What Influences Successful Communication? An Examination of Cognitive Load and Individual Differences

[1]Ester Navarro, [2]Brooke N. Macnamara, [3]Sam Glucksberg, [1]Andrew R. A. Conway

[1]Claremont Graduate University

[2]Case Western Reserve University

[3]Princeton University

Author Note

Ester Navarro, Department of Psychology, Claremont Graduate University.

Correspondence concerning this article should be addressed to Ester Navarro Garcia, Department of Psychology, Claremont Graduate University, 150 E 10th St, Claremont, CA. Contact: ester.navarrogarcia@cgu.edu

Abstract

The underlying cognitive mechanisms explaining why speakers sometimes make communication errors are not well understood. Some scholars have theorized that audience design engages automatic processes when a listener is present; others argue that it relies on effortful resources, regardless of listener presence. We hypothesized that (a) working memory is engaged during communicative audience design and (b) the extent to which working memory is engaged relies on individual differences in cognitive abilities and concurrent amount of resources available. In Experiment 1, participants completed a referential task under high, low, or no cognitive load with a present listener, whose perspective differed from the speaker's. Speakers made few referential errors under no and low load, but errors increased when cognitive load was highest. In Experiment 2, the listener was absent. Speakers made few referential errors under no and low load, but errors increased when cognitive load was highest, suggesting that audience design is still effortful under high cognitive load, regardless of the presence of a listener. Experiment 3 tested whether cognitive abilities predicted communication performance. Participants with higher fluid intelligence and working memory capacity made fewer communication errors. Our findings suggest that communication relies on available cognitive resources, and therefore errors occur as a function of factors like cognitive load, and individual differences.

Keywords: Individual differences, communication planning and monitoring, audience design, working memory capacity, fluid intelligence

What Influences Successful Communication? An Examination of Cognitive Load, Listener Interaction, and Individual Differences

Imagine the following scenario: Rachel, a first-year graduate student, is giving a presentation. She is under high cognitive load. As she attempts to describe a slide in her presentation to the audience, she points to her computer screen instead of the screen facing the audience, even though she should know they cannot see where she is pointing. Why would this occur? To communicate effectively, speakers and listeners engage in processes that guide each other's language use, such as perspective taking (Wardlow, 2013). This is often termed audience design, that is, the speaker's ability to produce utterances that are easy to understand, and therefore more likely to be understood by their conversation partner (Ferreira, 2019). This design ranges from organizing the components of a sentence to communicate who did what to whom (e.g., "I kissed her" and "She kissed me" convey different relationships about an event) and selecting words that convey the intended meaning of a sentence ("it's sunny" and "it's cloudy" have different meaning), to avoiding garden path sentences ("The horse raced past the barn fell" vs. "The horse that *was* raced past the barn fell").

Audience design assumes that speakers choose their words based on the listener's ability to understand those words (Kraus & Fussell, 1991) and that speakers develop a knowledge of the conversation that allows them to learn what information is shared (and what information is not) by both conversation partners. This shared information is known as *common ground* and is thought to increase as the conversation continues and partners exchange information. When the information is only known to the speaker (i.e., the addressee has no knowledge of it), this is referred to as *privileged ground*. Through collaborative learning, conversation partners can distinguish common ground from privileged ground (Gorman et al., 2013).

A number of scholars have theorized about how speakers consider common and privileged ground, and how they plan communication utterances for their audiences. Classic views of communication (Grice, 1957, 1989) suggested that speakers and listeners rely upon a "meta-communicative module" that utilizes relatively automatic processes that rely on heuristic mechanisms. Clark's model of optimal design (Clark, 1992; Clark & Marshall, 1981) proposed that speakers consider their audience's perspective and state of mind while planning and monitoring utterances through the use of common ground, adjusting references appropriately. According to Clark and colleagues, the presence of a listener engages relatively automated processes for communication, leading to effortless communication. Similarly, Pickering and Garrod's (2004) interactive alignment account argued that syntactic repetition (i.e., repetition of linguistic features) of speech (Branigan et al. 2000) is a form of audience design in which conversational partners communicate effectively by engaging alignment processes. However, these accounts do not directly address the underlying mechanisms that make audience design effective.

More recent research suggests that the automaticity of audience design in communication might be due to general memory mechanisms. Specifically, conversation partners serve as memory cues for the automatic retrieval of speakers' shared information from memory (Horton, 2007; Horton & Gerrig, 2005). This ordinary-memory view (Brown-Schmidt & Horton, 2014) is based on domain general models of recognition memory (Gillund & Shiffrin, 1984) and is rooted in the idea that that common ground relies on basic memory processes (e.g., Horton & Gerrig, 2002; Horton & Gerrig, 2005). Generally, this view proposes that communication relies largely on basic memory mechanisms and leads to a relatively effortless communication design.

While these models suggest that communication is largely automatic and relies on general memory, it is unclear how they can accurately explain situations in which communication fails, such as when a speaker's utterance does not take into account the listener's perspective. Everyday experience and previous research suggest (e.g., Keysar, Lin, & Barr, 2003) that communication errors occur on a daily basis. For example, as in the scenario presented at the beginning, speakers can forget that some information is accessible to them but inaccessible to their conversation partner, therefore not engaging in audience design. For this reason, some researchers have proposed that oral communication might be cognitively effortful and rely on domain-general processes.

Specifically, Horton and Keysar (1996)'s monitoring-and-adjustment model proposed that speakers' utterances are egocentrically anchored (i.e., from the speaker's perspective). This dual-process theory suggests that communication begins egocentrically but is monitored and adjusted as the speaker recognizes egocentric communication plans via controlled processes. The theory was based on Horton and Keysar (1996)'s findings showing that manipulating time pressure increased references to information that was not accessible to the listener, even when the speaker was aware that the information was not accessible to the listener. That is, participants used their privileged ground and failed to engage in audience design. Their finding suggested that speakers consider communicative perspectives following a two-stage process, where the first planning step is automatic and egocentric, and the second step is controlled via executive control mechanisms. According to this model, errors occur when there is not time to monitor and adjust one's utterances. In this case, the initial plan, which is egocentrically anchored, is uttered.

Horton and Keysar's (1996) two-stage process model is further supported by Ferreira's (2019) framework of feedforward audience design. In Ferreira's framework, executive control is thought to be necessary for message-encoding processes where speakers need to decide the most adequate linguistic and pragmatic features of the message to adequately communicate their intentions. According to this framework, executive control is engaged to create a listener-appropriate message and is therefore essential for audience design. That is, encoding information that is adequately adapted to the communicative situation is crucial for adequate audience design and therefore effective communication. Ferreira's framework extends Horton and Keysar's (1996) idea that speaker communication is divided into automatic and controlled processes and opens the possibility to explaining *when* in the communication process errors occur. However, the framework does not directly address *why* communication is sometimes not successful.

Recent research has begun examining the question of why communication is sometimes fallible. For example, Hawkins, Gweon, and Goodman (2018) examined listener performance on the director task (Dumontheil, Apperly, & Blakemore, 2010) and proposed that the reason why listeners do not always adjust their egocentric utterances when the perspective of the speaker differs from their own is due to a resource-efficient mechanism (i.e., "joint effort") that allows listeners to choose when and how much effort they must allocate for taking the other person's perspective. According to the authors, this decision is dependent on the nature of the speaker; when the speaker is a confederate, there is a mismatch in the expected load that each conversational partner must assume (i.e., the listener expects the speaker to assume a certain amount of load that the confederate does not actually assume as they normally would in a conventional conversation). This leads to greater error rates in participants' performance. Instead, when the speaker is not a confederate, the load is distributed more evenly, helping the listener reduce errors. However, this account focuses on listeners' behavior when a speaker is present. It is unclear whether this joint effort also applies from the perspective of a naive speaker as well as to settings in which the listener is not actually present (and therefore joint load is ambiguous). Given the increasing number of daily activities that require remote conversations (e.g., asynchronous online lectures, phone interviews), it is important to explore the extent to which Hawkins, Gweon, and Goodman's (2018) account applies to these other scenarios.

