaker ignorance, even if cognitive load makes it harder to compute SSIs, we may expect them to occur more frequently than expected if load also impairs the ability to integrate contextual information concerning speaker ignorance. We return to this issue, and how results might be explained by different theories of scalar implicature in the Discussion.

2 Methods

2.1 Participants

We tested 60 English-speaking participants, ranging in age from 18 to 45 years old (mean 23.48; 20 male) at Concordia University, Canada. Of the 60 participants, 17 reported 100% of their daily communication to be in English, 14 reported 80% to be in English, and 6 reported 50% or less to be in English. An additional 23 participants did not provide this information, as it was initially an optional question.

2.2. Design

Methods were adapted from Hochstein et al. (2018), which included younger participants. Given that the current study tested only adults, participants were warned that the experimental tasks were designed for “little kids” and would be “very straightforward.” Participants were asked to go with their “natural instinct and try not to overthink anything”.

Each participant performed a warm-up task and then completed a scalar-implicature task. Whereas the warm-up task was the same for all participants, participants were divided into two groups for the scalar implicature task resulting in a between subjects design. The Cognitive Load group completed the scalar implicature task while also performing an interference task that aimed to place participants under cognitive load. The Control group completed the same scalar implicature task but without any additional cognitive load.

2.2.1 Warm-up Task

The warm-up task was designed to focus participants on the knowledge states of speakers. Details and results of this task are given in the Appendix.

2.2.2 Scalar Implicature Task

The Scalar Implicature task was identical for participants in both the Control group and the Cognitive Load group. In this task, the experimenter placed three brown circular boxes with lids onto the table and re-introduced participants to Farmer Brown (the non-blindfolded character from the previous warm-up task). The experimenter then told the participants the following.

> *Farmer Brown doesn't know what is inside the boxes, but he is going to look, sometimes in all of them, and sometimes only in two of them. Then he is going to tell you what he knows. He's trying to help you. He's not trying to trick you or anything. Then, I'm going to ask you some Yes/No questions about the boxes, and you can use what you saw and what Farmer Brown knows to make an educated guess. Then you can answer either "Yes" or "No". But sometimes if you have no idea, or you can't possibly know what's inside the box, then you should just say, "I don't know". So your options are "Yes," "No," and "I don't know."*

In each trial, Farmer Brown opened and looked in either two or three of the boxes and then made a statement about what was inside them. Farmer Brown always opened Box 1 and Box 2 in such a way that the contents of these boxes were fully visible to the participant. Box 1 and Box 2 always had two foam cubes inside, all of them the same color. Critically, the participants were unable to see what was in the third box. On some trials, Farmer Brown looked into the third box (hereon, *full-knowledge* trials), on others he did not (hereon, *partial-knowledge* trials).

Once Farmer Brown had finished looking inside the appropriate boxes for a given trial, the experimenter then pointed to the third box and asked the participant, "Does Farmer Brown know what's in this box?" Participants were expected to answer *yes* in the full-knowledge trials and *no* in the partial-knowledge trials. If the participants gave the wrong response (N=3 out of 540 trials), the experimenter repeated the question at which point all participants were able to provide correct responses. Once participants had responded correctly, the experimenter then said: "Now Farmer Brown is going to tell you what he knows."

The statement made by Farmer Brown varied between trials. In Full-Knowledge+All trials (FK+All), Farmer Brown peeked into the third box and made a statement using the quantifier *all* (e.g., “All of the boxes have red cubes”). In such trials, participants were expected to infer that, like the two open boxes, the third box contained, e.g., red cubes. In Full-Knowledge+Some trials (FK+Some), Farmer Brown peeked inside the third box and made a statement using the quantifier *some* (e.g., “Some of the boxes have red cubes”). In such trials, participants were expected to infer that the third box did not contain any, e.g., red cubes. In the Partial Knowledge+Some trials (PK+Some), Farmer Brown looked into the first two boxes but not the third, and made a statement using the quantifier *some* (e.g., “Some of the boxes have red cubes”). In such trials, participants were not expected to infer anything about the content of the third box (since Farmer Brown did not know what it contained). For all three trial types, after Farmer Brown made his statement the experimenter pointed to the third box and asked, e.g., “Do you think there are red cubes in this box?”. We expected that participants would respond “yes” in the *FK+all* trials, “no” in the *FK+some* trials, and “I don’t know” in the *PK+some* trials.