Another factor that might influence egocentric errors in communication might reside in the mechanism behind cognitively controlled communication. While Horton and Keysar's monitoring-and-adjustment model states that communication planning taxes cognitive resources, it focuses on the temporal aspect of communication processes and the effects of time pressure. However, a study by Müller, Großmann-Hutter, Jameson, Rummer, and Wittig (2001) suggested that manipulating time pressure and cognitive load results in additive effects. This suggests that the *availability* of cognitive resources is important for communication planning and monitoring. That is, reducing general resources (e.g., working memory) via increased cognitive load might result in increased egocentric errors, just like when time pressure is applied.

This is supported by evidence showing that cognitive abilities affect communication performance. Specifically, Wardlow (2013) assessed the relationship between participants' perspective-taking skills on a referential communication task that had competing alternatives, short-term memory, and inhibitory control. Wardlow reported that responses to these tasks were predictive of perspective taking, and therefore that working memory (WM) was a predictor of perspective-taking. However, the tasks Wardlow (2013) employed (i.e., digit span task and flanker task) are not measures of WM, thus hindering the interpretation of the findings. In another study, Ryskin, Benjamin, Tullis, and Brown-Schmidt (2015) tested the relationship between perspective taking skills and a battery of inhibition, WM, and WM capacity tasks. In a language production task, the researchers found that WM capacity positively predicted appropriate information utterances and negatively predicted inappropriate information utterances (i.e., using an adjective to describe an object when it was not necessary). These findings suggest that higher WM capacity is necessary for successful perspective-taking in communication. While the evidence from these studies suggests that WM might be relevant for communication, it is not clear whether (1) WM is always engaged in communication, or (2) cognitive load moderates the effect of WM on communication.

WM is necessary to overcome prepotent (i.e., salient but irrelevant) responses to the environment, such as referring to a salient contextual cue that is inappropriate for the listener. WM is also necessary for maintaining information about the listener's state of mind and remembering non-salient but adequate cues (Kane, Hambrick, & Conway, 2005). This suggests that WM is likely a relevant mechanism for processes required for audience design, like active maintenance and retrieval of information (Unsworth & Engle, 2007). Further, the ability to infer perspective has been shown to vary among speakers and listeners, provoking occasional failures to engage perspective-taking to communicate (Wardlow-Lane, & Ferreira, 2008). Therefore, although speakers commonly engage in audience design when a conversation partner is present by using common ground to design partner-specific utterances (e.g., Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Horton & Gerrig, 2002; Horton & Gerrig, 2005), this might not be the case when the situation involves information that is cognitively challenging or when common ground cannot be easily established.

In the current set of studies, we investigated communication when common ground was limited, and speakers attempted to overcome their privileged ground to avoid miscommunication or communication errors. If speakers are mindful of their privileged ground, they should be able to communicate effectively without making errors. However, as illustrated in the opening scenario, there are situations where speakers fail to consider the perspective of the listener, resulting in egocentric errors. Across these studies, we manipulated cognitive load (Experiments 1 and 2) with a present listener (Experiment 1) and without (Experiment 2) and then tested whether individual differences in cognitive abilities predicted communication performance (Experiment 3). Given that different communication accounts propose that the use of audience design is either automatic (e.g., Horton & Gerrig, 2005) or controlled (e.g., Ferreira, 2019), it is of interest to examine the extent to which these accounts explain communication errors to better understand which of these frameworks adequately explain communication processes.

Experiment 1: Cognitive Load with a Present Listener

Current accounts suggest that the processes behind audience design in communication engage general memory resources and domain-specific processes (e.g., Clark, 1992; Horton, 2007; Horton & Gerrig, 2002; Horton & Gerrig, 2005; Pickering & Garrod, 2004). These models have in common the claim that communication among conversation partners engages largely automated processes for communication planning either via priming mechanisms or cuing general memory. In turn, other accounts (e.g., Horton & Keysar, 1996; Ferreira, 2019) propose that some steps of the communicative process depend on controlled mechanisms and are therefore crucial for adequate audience design. We hypothesized that despite the presence of a listener, audience design will nonetheless rely on working memory (WM) mechanisms (i.e., information processing and maintenance). More specifically, we predict that failures to keep in mind listener-appropriate information (and therefore common ground) are likely to occur during highly cognitively demanding situations. That is, when cognitive load is low, speakers should allocate more cognitive resources to keeping in mind factors such as common ground, thus reducing errors. In contrast, the demands induced by highly demanding tasks, like public speaking or time constraints (Oei, Everaerd, Elzinga, van Well, & Bermond, 2006; Schoofs, Preuß, & Wolf, 2008), can tax WM resources (Porcelli et al., 2008; Qin, Hermans, van Marle, Luo, & Fernández, 2009) that are needed to keep relevant information in mind. If communication processes needed for adequate audience design rely on WM, then cognitive load should negatively impact it, especially when the communication task is difficult and when the contextual cues available to the speaker are not appropriate references for the listener. We predicted that even with a present listener reducing load demands, cognitive load would still negatively impact communication by increasing egocentric errors on a communicative referential task.

Method

Participants and Design

One hundred eight native-English speaking undergraduate students participated as part of a dyad, half randomly assigned to the role of the speaker, the other half as the listener. Cognitive load was manipulated between subjects: 22 participants in the role of the speaker were randomly assigned to the no load condition, 18 were randomly assigned to the low load condition, and 16 were randomly assigned to the high load condition. Two dyads were removed due to missing data, both in the no load condition leaving 20, 18, and 16 dyads, respectively.

Materials

Two identical 24-inch X 24-inch X 24-inch apparatuses (see Figure 1) were created for these studies by a professional art designer. The apparatus contained 15 asymmetrical, abstractly-shaped compartments. The compartment shapes were designed to be difficult to describe. Each compartment had at least one other "sister shape"—another compartment within the apparatus that had similar but not identical features to decrease participants' use of vague communicative utterances. Four different colored “placement pieces” —small pegs placed into the compartments —were used to identify the to-be-described compartments. Experimental sessions were audio recorded via computer software or a portable digital recorder.

A divider was used to prevent participants from viewing each other’s apparatus and placement pieces. Listeners were instructed to move the pieces based on the speaker’s instructions. Listeners were allowed to ask follow-up questions but were not allowed to describe their orientation. Before the listeners arrived, speakers were instructed to view and recall 3 digits (if in the low load condition) or 6 digits (if in the high load condition) and small slips of paper were provided to the speaker to recall the digits in order by writing them on the paper after each trial. To keep the listener-participant blind to the condition of the speaker’s load, digits were presented on a computer screen visible only to the speaker.

Procedure

Speakers entered a room with the apparatus flat on a table. The apparatus was always in the same initial position. The decision of what side should be the initial position was done randomly. Participants were asked to familiarize themselves with it and instructed to view it from all perspectives. Once participants indicated that they were finished familiarizing, the experimenter provided the following instructions for the task:

> I’ll place these four different-colored pegs in four compartments and we’ll audio record you. We’ll be bringing in another participant who will see the same board, but they’ll have the colored pegs next to their board. Your job is to describe where the pegs are, so that the other participant can place the pegs in the right compartments—where you see them. However, the other participant will *not* be viewing the board from the same perspective as you, so describing based on your perspective won’t be helpful. You are

also not allowed to describe your orientation or tell the other participant to change their orientation. I'll place the four pegs, you'll describe them in this order [indicated the sign in which the colors were labeled, 1st, 2nd, 3rd, 4th that remained in view], then I'll rotate the board, place the pegs in compartments again, let you describe them, and rotate again so that we'll have 4 rounds with 4 pegs each round. Each round you'll be speaking to a new person. [*If in the low load or high load condition:*] Also, you will see 3 [6] digits before each round and ask you to recall the numbers in order after you've finished describing the 4 pegs. I'll read you the numbers and then, immediately begin speaking. Do you have any questions?