Each participant completed 9 trials, three of each type. Within each trial, the same color of cubes was displayed within the open boxes. However, the color of cubes (and the color word used) varied between trials. Each participant was given one of two counter-balanced orders. Order 1 was as follows: 1) FK-All trial, 2) FK+Some trial, 3) PK-Some trial, 4) FK+Some trial, 5) FK+All trial, 6) PK+Some trial, 7) FK+Some trial, 8) PK+Some trial, 9) FK-All trial. Order 2 was the reverse of Order 1.

The Control group completed this Scalar Implicature task without any additional cognitive load. The Cognitive Load group, however, performed an additional task prior to each Scalar Implicature trial. Specifically, prior to each implicature trial, participants were asked to look at a computer screen and to memorize a display that was presented for three seconds. They were told that they would be tested on their ability to recall the display later in the task. The display presented three boxes, each divided into four equal quadrants that varied randomly with respect to whether they contained a dot (see Figure 1).

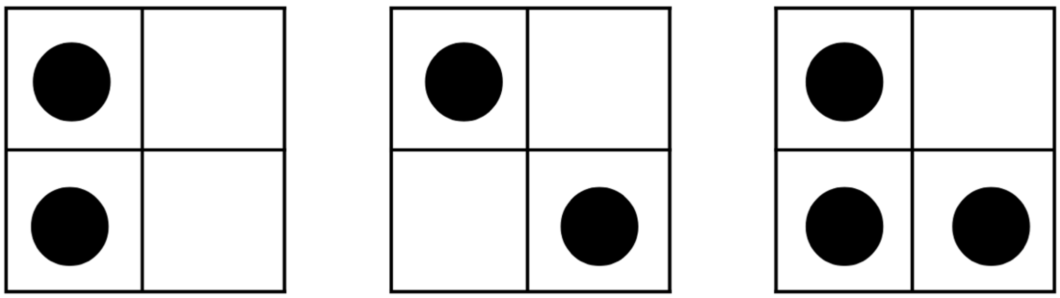

Figure 1. An example of a dot display used in the memorization task.

After viewing the display, participants then completed the Scalar Implicature trial as described above. After the implicature trial, they were then asked to report their memory of the display on a sheet of paper containing three empty boxes.

3. Results

Our first analysis asked whether participants engaged in the cognitive load task, in order to ensure that there was a true difference between the two conditions (which would not be the case if participants in the Cognitive Load condition failed to engage with the task). Participants in the Cognitive Load condition were assigned a score out of 12 on the dot-recall task, where a point was given for each of the 12 quadrants that was correctly replicated. Overall, participants responded correctly on 71.9% of trials (minimum 8.3%, maximum 100%, SD=21.0), which is significantly greater than expected by chance according to a one-sample t-test ($t = 17.1$, $df = 269$, $p < 0.001$).

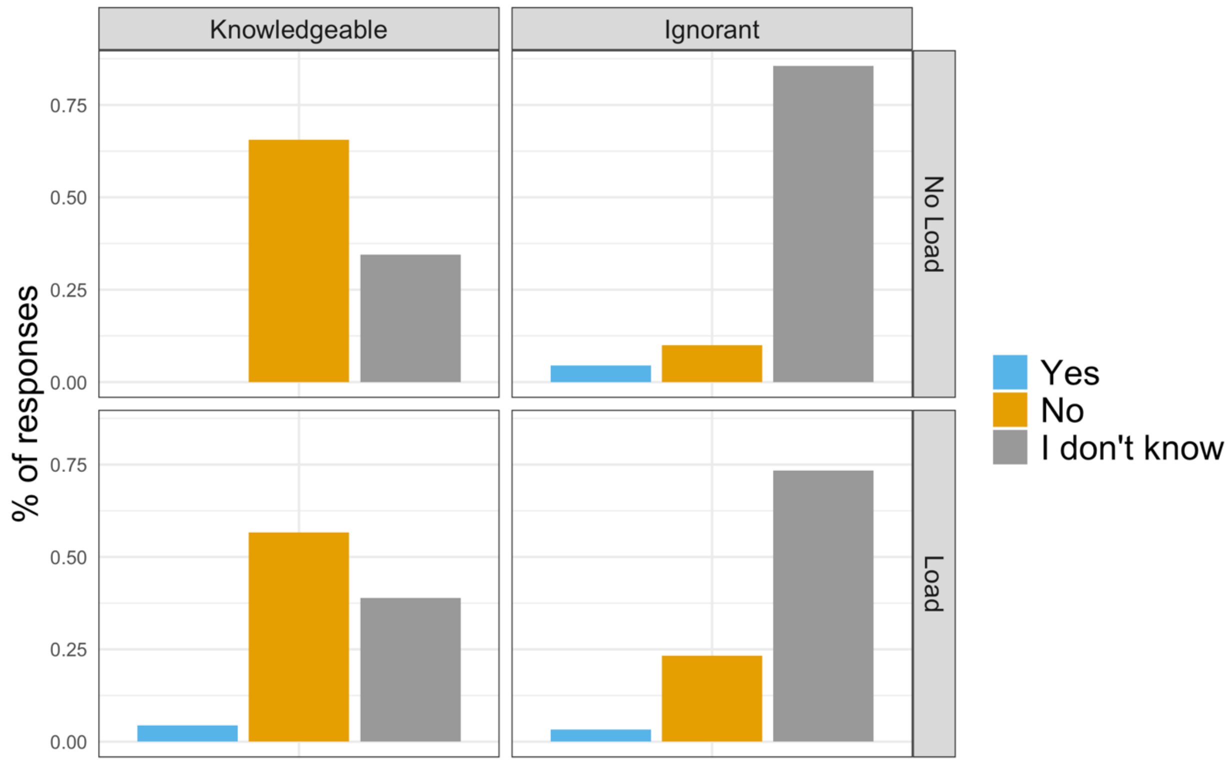

Figure 2. Proportion of “Yes”, “No”, and “I don’t know” responses in trials with *some*.

Participants performed at ceiling (i.e., they responded “Yes”) in trials where *all* was used, as expected. To assess whether cognitive load impacted how participants computed scalar implicatures when speakers were either ignorant or knowledgeable, we constructed two generalized linear mixed-effects models with the *glmer* package (Bates et al., 2015) in R (R Core Team, 2022).[4] Our first model predicted the proportion of “no” responses as a binary outcome in

[4] With the number of participants and items in this experiment, including a further random effect such as random slopes for subjects and random intercepts/slopes of items resulted in the models not converging, and thus only the random intercepts for participants are included.

some trials from Cognitive Load, Knowledge State, and their interaction. The model also included random intercepts of subject and item. This model revealed a significant effect of Knowledge State (β =-5.06, SE = 0.8, $p < .001$), and more importantly, an interaction between Cognitive Load and Knowledge State (β = 2.62, SE = 0.86, $p < .01$). As shown on Figure 2, this was due to two facts. First, on Partial Knowledge+Some trials, participants in the Cognitive Load condition provided more "no" responses than participants in the Control condition (10% to 23.3%; $t = -2.43$, $df = 178$, $p = 0.02$), compatible with making more scalar implicatures. Second, and in contrast, there was no increase in "no" responses due to cognitive load on Full Knowledge+Some trials. Instead, there was a modest, non-significant, reduction in scalar implicatures (65.6% to 56.7%; $t = 1.22$, $df = 178$, $p = 0.22$) when participants were under load, a difference (8.9%) that was similar in size to reductions seen in previous studies that report fewer scalar implicatures under load (De Neys & Schaeken, 2007; Dieussaert et al., 2011; Marty & Chemla, 2013; Marty et al., 2013; Cho, 2020). Furthermore, the reduction in scalar implicatures did not coincide with a significant increase in "I don't know" responses (34.4% to 38.9%; $t = -0.62$, $df = 178$, $p = 0.54$), indicating that the effect of load was not simply to induce confusion and uncertainty.