The task did not begin until participants clearly understood the instructions and all questions were answered. In addition, the board was rotated so that participants could see that compartments look differently in different perspectives. At the end of each trial, the board was rotated clockwise. There were four total rotations (i.e., one per trial) and participants always viewed each side of the board in the same order. The positioning of the pegs in each trial was initially decided randomly by using a random number generator. The positioning of the pegs in each trial was identical across participants. The apparatus design, positioning, and rotation, and the fact that the listener always viewed the board from a different perspective were designed so that directional references to a shape's location relative to another shape's location (e.g., "under," "left of") would be egocentric. That is, no compartment had an absolute directional relative position: "under" or "left of" another compartment would only be correct from the speaker's perspective and never the listener's perspective. However, references to the aspect of a compartment relative to itself (e.g., "small", "pointy") or its location on the board (e.g., "on a corner", "on an edge") were considered correct utterances because these features remained constant regardless of the perspective.

Those assigned to the speaker role arrived in the experimental room before those participants assigned the listener role. The experimenter explained to the speaker that she or he would 1) view 3 digits (if in the low load condition) or 6 digits (if in the high load condition) on a screen when visually prompted by the experimenter, 2) complete a communication task, then 3) write the numbers down in order when prompted by the experimenter. Speakers in the no load condition were not given instructions about digits. The number of digits were selected based on the number of items that would cause a mental load for individuals (Baddeley & Hitch, 1974). When the listener arrived, they were instructed on their role; the listener was tasked with placing the placement pieces in the corresponding compartments within their identical but differently-oriented apparatus based on the speaker' instructions (see Figure 1, Panel A). To be included in the analysis, participants were required to correctly recall over 50% of digits in order.

Results

The effect of cognitive load was assessed examining the number of egocentric utterances (*egocentric errors*) emitted and uncorrected even with the knowledge that the listener's perspective would be different than their own and that therefore egocentric utterances would not benefit and likely hinder the listener (e.g. "under", "to the left of", "above"). In addition, two other dependent variables were measured to examine whether cognitive load affected these measures differently: the average time spent describing in seconds (*time describing*), and (b) the average number of descriptive utterances (*descriptors*), defined as any utterance that described the shape or location of a placement piece without using perspective-dependent information (e.g. "small", "curvy", "on a corner"). Given the instructions that participants were given (see

Procedure), we expected to find a tradeoff between egocentric errors and descriptors. In addition, we expected participants under high load to spend less time describing than those in the other conditions in an effort to offload more quickly the additional load, and because there was less capacity available to effectively monitor communication clarity. This behavior would reflect the difficulty to complete the primary task (i.e., descriptions) while performing at the highest load level of the secondary task (i.e., digit maintenance).

Egocentric Errors

The total number of uncorrected egocentric utterances per participant was recorded. Errors were not normally distributed according to Shapiro-Wilk tests of normality (Royston, 1982), therefore independent samples Kruskal-Wallis for non-normal distributions were used. The Kruskal-Wallis test revealed that egocentric errors were significantly different across levels of load, $p = .001$, indicating that speakers' errors varied as a function of load. A Dunn test (Dunn, 1961) for multiple comparisons revealed that the effect was driven by the high number of errors produced by speakers in the high load condition ($M = 10.31$, $SD = 7.39$, 95% CI [7.41, 13.21]). Specifically, speakers in the high load condition produced significantly more egocentric errors than speakers in the low load condition ($M = 3.39$, SD = 3.04, 95% CI [1.98, 4.79]), $p = .005$, $d = -1.25$, and speakers in the no load condition ($M = 3.65$, SD = 4.64, 95% CI [1.62, 5.68]), $p = .002$, $d=-1.11$. Participants in the low load condition produced similar egocentric errors compared to participants under no load, $p = >.10$, $d = .066$. See Table 1. The internal reliability of the egocentric errors was calculated by conducting a split half correlation analysis between the four trials. Internal reliability of egocentric errors in Experiment 1 was $r = .61$.

Time describing

Time describing was also non-normally distributed according to the Shapiro-Wilk test. A Kruskal-Wallis test demonstrated that participants spent approximately the same amount of time describing regardless of cognitive load, F (1, 52) = .05, p > .10. See Table 1. The internal reliability of the time variable was r = .94.

Descriptors

The number of descriptors was normally distributed in this sample. A one-way ANOVA indicated that cognitive load did not significantly affect the number of descriptors, F (1, 52) = .72, p > .10. See Table 1. The internal reliability of the descriptors variable was r = .89.

Discussion

Experiment 1 examined whether cognitive load negatively impacted communication in a communicative referential task with a naïve listener. It is possible that the presence of the listeners might have allowed speakers to properly engage in audience design and therefore avoid egocentric errors when under no and low load. However, speakers frequently made egocentric errors when under high cognitive load. Time and descriptors did not differ across levels of load. These results suggest that if WM resources are tapped during communication, they are only taxed under high amounts of concurrent load.

Previous research (Kuhlen & Brennan, 2010) has found that speakers interpret the lack of feedback from the listeners as a sign of lack of attention. While this might have affected the information uttered by the speaker, it is unlikely that lack of feedback was responsible for the results of Experiment 1, since all load conditions would have been affected in a consistent manner. More specifically, Kuhlen and Brenann (2010) asked participants to tell jokes, instead of describing visuo-spatial positions. It seems possible that telling jokes might prime speakers to expect a response from the listeners, while visuo-spatial descriptions might not create such an expectation.

To further investigate the effect of cognitive load on communication, we set to study situations in which listeners' feedback was not available, and therefore privileged ground was most salient. For this purpose, we next present an identical experiment, with the exception of the present listener.

Experiment 2: Cognitive Load Without a Present Listener

Experiment 2 assessed whether the pattern of results found in Experiment 1 would differ in the absence of a listener. Specifically, the presence of a listener is thought to reduce cognitive demands by providing memory cues about shared common ground to the speaker. However, in Experiment 1 we observed that increased load increased egocentric errors even without the presence of the listener. If cognitive load affects audience design and communication performance regardless of the presence of a listener, then the same pattern of findings should be observed when a listener is not present.

Method

Participants and Design

Seventy-three native English-speaking undergraduate students participated in this study in exchange for partial course credit. Cognitive load was manipulated between subjects: 25 participants were randomly assigned to the condition in which there was no additional cognitive load applied (*no load*), 26 were randomly assigned to the low cognitive load condition (*low load*), and 25 were randomly assigned to the high cognitive load condition (*high load*). Three subjects were removed as outliers, one in the low load condition and two in the no load condition, so that the final sample was 23 in the no load condition, 25 in the low load condition, and 25 in the high load condition.

Materials

The same apparatus used in Experiment 1 was used in Experiment 2, except only one apparatus board was used (see Figure 1, Panel A).

Procedure

The procedure was identical to Experiment 1, except a second participant serving as the listener was not present. The speaker participant was told that another participant would come later, listen to the recording, and attempt to place the placement pieces that the speaker described. In reality, a listener never listened to the descriptions. Participants were debriefed on the deception at the end of the session.

Results

The outcome variables were the same as in Experiment 1: egocentric errors, time describing, and descriptors. See Table 1 for descriptive statistics.

Egocentric Errors

Shapiro-Wilk tests revealed that errors were non-normally distributed. Kruskal-Wallis tests revealed that the distribution of total egocentric errors was not significantly different across the levels of cognitive load, $p > .10$. Overall, participants made a similar number of egocentric errors than in Experiment 1, where a listener was present (see Table 1). A Dunn test for multiple comparisons revealed that participants under high load ($M = 6.24$, SD = 7.18 95% CI [3.43, 9.05]) committed a similar number of egocentric errors as participants under low load ($M = 5.44$, SD = 5.82, 95% CI [3.16, 7.72]), $p > .10$, but significantly more errors than participants in the no load condition ($M = 2.39$, SD = 2.21, 95% CI [1.49, 3.29]), $p = .035$, $d = .64$. The number of errors did not differ between participants in the low load and no load conditions (*M* = 2.39, SD = 2.21, 95% CI [1.49, 3.29), *p* >.10. See Table 1. The internal reliability of the egocentric errors was calculated by conducting a split half correlation analysis between the four trials (*r* = .59).

Time describing

A Kruskal-Wallis test for non-normally distributed data revealed a significant effect of cognitive load on time describing, *p* = .045. Follow-up tests revealed that participants in the high load condition spend significantly less time describing than those in the low load condition (high: *M* = 21.89, *SD* = 8.55, 95% CI [17.31, 26.46]; low: *M* = 23.44, *SD* = 10.05, 95% CI [18.87, 28.02]), *p* = .034, *d* = .56 and less time than in the no load condition (*M* = 33.20, *SD* = 70.74, 95% CI [24.73, 41.67]), *p* = .012, *d* = .72. See Table 1. The internal reliability of the time variable was *r* = .60.

Descriptors

A Kruskal-Wallis test for non-normally distributed data revealed a significant effect of cognitive load on the number of descriptors, *p* = .013. Follow-up tests revealed that participants in the high load condition used fewer descriptors than those in the no load condition (*M* = 29.35, SD = 14.58, 95% CI [23.63, 35.06]), *p* = .004, *d* = .86. There were no differences in terms of descriptors between participants in the high load and low load, *p* > .10, *d* = .43, and between participants in the no load and low load conditions *p* = .086 *d* = .51. See Table 1. The internal reliability of the descriptors variable was *r* = .30.

Discussion

The results of Experiment 2 indicated that there was no main effect of load on egocentric errors, though pairwise comparisons revealed that participants in the low load condition had significantly more errors than those in the high load condition. This, along with the significant main effects of cognitive load on number of descriptors and time describing, suggests that high amounts of cognitive load may be harmful. We next examined whether individual differences in general cognitive abilities affected egocentric errors.

Experiment 3: Cognitive Load and Individual Differences

The results of Experiments 1 and 2 showed that increasing cognitive load caused a decrease in overall communication performance, suggesting that WM is a relevant factor for communication. It is nevertheless possible that aspects other than cognitive load affect communication performance and adequate audience design. As Hawkins, Gweon, and Goodman (2018) discuss, it is unclear the extent to which the allocation of cognitive resources for perspective taking is a fixed strategy or whether it can be adjusted based on the context or individual differences of the speakers. In particular, studying individual differences in cognitive abilities can help understand some of the factors that affect the allocation of resources for perspective taking.

While WM is the ability to maintain a goal in mind while manipulating information in the face of concurrent action (Baddeley & Hitch, 1974), WM capacity (WMC) is an individual difference measure that predicts performance on other measures, such as general intelligence (Conway, Kane, & Engle, 2003) and reading comprehension (McVay & Kane, 2012). In a conversation, egocentric errors should be controlled via cognitive mechanisms that avoid interference and inhibit prepotent responses. For this reason, taxing WM, as was done in Experiments 1 and 2, was expected to be associated with increased numbers of egocentric errors. If communication processes rely on WM, then individual differences in WMC will interact with cognitive load, allowing individuals with higher WMC to avoid the negative effect of cognitive load by allocating more resources to keeping relevant information in mind (Conway et al., 2005).

Generally, relative to people with lower WMC, people with higher WMC are better able to focus their attention and avoid interference (Conway, Cowan, & Bunting, 2001; Kane, Bleckley, Conway, & Engle, 2001). We hypothesized that these abilities are needed to avoid making egocentric communication errors. There is some evidence that WM plays a role in certain aspects of perspective taking. However, some of these studies used WM (as opposed to WMC) tasks to assess individual differences (Wardlow, 2013) or used only one complex span task (i.e., operation span) to measure WMC (Ryskin, Benjamin, Tullis, & Brown-Schmidt, 2015), which does not allow to correctly capture the WMC construct (Foster et al., 2015). This is likely the reason why these studies have found mixed results on the predictive power of the measures. Therefore, it is unclear if and under what circumstances WMC is related to communication errors. That is, individual differences might have a general effect on communication, but they could also affect performance as a function of cognitive load.

In addition to WMC, we also hypothesized that communication processes may require general cognitive ability and novel reasoning. Specifically, individual differences in fluid reasoning (Gf) capture the ability to reason across a variety of tasks, especially novel ones, allowing novel problem-solving independent of previous knowledge (Cattell, 1963). Higher Gf is related to better novel problem-solving and an increased ability to adapt to new situations, regardless of verbal ability (Engle, Tuholski, Laughlin, & Conway, 1999). Taking the perspective of another person in novel situations, such as remote learning and teaching, as opposed to familiar situations, such as traditional in-person communication, may require additional effort to identify the need to take the other person's perspective (Apperly, 2010, p. 79). This suggests that higher Gf ability is a possible predictor of communication errors when the communicative situation is not familiar, such as in the referential task used in this study. In fact, a number of studies have reported that taking into account other people's perspectives and relevant information in novel situations require more effortful processing, whereas familiar situations are less cognitively effortful and more experience-dependent (Apperly, Back, Samson, & France, 2008; Qureshi, 2009). Despite these findings, no studies have examined whether individual differences in Gf ability are relevant for efficient communication.

Because of their shared processes, both WMC and Gf are strongly related constructs and predict a number of real-life outcomes (Unsworth & Engle, 2007). More importantly, higher scores in these abilities predict better performance in a number of tasks and measures that involve inhibiting prepotent but inappropriate information (Daneman & Carpenter, 1980; Kyllonen & Christal, 1990; Unsworth et al., 2014). Thus, individual differences in WMC and Gf could be related to abilities needed for communication tasks because they might affect (a) the individual's ability to actively maintain information and (b) retrieve relevant information as opposed to irrelevant information from memory (Unsworth and Engle, 2007). Hence, in Experiment 3 we assessed the relationship between WMC and Gf, and communication errors.

Method

Participants and Design

One hundred and one English-speaking undergraduate students were included in the study. Participants participated in two sessions and were paid $10 for each session. An *a priori* power analysis was conducted to ensure that power was sufficient for an individual differences study. Assuming a moderate effect size (R^2 = .10) for a multiple regression analysis with a change in R^2 of 10% a power analysis indicated that at least *N* = 100 participants were necessary to achieve power = .80. The final sample was *N* = 100 after one case was removed due to not following the instructions of the communication task. The study was a correlational design where all participants completed two WMC tasks and one Gf task in two between-subjects conditions of cognitive load. Half of the participants were randomly assigned to the high load condition (n = 50) and half were assigned to the no load condition (n = 50). We used only two load conditions in this experiment to increase power. Both Gf and WMC were measured as two continuous variables. As in the previous experiments, egocentric errors, descriptors, and time describing were used as dependent variables.

Materials and Equipment

Cognitive tasks. WMC was assessed using two complex span tasks: rotation span and reading span. These tasks are designed to tax both the processing and storage components of WM. For each task, participants completed three blocks of trials, which are described in more detail below. This configuration of tasks and number of blocks provides a reliable and valid assessment of WMC in a timely manner (Foster et al., 2015). Together, the two tasks took approximately 45 minutes to complete. Both tasks were automated using E-prime software. Stimuli were presented on a computer screen and responses were collected through mouse clicks.