The analyses presented above indicate that participants performed differently when under load than when not under load. However, they do not test whether accuracy on the load task was related to the likelihood of computing implicatures across different conditions. To probe this, a final, exploratory, analysis asked whether performance on the Cognitive Load (dot memorization) task was related to judgments on the Scalar Implicature task. A glmer model predicting "correct" responses on the Scalar Implicature task (i.e., computing SSIs in the Full Knowledge condition, and not computing them in the Partial Knowledge condition) from score on the Cognitive Load task found that when participants correctly recalled dot arrays, they were more likely to provide correct responses on the Scalar Implicature Task ($\beta = 1.4$, SE = 0.71, p = 0.05).

4. Discussion

We investigated whether listeners make the default assumption that speakers are competent when computing scalar implicatures. To do this, we tested the ability of participants to draw on contextually available speaker knowledge during a test of scalar implicature while they were placed under cognitive load. Compatible with previous tests of the SSI-by-Default hypothesis, we found that when participants were placed under cognitive load in the context of a knowledgeable speaker, they computed slightly fewer SSIs relative to participants who were not under load. New to our study, however, was that when participants interpreted the utterance of a speaker who was ignorant, they computed significantly more implicatures when placed under cognitive load.

These results are similar to those reported by Hochstein et al. (2018) who found that adolescents with ASDs over-computed SSIs in contexts where the speaker was ignorant. In their study, Hochstein et al. raised the possibility that adolescents on the autism spectrum compute SSIs by default without engaging in any kind of epistemic reasoning. A weakness of this account, however, is that it fails to account for mounting evidence that SSIs are not computed by default in neurotypical participants, and would therefore require a different theory of implicature for individuals with ASDs. Although one previous study argues that neurotypical participants overcompute SSIs in response to utterances from ignorance speakers (Dieuleveut et al., 2019), other work argues against SSIs by default, noting that participants take more time to process

SSIs relative to literal meanings (see Noveck & Posada 2003; Bott & Noveck, 2004; Breheny et al., 2006; Huang & Snedeker, 2009; Tomlinson et al., 2013; Politzer-Ahles et al., 2013; Zhao et al. 2015, 2021; among others) and that placing neuro-typical participants under load leads to fewer SSIs, not more, in contexts that are either neutral or favor a strengthened interpretation (see De Neys & Schaeken 2007; Dieussaert et al. 2011; Marty & Chemla 2013; Marty et al. 2013; Cho 2020). In our study, although we did not find a statistically significant reduction in SSIs in the knowledgeable speaker condition, the magnitude of difference between the load and no-load conditions was similar to that found in previous studies—we found a 9% decrease whereas De Neys & Schaeken found a 5% decrease, Dieussaert et al. 4%, and Marty et al. 8% .[5] Finally, as noted in the Introduction, although 30% of participants appeared to have overcomputed SSIs in the ignorant speaker condition of Dieuleveut et al., 40% appeared to undercompute SSIs in the knowledgeable speaker condition, a result that is hard to reconcile with an SSI-by-Default view.

In the Introduction we noted that, rather than computing SSIs by default, listeners may instead make a Competence-by-Default assumption. On this view, individuals with ASDs may overcompute SSIs because they are unable to override a default Competence Assumption using contextually available information about the speaker's knowledge state. Based on this, we reasoned that neurotypical participants placed under load might also overcompute SSIs under conditions of speaker ignorance. Specifically, if cognitive load interferes with their ability to integrate the content of linguistic inferences with contextual information about speaker knowledge states, then participants might compute SSIs even when speakers are clearly ignorant.