Rotation complex span task. In this task, participants are shown a letter that can be presented normal or mirror reversed. In each trial, participants need to indicate as fast as possible whether the letter is normal, or mirror reversed by clicking on the Yes or No button. This part of the task taxes the processing component of WM. After each judgment, participants were presented with a new screen. In each new screen, participants saw an arrow that varied in length and direction. Participants had to remember the length and direction of the arrow for subsequent recollection at the end of the trial list. This part of the task taxes the storage component of WM. There are 15 randomly presented trial lists and each consists of an alternating list of processing and storage stimuli. At the end of each trial list, participants need to recall all the storage stimuli they have been presented in serial order (i.e. arrows in the correct length and direction). Sets sizes for the storage component part ranged between two to six items. In total, each participant responded to 42 items.

Reading complex span task. In this task, participants are shown sentences that can be meaningful or senseless. In each trial, participants need to indicate as fast as possible whether the sentence makes sense or not by clicking on the Yes or No button. This part of the task taxes the processing component of WM. After each judgment, participants were presented with a new screen. In each new screen, participants saw a letter in the center of the screen. Participants had to remember the letter in the order in which it was presented at the end of the trial list. This part of the task taxes the storage component of WM. There are 15 randomly presented trial lists and each consists of an alternating list of processing and storage stimuli. At the end of each trial list, participants need to recall in serial order all the storage stimuli they have been presented (i.e. letters). Sets sizes for the storage component part ranged between three to seven items. In total, each participant responded to 75 items.

To obtain the WMC score, a partial credit unit score was calculated for each task and a standardized score was then obtained. The scoring counted all items equally, such that for each trial the proportion of correctly recalled items (i.e., words or arrows) per trial list regardless of list length was calculated. Then, the proportions were averaged. For example, recalling one item (e.g., “E”) correctly from a two-items list (e.g., “E” and “S”) counted as much as recalling two items (e.g., “F” and “R”) from a four-item list (e.g., “F”, “R”, “G”, “A”). Thus, if a participant recalled one item from a two-items list and two items from a four-items list correctly, her score in the task would be (1+2)/6 = .5. Finally, to create a WMC score, the proportion scores obtained in each task were averaged (Conway et al., 2005). By using two complex span tasks with different domain-specific properties, a general WMC construct can be better captured (Engle, Tuholski, Laughlin, & Conway, 1999).

Raven's progressive matrices (RAMP). A short version of Raven's figural inductive reasoning task was used to measure fluid intelligence (Raven, 1938). The decision to use Raven's progressive matrices was based on its high reliability and predictive validity as a measure of fluid reasoning. This task is also highly correlated with WMC (Engle, Tuholski, Laughlin, & Conway, 1999). A total of 18 items from the RAMP Set II were used (Hamel & Schmittmann, 2006). In this task, each item is part of a pattern of eight black and white figures arranged in a 3x3 matrix in which the last bottom right figure is missing. At the bottom of the matrix is a list of eight possible figures to choose from. Only one of those figures is the correct answer that best completes the pattern of the missing piece in the matrix. Figures range from simpler geometrical shapes to complex patterns. In each item there are a series of rules that the participant needs to find and keep in mind to find the right answer. Participants had unlimited time to complete the task, but on average participants spent 20 minutes. Participants were given a short practice set from RAMP Set I before completing the task. A standardized score of correct responses was calculated.

Communication task. The same 24-inch x 24-inch x 24-inch apparatus was used to assess communication following the instructions used in Experiment 2 (see Figure 1, Panel B).

Procedure

The third experiment consisted of two sessions. In the first 45-minute session, participants first completed the WMC tasks. Both tasks were computerized and administered in the same order to all participants. In the second session, participants completed the communication and Gf tasks. Each participant was randomly assigned to one of the two conditions (i.e., high load or no load).

When participants arrived at the testing location, they sat at a table where the apparatus was always placed in the same initial orientation (see Figure 1, Panel B) and the four placement pieces were next to the board. Participants were told to look at the apparatus board and familiarize themselves with it. Next, participants were given the instructions for the communication task (see Experiment 1). The instructions were identical to those used in Experiment 2. Cognitive load was manipulated before the communication task by asking participants in the high load condition to look at 6 digits displayed in the computer screen and remember them in order at the end of the trial. Participants in the no load condition were not asked to remember any digits.

Finally, following the communication task, participants were administered the RAMP test. Participants were given instructions and completed the RAMP I Set as practice. Then, they were administered a paper-based version of the RAMP II. Overall, the second session lasted between 40-50 minutes. At the end of the second session, participants were debriefed of the deception used in the communication task.

Results

Normality tests showed that egocentric errors were non-normally distributed and were assessed using nonparametric Mann-Whitney U tests for independent samples. Descriptors and time describing were normally distributed and therefore were tested using a standard independent *t*-tests.

Egocentric Errors, Time describing, and Descriptors.

A Mann-Whitney *U* test was used to assess differences between high load and no load groups. Results revealed that the effect of load on egocentric errors was significant (U = 951, *p*

= .039, *d* = .58); participants in the high load condition made more egocentric errors (*M* = 14.04, *SD* = 13.72, 95% CI [10.24, 17.84]) than in the no load condition (*M* = 7.94 , *SD* = 5.65, 95% CI [6.37, 9.51]). The internal reliability of the egocentric errors was calculated by conducting a split half correlation analysis between the four trials ($r = .78$.) In terms of time, there were no significant differences by cognitive load condition, $t(98) = -1.06$, $p > .10$, $d = -.21$. The internal reliability of the time variable was $r = .43$. Finally, the effect of cognitive load on descriptors was significant ($t(98) = -3.24$, $p = .002$, $d = -.65$); participants in the high load condition (*M* = 18.35, SD = 10.59, 95% CI [15.42, 21.28]) used fewer descriptors than the no load condition (*M* = 25.19, SD = 10.54, 95% CI [22.27,28.12]). The internal reliability of the descriptors variable was $r = .84$.

In sum, the results for the effect of cognitive load on the dependent variables showed that high load impacted performance: more egocentric errors were made and fewer descriptors were used than in the no load condition (see Table 1), but no significant difference was found for the time variable. Overall, the results largely replicated the findings of Experiment 1 and 2.

Individual Differences Analyses

Correlation analyses (Table 2) showed that egocentric errors were significantly negatively correlated with both Gf ($r = -.37$, $p = <.001$) and WMC ($r = -.30$, $p = .002$). The distribution of egocentric errors was positively skewed and non-normally distributed, therefore a regression model using a Poisson regression for positive count data was used to examine the effect of cognitive abilities on egocentric errors. Three multiple regression analyses were conducted using a Poisson regression. Model 1 examined the effect of load condition, Gf*,* and WMC on egocentric errors. Model 2 analyzed the interaction between Gf and load condition. Model 3 analyzed the interaction between WMC and load condition. WMC and Gf were analyzed separately because even though they are highly correlated, they represent different constructs. Results for all three models are presented in Table 3.

The estimates of Model 1 show that both lower Gf, lower WMC, and condition (i.e., cognitive load) significantly predicted egocentric errors. In addition, Model 2 estimates indicate that the interaction between Gf and condition significantly predicts egocentric errors and Model 3 estimates show that the interaction between WMC and condition also significantly predicts egocentric errors. Figure 2 shows the number of errors as a function of Gf and WMC, respectively. Results indicated that as Gf and WMC scores increased, the number of egocentric errors decreased in the high load condition, while the number of errors remained constant in the no load condition. The results suggest that individual differences play a role in communication performance, especially when cognitive load is high.

Discussion

The results of Experiment 3 largely replicated the results of Experiments 1 and 2; cognitive load increased egocentric errors. Further, the results showed that individual differences in Gf and WMC moderated errors as a function of load. Specifically, in the high load condition, error rates were lower among participants with higher versus lower Gf and WMC. These results are further discussed in the general discussion.