[5] Given the similar effect sizes, the difference in significance was likely due to power. Although the number of participants in our study was similar to that of previous studies, we used a between subjects design whereas others were within subjects.

To make this concrete, consider the case in our experiment where a speaker utters the statement in (3) when they are ignorant about the contents of the third box.

(3) Some of the boxes have red cubes.

As noted in the Introduction, prior to hearing the utterance in (3), the listener (i.e., participant in our study) is aware that the speaker is ignorant about the contents of the covered box. However, at this point they do not yet know which utterance the speaker will make, and they therefore have no reason to doubt the speaker's competence about things they might say in the future. Following a standard neo-Gricean analysis of SSI, upon hearing the utterance in (3), three things happen. First, the listener generates the alternative containing "all", in (4). Second, from this they compute the weak implicature in (5) via standard Gricean reasoning (Maxim of Quantity), which is compatible both with an SSI and also with speaker ignorance. This inference reflects the belief that if the speaker had known that all of the boxes had red cubes, then they should have said so. Third, because the listener has assumed that the speaker is competent, the proposition in (6) follows—i.e., that the speaker knows what is contained in all of the boxes.

(4) All of the boxes have red cubes.

(5) It is not the case that the speaker knows that all of the boxes have red cubes.

(6) The speaker knows whether or not all of the boxes have red cubes.

(7) The speaker believes that not all of the boxes have red cubes.

Critically, only after reaching the inference in (6) is the listener in a position to notice a contradiction between the linguistic inference just reached and their prior knowledge gained from the non-linguistic context (i.e., that the speaker is ignorant regarding the contents of the third box). Thus, at this point a participant who is *not* under load should draw upon their contextual knowledge to reject the inference in (6). However, if they do not integrate contextual knowledge in this way, then the SSI in (7) arises—i.e., the speaker believes that not all of the boxes have red cubes. This integration of linguistic and non-linguistic knowledge, we suggest, is the process that is impaired when participants are placed under cognitive load, and that may also be impaired in participants with ASDs. Notably, if listeners do not assume Competence-by-Default, then they should not generate the inference in (6), precluding the possibility of generating the SSI in (7).

So far, we have discussed the data through a Gricean perspective of SSIs. However, as mentioned in the Introduction, an alternative approach to implicatures hypothesizes a purely grammatical analysis of strengthened meanings (see Chierchia, 2004, 2006; Fox, 2007; Chierchia et al., 2012; Meyer, 2013). According to such accounts, SSIs are derived via a phonologically null exhaustivity operator (*Exh*), which has a meaning similar to the adverb *only*. When this operator is inserted into an utterance like "Some of the boxes have red cubes," the resulting meaning is similar to that of the sentence in (8).

(8) Only some of the boxes have red cubes.

Thus, a "some but not all" meaning (i.e., the SSI) is encoded entirely via grammatical representations. On this view, in order to explain the fact that implicatures are not always computed, a sentence such as, "Some of the boxes have red cubes" must permit at least two grammatical parses: one that includes the *Exh* operator and yields an SSI, and another that does not contain such an operator, or at least cancels its effect, and thus is compatible with the speaker being ignorant about alternative statements like "All of the boxes have red cubes" (see the discussion in Singh et al. 2016).[6] On the grammatical view, the choice of an exhaustified parse

[6] Technically, one could achieve the effect of not having the Exh operator by pruning the relevant alternatives that are used to compute an exhaustified reading. For example, all that is needed to block the derivation of the SSI for (9) is the elimination of "All of the boxes have red cubes" as a formal alternative. Thus, the relevant parse could be one that prunes alternatives

yields the same result as adopting the Competence Assumption whereas the choice of a non-exhaustified parse is compatible with speaker ignorance.