General Discussion

Effective communication is an interactive process where speakers need to tailor the information to the knowledge shared with the listener (i.e., common ground). The process of tailoring information is known as audience design or the speaker's ability to produce utterances that are easy to understand (Ferreira, 2019). Audience design is assumed to involve natural, automatic selection of information that is appropriate for the listener (Kraus & Fussell, 1991).

While much is known about the generation of audience design, the reasons why sometimes speakers do not engage in audience design, creating utterances that are not adequately accessible by the listener, are not clear. For this reason, the current study examined the factors that affect communication when speakers cannot rely on common ground.

While a number of communication models have found evidence for their claims, some of them are contradicting. Classic as well as more recent models suggest that the presence of a listener allows speakers to adequately engage in audience design by either providing constant memory cues or through interactive alignment of syntactic structures (Clark, 1992; Pickering & Garrod, 2004; Horton & Gerrig, 2002). However, other models propose that different steps of the communication process are influenced by controlled processes and are therefore prone to failure if these processes are not available (Horton and Keysar, 1996; Ferreira, 2019). While these models attempt to explain how communication occurs, they do not directly answer the question of why communication fails and what factors influence the failure of the communicative process.

We hypothesized that cognitive load impairs communication (i.e., audience design) by taxing WM. To address the different predictions made by models of communication, we first examined the impact of cognitive load in egocentric errors in the presence of a listener (Experiment 1). We found that cognitive load impaired communication (i.e., egocentric errors), such that speakers frequently made egocentric errors under the highest levels of cognitive load. Our findings suggest that despite the presence of a listener, high levels of load can impact proper audience design. That is, while the listener might provide a reminder of the need to tailor the target message to the listener (Horton & Gerrig, 2002; Horton & Gerrig 2005) when cognitive load is low, our results suggest that listener presence does not eliminate egocentric errors completely. Following the results of Experiment 1, in Experiment 2 we found that cognitive load impaired communication when the listener was absent as well. These findings suggest that when WM is taxed, there are fewer cognitive resources available to overcome egocentric responses, such as keeping in mind appropriate utterances and ignoring inappropriate ones. This set of findings supports the claims made by the monitoring and adjustment model (Horton & Keysar, 1996) and the feedforward audience design model (Ferreira, 2019).

Additionally, these findings are in line with previous research that examines how speakers take the listeners' knowledge into account when the listener is not a confederate. Specifically, Lockridge and Brennan (2002) found that speakers used more efficient utterances (e.g., early in the sentence) when naïve listeners lacked visual aids for abstract, difficult-to-describe instruments compared to when they had a visual aid, suggesting that when speakers realize the naïve listener's needs, speakers are able to adjust information more efficiently. Our findings suggest that relative to participants under cognitive load, individuals not under cognitive load used more descriptive words (Experiment 2 and 3) and made fewer errors (Experiments 1-3) when speakers described abstract, difficult-to-explain shapes to naïve listeners. Since participants in our study did not know whether listeners would be looking at the correct shape in the board (i.e., lacked knowledge about the listeners' access to visual aid), this could have reduced speakers' ability to address the listeners' needs, leading to errors. However, our findings suggest that this only occurred when cognitive load was taxed.

Our findings also suggest that selective attention might not be sufficient to avoid making egocentric errors, as previously suggested (Rubio-Fernandez, 2017). In particular, Rubio-Fernandez (2017) proposed that participants' egocentric eye fixations in the inappropriate location of a shelf do not necessarily represent failure to take the perspective of the speaker, but instead constitute a way to track the perspective of the speaker via selective attention. Selective attention was not assessed in the current study, so it is unclear whether speakers could have used selective attention to correctly describe the abstract shapes. However, speakers in our study failed to remember the perspective of the listener when cognitive load was taxed. This indicates that even if selective attention could have helped speakers track the perspective of the listener when there is little cognitive load, such as in Rubio-Fernandez (2017)'s simplifiedx director task, egocentric errors still occur when cognitive load is taxed.

In line with the experimental findings, individual differences in cognitive resources predicted communication performance and therefore adequate audience design, especially under high cognitive load. Both WMC and Gf predicted communication performance. Our findings suggest that when cognitive load is high, participants with higher WMC and Gf are able to avoid prepotent egocentric responses more often in a communication task that relies on privileged ground. Thus, WMC and Gf are relevant processes for monitoring information and engaging in appropriate audience design. Individuals with higher WMC might be able to allocate more resources to monitor egocentric responses even when cognitive load is high; individuals with higher WMC can focus more attention and avoid more interference than lower WMC individuals (Conway et al., 2001; Kane et al., 2001). In addition, individuals with higher Gf may be able to resolve the novel problem posed by the communication task and by the unknown listener's perspective more efficiently (i.e. avoid privileged ground). The findings suggest that speakers with higher Gf provided more perspective-free utterances by engaging general reasoning abilities to follow the instructions provided by the experimenters regarding the listener's inability to understand perspective-dependent utterances.

Overall, these findings add to theories that propose that speakers might engage in a cognitive *joint effort* to use perspective-taking to varying degrees (Hawkins, Gweon, &

Goodman, 2018). Specifically, our findings suggest that one factor that might contribute to this decision is the amount of resources available for each individual. That is, when cognitive resources are taxed, individual differences in the capacity available to allocate resources plays a role in the extent to which individuals take the perspective of their speaker. Our results indicate that when cognitive load is high, and therefore more cognitive resources are necessary, higher WMC and higher Gf correspond to fewer egocentric errors. This finding also provides support for the view that communication can be cognitively effortful and depend on domain-general mechanisms when the communicative situation is not ideal (i.e., lack of common ground, high cognitive load).

The results of the studies presented here potentially explain why models with competing predictions are nonetheless supported by various research findings. This study does not discredit any of the previously established models. Instead, it expands all models by studying the factors behind communication failure. For example, the interactive alignment account and ordinary memory view suggest that listener presence engages relatively automatic communicative processes. Our results indicate that factors like low cognitive load indeed tap relatively resource-free mechanisms, independently of the presence of a listener. This finding is especially relevant for studies that have examined the domain-general vs. domain-specific dichotomy but have not taken into account the amount of cognitive load that participants were experiencing. In turn, models that account for controlled processes, such as the monitoring-and-adjustment model and feedforward audience design suggest that communication relies on different processing steps and controlled processes are needed when individuals need to adjust the message for a listener. Our results suggest that failures to ignore salient but inappropriate information (i.e., the speaker's own perspective when this is not shared with the listener) occur when the circumstances are more cognitively challenging, such as when cognitive load is high and when individuals have lower WMC and Gf. The difference between studying communication under ideal compared to challenging circumstances should be considered when conducting research on audience design, communication, and perspective taking.

The findings of the study can have wider implications. In particular, this set of studies provides evidence that communication can be effortful. While many instances of “real-life” communication are likely not too cognitively demanding, there are many circumstances where communication is effortful, and errors might lead to failures in communication. As remote teaching, learning, and working become increasingly common, many will be challenged with communicating under higher cognitive load (such as navigating novel online interfaces) in the absence of other individuals to provide feedback and cues (i.e., establish common ground). According to our findings, it is likely that these circumstances increase egocentric errors, hindering communication. For this reason, learning about the factors that affect communication under less-than-ideal circumstances can help understand how to better alleviate the load and reduce errors.

This study was not without limitations. Experiment 1 and Experiment 2 were conducted separately, therefore direct comparisons cannot be drawn between the two experiments and future research should examine egocentric errors within the same experiment to compare whether the presence or absence of listener constitutes another factor that affects communication. In addition, the sample size for Experiment 2 was reduced due to the need to form dyads, decreasing the number of final participants.

In conclusion, our results provide a framework that expands current theories of communication by examining communication errors. First, the communication context, such as whether the speaker has available cognitive resources, matters. Second, communication processes tap cognitive resources, but only *tax* these resources under certain circumstances. That is, while communication can seem relatively effortless under ideal circumstances, such as when common ground is established, communication processes are cognitively effortful, and errors are affected by interference from privileged ground due to factors like cognitive load, and individual differences.

Acknowledgements

We thank Mark Thornton and Erin Sherman for providing data collection assistance and Anthony Ryan for designing and creating the apparatus used in the study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Apperly, I. (2010). *Mindreaders: the cognitive basis of" theory of mind"*. Psychology Press.

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, *106*(3), 1093–1108. http://doi.org/10.1016/j.cognition.2007.05.005

Baddeley, A., & Hitch, G. J. (1974). Working Memory. In GA Bower (Ed.) *Recent advances in learning and motivation.* pp. 47–89. New York: Academic.

Branigan HP, Pickering MJ, Cleland AA. 2000. Syntactic co-ordination in dialogue. Cognition 75(2):B13–25

Brown-Schmidt, S., & Horton, W. S. (2014). The influence of partner-specific memory associations on picture naming: A failure to replicate Horton (2007). PloS one, 9(10), e109035.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of educational psychology*, *54*, 1-22.

Clark, H. H. (1992). *Arenas of language use*. Chicago, Ill.: University of Chicago Press.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, I. A. Sag, & B. L. Webber (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge: Cambridge University Press.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, *8*, 331-335. doi: doi.org/10.3758/BF03196169

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*, 769-786. doi: 10.3758/BF03196772

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in cognitive sciences*, *7*(12), 547-552.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Memory and Language, 19(4), 450.

Dunn, O. J. (1961). Multiple comparisons among means. Journal of the American statistical association, 56(293), 52-64.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental science*, *13*(2), 331-338.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*, 309-331.

Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual review of psychology*, *70*, 29-51.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition*, *43*, 226-236. doi: 10.3758/s13421-014-0461-7

Long, J. Scott. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks: Sage Publications, 1997.

Gillund G, Shiffrin RM (1984) A retrieval model for both recognition and recall. *Psychological Review*. 91: 1–67.

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 843.

Grice, H. P. (1957). Meaning. *The philosophical review*, *66*, 377-388. doi: 10.2307/2182440

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

Hamel, R., & Schmittmann, V. D. (2006). The 20-minute version as a predictor of the Raven Advanced Progressive Matrices Test. *Educational and Psychological measurement*, *66*, 1039-1046. doi: 10.1177/0013164406288169

Hawkins, R. X., Gweon, H., & Goodman, N. D. (2018). Speakers account for asymmetries in visual perspective so listeners don't have to. *arXiv preprint arXiv:1807.09000*.

Horton W.S., Gerrig R.J. (2002) Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*. 47: 589–606.

Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. Language and Cognitive Processes, 22(7), 1114-1139.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. Cognition, 96(2), 127-142.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*, 91-117. doi: 10.1016/0010-0277(96)81418-1

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169-183. doi: 10.1037//0096-3445.130.2.169

Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin. 131,* 66-71. doi: 10.1037/0033-2909.131.1.66

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25-41. http://dx.doi.org/10.1016/S0010- 0277(03)00064-7

Kraus, R.M. and Fussell, S.R. (1991). Constructing Shared Communicative Environments. in Resnick, L. Levine, J. & Teasley, S. Perspectives on Socially Shared Cognition. Pittsburgh: American Psychological Association.

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, *47*(7), 567-587.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. Intelligence, 14(4), 389-433.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. Psychonomic bulletin & review, 9(3), 550-557.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of experimental psychology: general*, *141*(2), 302.

Müller, C., Großmann-Hutter, B., Jameson, A., Rummer, R., & Wittig, F. (2001). Recognizing time pressure and cognitive load on the basis of speech: An experimental study. *User Modeling*, 24-33. doi: 10.1007/3-540-44566-8

Oei, N. Y., Everaerd, W. T., Elzinga, B. M., van Well, S., & Bermond, B. (2006). Psychosocial stress impairs working memory at high loads: an association with cortisol levels and memory retrieval. *Stress*, *9*, 133-141. doi: 10.1080/10253890600965773

Patrick Royston (1982). Algorithm AS 181: The *W* test for Normality. *Applied Statistics*, **31**, 176–180. doi: 10.2307/2347986.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-190. doi: 10.1017/S0140525X04000056

Porcelli, A. J., Cruz, D., Wenberg, K., Patterson, M. D., Biswal, B. B., & Rypma, B. (2008). The effects of acute stress on human prefrontal working memory systems. *Physiology & behavior*, *95*, 282-289. doi: 10.1016/j.physbeh.2008.04.027

Qin, S., Hermans, E. J., van Marle, H. J., Luo, J., & Fernández, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological psychiatry*, *66*, 25-32. doi: 10.1016/j.biopsych.2009.03.006

Qureshi, A. (2009). The cognitive basis of mental state reasoning in adults. Unpublished doctoral thesis, University of Birmingham, UK.

Raven, J. C. (1938). *Raven's progressive matrices*. Los Angeles, CA: Western Psychological Services.

Rubio-Fernández, P. (2017). The director task: A test of Theory-of-Mind use or selective attention?. *Psychonomic bulletin & review*, *24*(4), 1121-1128.

Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General, 144*(5), 898-915. http://dx.doi.org/10.1037/xge0000093

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology*, *33*, 643-653. doi: 10.1016/j.psyneuen.2008.02.004

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. P*sychological Review, 114*, 104-132. doi: 10.1037/0033-295X.114.1.104

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. Cognitive psychology, 71, 1-26.

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1466–1481. doi: 10.1037/a0013353

Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic bulletin & review*, *20*, 766-772. https://doi.org/10.3758/s13423-013-0396-1

Table 1.

Descriptive Statistics for Experiments 1, 2, and 3.

Experiment 1. *Descriptive statistics including skewness and kurtosis by load condition (N =54)*

	No Load					**Low Load**					**High Load**			
	N	***M***	***SD***	**Skew**	**K**	***N***	***M***	***SD***	**Skew**	**K**	***N***	***M***	***SD***	**Sk**
Errors	20	3.65	4.64	1.86	3.03	18	3.39	3.04	.60	-1.28	16	10.31	7.39	.77
Time	20	29.44	9.38	.70	-.30	18	30.84	12.72	1.43	1.81	16	30.24	11.76	.1
Descrip	20	34.45	9.59	.98	1.20	18	36.29	10.81	.36	-1.16	16	37.53	12.83	.1

Experiment 2. *Descriptive statistics including skewness and kurtosis by load condition (N = 73)*

	No Load					**Low Load**					**High Load**			
	N	***M***	***SD***	**Skew**	**K**	***N***	***M***	***SD***	**Skew**	**K**	***N***	***M***	***SD***	**Sk**
Errors	23	2.39	2.21	.39	-1.4	25	5.44	5.82	.85	-.70	25	6.24	7.18	1.
Time	23	33.20	20.74	1.65	2.61	25	23.44	10.05	1.14	.72	25	21.89	8.55	.9
Descrip	23	48.24	27.29	1.22	1.10	25	36.06	16.30	1.51	2.80	25	29.35	14.58	.7

Experiment 3. *Descriptive statistics including skewness and kurtosis by load condition (N = 100)*

	No Load					**High Load**						
	N	***M***	***SD***	**Skew**	**K**	***N***	***M***	***SD***	**Skew**	**K**	***t /U***	***d***
Errors	50	7.94	5.65	1.06	1.33	50	14.04	13.72	1.57	1.86	951*	.5
Time	50	39.13	18.75	1.05	.39	50	35.11	19.10	1.82	4.76	-1.06	-.2
Descrip	50	25.19	10.54	.31	-.36	50	18.35	10.59	.94	.94	-3.23**	-.6
WMC	50	.04	.69	-.35	-.02	50	-.04	.93	-.62	-.72	.50	-.0
Gf	50	.13	.91	-.25	-.47	50	-.13	1.08	-.32	-.54	-.25	-1.