How might the grammatical view explain data like those presented in our study? Although it does not commit to default strategies, some proponents of the grammatical view have argued that listeners should default to the strongest parse possible (i.e., the "only some" parse) and thus compute SSIs by default (e.g., see the discussions of SSIs in Chierchia et al. 2012 and Singh et al. 2016; see also Dalrymple et al.'s, 1998, "Strongest Meaning Hypothesis"). On this view, if participants defaulted to the grammatically strengthened meaning (i.e., the parse with the *Exh* operator), then they might fail, under load, to detect a contradiction between the exhaustified meaning and their prior knowledge that the speaker was ignorant about the contents of the closed box. Thus, rather than assuming Competence-by-Default, the grammatical view could explain our data via a parsing preference by adopting a form of the SSI-by-Default view.

However, as already noted, the hypothesis that listeners adopt a strong default parse struggles to explain the fact that listeners compute fewer SSIs under load in full-knowledge trials (as reported in De Neys and Schaeken 2007; Dieussaert et al. 2011; Marty and Chemla 2013; Marty et al. 2013; Cho 2021). If the strongest meaning is the default parse, then one would expect either more SSIs under load, or at least an equal amount. Also, it cannot explain why it takes longer for participants to compute SSIs than to not do so. If SSIs are computed by default, then we might expect a process in which the grammatical *Exh* operator is introduced, an SSI is realized, and then the SSI is canceled when contextual information is considered. To avoid such a consequence, proponents of the grammatical view might insist that implementation of the *Exh* operator is decided by contextual factors that are available to the listener prior to hearing the

---

instead of dropping the Exh operator completely. See the discussion in Crnic et al. 2015 for details.

utterance. However, this confronts the problem mentioned earlier, which is that a listener cannot know the content of an utterance before it is heard, and therefore cannot decide whether to include the *Exh* operator in their parse of a sentence until the semantic content of making this decision becomes apparent. Critically, whereas on the neo-Gricean view this occurs before the SSI is computed (at stage 6 above), on the grammatical view this could only occur after computing an SSI, since the grammatical view does not posit the intermediate steps in (5) and (6) above. Short of adding additional steps such as these (e.g., by positing that a Gricean consultation of context occurs between hearing an utterance and choosing a parse), it is unclear how the grammatical view could account for the processing facts presented here and in previous studies.

Thus far, we have argued that the Competence-By-Default hypothesis is compatible not only with our data, but also with previous studies of implicature. In particular, we have argued that cognitive load leads to a generalized impairment of inferential processes, which should result in fewer SSIs when speakers are known to be knowledgeable (since load impairs the ability to engage in inferences like generating and negating alternatives), and more SSIs than expected when speakers are known to be ignorant (since load impairs the ability to integrate the output of an SSI with prior contextual knowledge about speaker knowledge states). Note, however, that for the Competence-by-Default view to be true, it must also be true that cognitive load has a larger effect on reasoning about speaker knowledge than on generating and negating alternatives. In particular, if the magnitude of these two effects were equal, then we would expect them to cancel out in the ignorant speaker condition, resulting in no difference between the load vs. no-load conditions. SSIs should be more likely if listeners are unable to reason about a speaker's ignorance, but less likely if they are unable to generate and negate alternatives.

Meanwhile, in the knowledgeable speaker condition, we would expect a net reduction of SSIs if these effects were equal. In this case, the Competence-By-Default assumption is compatible with contextual evidence that the speaker is knowledgeable, and thus there is no integration problem that can be impaired by load, leaving only SSI to be impacted.

Critically, although the account just offered posits a type of error—that participants fail to integrate contextual information prior to deriving SSIs—the particular form of error described is not simply noise, or a product of uncertainty. For example, one possibility is that when participants are under load they experience uncertainty, leading their judgments to regress toward 50% levels—with more SSIs in the partial knowledge condition and fewer in the full knowledge condition. However, our data suggests that this is not the case. Recall that in our study, the dependent variable was what participants themselves believed was inside of the covered box. In particular, participants were told that if they didn't know what was inside the box, then they should just reply "I don't know". Consequently, an increase in uncertainty should have resulted in a higher rate of "I don't know responses". However, when participants were placed under load, this is not what we found. Instead, in the ignorant speaker condition, there was a reduction of "I don't know" (from 85.6% to 73.3%) and an increase in "No" responses (from 10% to 23.3%). In the knowledgeable speaker condition, although there was a decrease in "No" responses, participants here changed their responses equally often to "I don't know" and "Yes" responses, such that there was a slight but non-significant increase in "I don't know" responses, rather than a decrease, a pattern more compatible with confusion.

In summary, by placing participants under cognitive load, we tested the hypothesis that listeners adopt the Competence Assumption by default. In support of this hypothesis, we found that, under load, participants compute more strong scalar implicatures than expected if a speaker is ignorant. This finding is compatible with the idea that listeners struggle to integrate the epistemic entailments of an utterance with contextually available information about speaker knowledge. These data can be explained by Gricean and neo-Gricean accounts of scalar implicature, but currently are more difficult to explain on grammatical accounts which omit epistemic reasoning processes from the core computations of implicature. Future studies should explore alternative theoretical solutions for integrating our findings to the grammatical view, and should probe other forms of implicature, to test whether similar results can be found in a more diverse range of phenomena.

Acknowledgments

To be completed after review.

Competing Interests

There are no competing interests.

Data Accessibility Statement

All data are available at https://osf.io/xe65y/?view_only=740ab55c261941fa9efb210985e33559

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

Bott, L. and I. A. Noveck. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. Journal of Memory and Language 51(3), 437-457.

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

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. Structures and Beyond 3, 39-103.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. Linguistic Inquiry 37(4), 535-590.

Chierchia, G., D. Fox and B. Spector. (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. Semantics: An International Handbook of Natural Language Meaning 3, 2297-2332.

Cho, J. (2020). Memory Load Effect in the Real-Time Processing of Scalar Implicatures. Journal of Psycholinguistic Research 49(5), 865-884.

Crnič, L., E. Chemla and D. Fox. (2015). Scalar implicatures of embedded disjunction. Natural Language Semantics 23, 271-305.

Dalrymple, M., M. Kanazawa, et al. (1998). Reciprocal expressions and the concept of reciprocity. Linguistics and Philosophy 21, 159-210.

De Neys, W. and W. Schaeken. (2007). When people are more logical under cognitive load. Experimental Psychology 54, 128–133.

Dieuleveut, A., E. Chemla and B. Spector. (2019). Distinctions between primary and secondary scalar implicatures. Journal of Memory and Language 106, 150-171.

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

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

Geurts, B. (2010). Quantity Implicatures. Cambridge: Cambridge University Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), Syntax and semantics 3: Speech acts, pp. 41–58. New York: Academic Press. Reprinted in and cited from Grice (1989: 22-40).

Hirschberg, J. B. (1985). A Theory of Scalar Implicature. Ph. D. Dissertation. University of Pennsylvania.

Hochstein, L., A. Bale and D. Barner. (2018). Scalar implicature in absence of epistemic reasoning? The case of Autism Spectrum Disorders. Language Learning and Development 14(3), 224-240.

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

Huang, Y. T. and J. Snedeker. (2009). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. Cognitive Psychology 58(3), 376–415.

Leech, G. N. (1983). Principles of Pragmatics. London/New York: Longman.

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

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

Marty, P. and E. Chemla. (2013). Scalar implicatures: working memory and a comparison with only. Frontiers in Psychology 4, 403.

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

Matsumoto, Y. (1995). The conversational condition on Horn scales. Linguistics and Philosophy 18, 21–60.

Meyer, M.-C. (2013). Ignorance and Grammar. PhD dissertation. MIT.

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

Politzer-Ahles, S., R. Fiorentino, X. Jiang and X. Zhou. (2013). Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. Brain Research 1490, 134-152.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sauerland, U. (2004). Scalar implicatures in complex sentences. Linguistics and Philosophy 27: 367–391.

Singh, R., K. Wexler, A. Astle-Rahim, D. Kamawar and D. Fox. (2016). Children interpret disjunction as conjunction: Consequences for theories of implicature and child development. Natural Language Semantics 24, 305-352.