Note. Errors = egocentric errors; Time = average time spent describing; Descrip = average number of descriptors; WMC = working memory capacity; Gf = fluid reasoning; Skew = skewness; K = kurtosis; * indicates $p < .05$. ** indicates $p < .01$. *t/U* = *t* statistic from *t*-test and *U* statistic from Mann-Whitney *U* test.

Table 2
Correlations with Confidence Intervals from Experiment 3 (N = 100)

Variable	1	2	3	4
1. Gf				
2. WMC	.45** [.29, .60]			
3. Time	.00 [-.19, .20]	.08 [-.12, .27]		
4. Descriptors	.16 [-.03, .35]	.14 [-.06, .32]	.76** [.66, .83]	
6. Errors	-.37** [-.53, -.18]	-.30** [-.47, -.11]	.30** [.11, .47]	.10 [-.10, .29]

Note. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < .05$. ** indicates $p < .01$. Gf = fluid reasoning; WMC = Working Memory Capacity; Time = time spent describing; Descriptors = total descriptors; Errors = egocentric errors.

Table 3
Summary of Multiple Regression Analysis with Poisson Distribution for Variables Predicting Egocentric Errors (N=100).

	Model 1			Model 2			Model 3		
Variable	**B**	***SE***	***β***	**B**	***SE***	***β***	**B**	***SE***	***β***
Gf	-.24	.03	-7.46***				-.21	.05	-3.92***
WMC	-.16	.04	-4.27***	-.16	.07	-2.25*			
Condition	.47	.06	-.7.32***	.51	.06	7.86***	.45	.07	6.86***
Gf *x* Condition							-.13	.06	-2.01*
WMC x Condition				-.17	.08	-2.14*			
Pseudo-R^2		.26			.20			.24	

Note. *** indicates <.001, ** indicates <.01, * indicates<.05. Gf = Fluid reasoning; WMC = Working Memory Capacity; Condition = high vs. no cognitive load conditions. Reference level for condition is No Load. Pseudo-R^2 was calculated following McFadden's Pseudo-R^2 technique (Long, 1997).

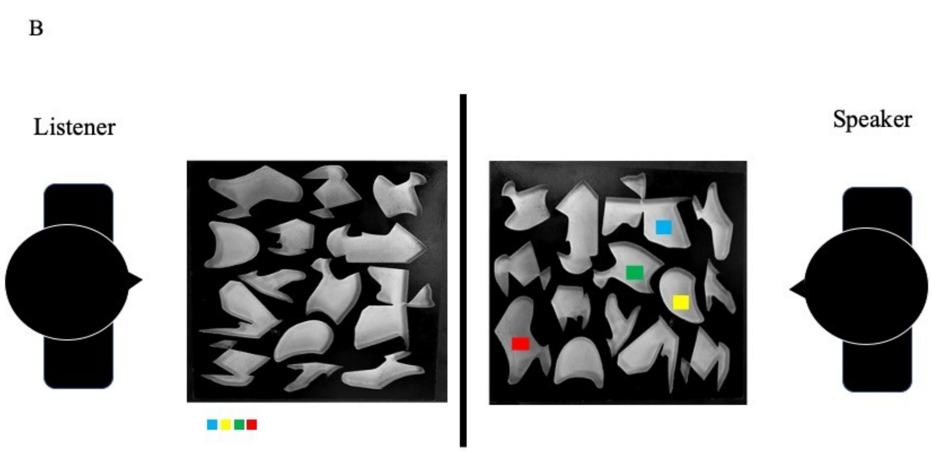

Figure 1. Apparatus set up in Experiments 1 (B), 2 (A), and 3 (A).

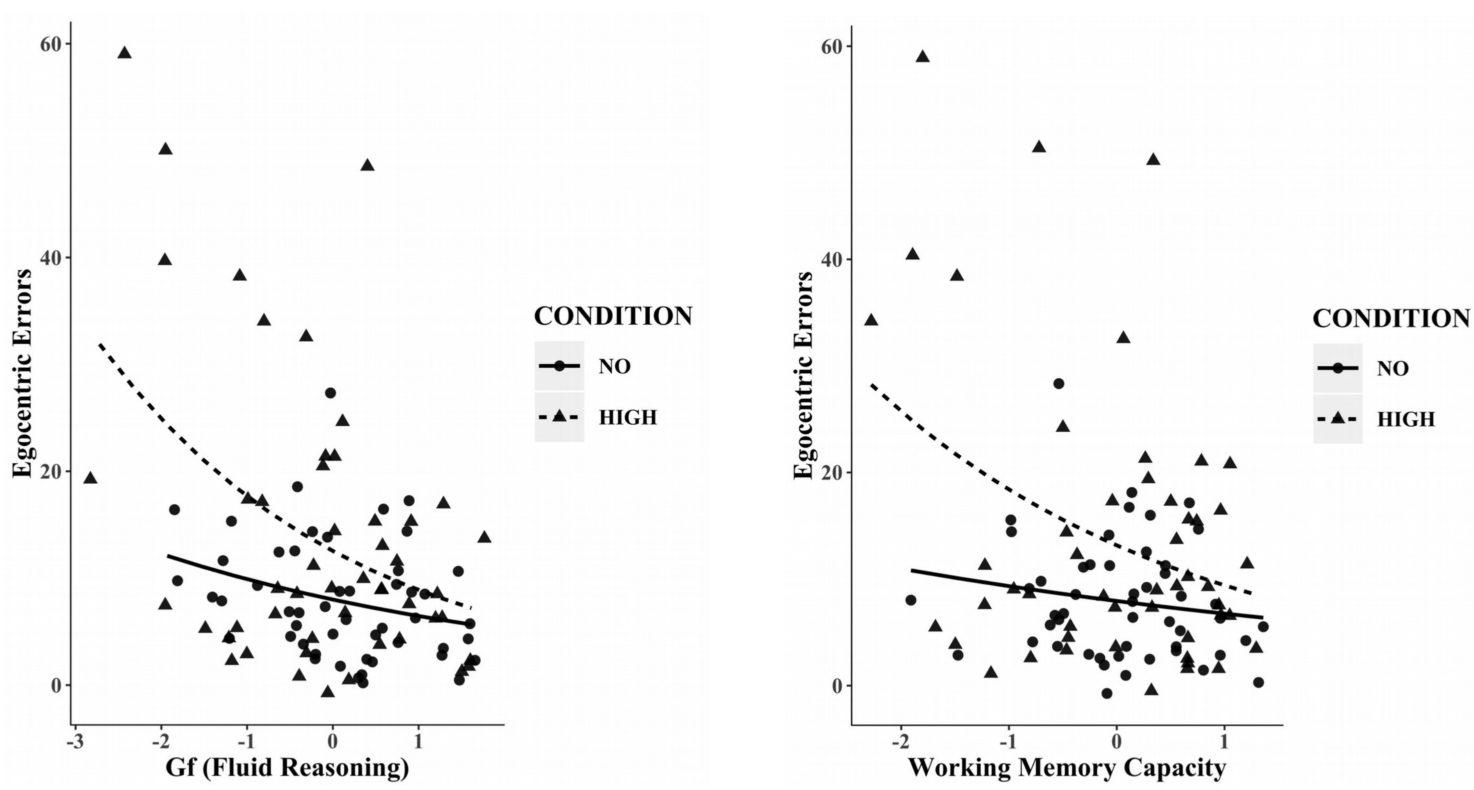

Figure 2. Individual differences in cognitive ability. Egocentric errors as a function of load condition (high vs. no load) for fluid reasoning (left) and WMC (right).