Soames, S. (1982). How presuppositions are inherited: a solution to the projection problem. Linguistic inquiry 13, 483–545.

Sperber, D. and D. Wilson. (1986/1995). Relevance: Communication and cognition. Oxford: Basil Blackwell.

Tomlinson Jr, J. M., T. M. Bailey and L. Bott. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. Journal of Memory and Language 69(1), 18-35.

van Rooij, R. and K. Schulz (2004). Exhaustive interpretation of complex sentences. Journal of Logic, Language and Information 13: 491–519.

Wilson, D. and D. Sperber. (2012). Meaning and Relevance. Cambridge: Cambridge University Zhao, M., T. Liu, G. Chen and F. Chen. (2015). Are scalar implicatures automatically processed and different for each individual? A mismatch negativity (MMN) study. Brain Research 1599, 137-149.

Zhao, M., X. Liu, X. Dai, S. Dong and Z. Han. (2021). Scalar implicature is not a default process: An ERP study of the scalar implicature processing under the effect of focus factor. Brain Research 1765, 147499.

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. Natural Language Semantics 8, 255–290.

Appendix: Warm-up task

A warm-up task was used to focus the attention of participants on the knowledge state of speakers. Participants were asked whether certain statements (about an event that had just happened) were uttered by a blindfolded character or a non-blindfolded character. Participants were expected to attribute accurate statements to the non-blindfolded character and inaccurate statements to the blindfolded character.

In this task, there were 5 trials presented in a fixed order. Before the trials, the participants were introduced to two characters in the form of small action figures, Farmer Brown and Cowboy Black. The experimenter explained to the participants that, "Cowboy Black has a blindfold on, so he cannot see what's happening… [he] might say something funny or not true because he cannot see. He can still hear, but he cannot see anything." Farmer Brown on the other hand, was not blindfolded, and the experimenter explained to participants that, "Farmer Brown will say the correct things because he can see and he can hear."

In each of the 5 trials, the experimenter placed one or two small foam cubes on the table (one in the first two trials and two in the remaining three). The experimenter then brought out a stuffed tiger who announced its intention to take something without naming what it wanted to take. The tiger then proceeded to take a cube (either the only cube present, or one of the two). The experimenter then told the participant, who was watching the scene, that someone said a certain statement. The participants were asked which character, the blindfolded Cowboy Black or the non-blindfolded Farmer Brown, said the statement. They were also asked to justify their response.

In the first two trials, the experimenter presented the participants with a sentence explicitly mentioning sight. In trial 1, the experimenter told the participant that someone said, “The tiger took something. I didn’t see what he took.” In trial 2, the experimenter told the participant that someone said, “I saw the tiger take the yellow cube.” It was expected that the participants would attribute the first statement to the blindfolded character (Cowboy Black) and the second statement to the non-blindfolded character (Farmer Brown).

For each of the remaining three trials (trials 3 to 5), the experimenter placed two small cubes on the table. In these trials, the stuffed tiger first named the colors of the two cubes on the table before announcing its intention to take something. For example, on one trial the tiger said, “It’s me, Tiger. Look, a green cube and an orange cube. Look what I’m taking.” The stuffed tiger then took one of the cubes. On each of the three trials, the experimenter told the participants that someone said that the tiger took a specific cube (e.g., “Someone said ‘The tiger took the orange cube’.”). In trials 3 and 5, the color mentioned by the experimenter matched the actual color of the cube that was taken. Thus, the statement presented to the participants was an accurate description of what happened. In trial 4, the color mentioned by the experimenter was a mismatch and, thus, the statement mentioned by the experimenter did not accurately describe what happened. It was expected that participants would attribute the accurate statements to the non-blindfolded character (Farmer Brown) and the inaccurate statement to the blindfolded character (Cowboy Black).

As expected, participants performed almost perfectly and responded correctly 99% of the time, with 3 unexpected responses out of 300, each given by different participants (1 in the control group, 2 in the group who were subsequently assigned to the cognitive load group).