IN PRESS AT *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: GENERAL*

Conveying and Detecting Listening During Live Conversation

Hanne K. Collins[1], Julia A. Minson[2], Ariella Kristal[3], Alison Wood Brooks[1]

[1]Harvard Business School

[2]Harvard Kennedy School

[3]Columbia Business School

All data and materials are available online:

https://osf.io/w4nf9/?view_only=db83cc05213c4964951ab88b1f9e5de0.

Abstract

Across all domains of human social life, positive perceptions of conversational listening (i.e., *feeling heard*) predict well-being, professional success, and interpersonal flourishing. But a fundamental question remains: Are perceptions of listening accurate? Prior research has not empirically tested the extent to which humans can detect others' cognitive engagement (attentiveness) during live conversation. Across five studies (total N = 1,225), using a combination of correlational and experimental methods, we find that perceivers struggle to distinguish between attentive and inattentive conversational listening. Though people's listening fluctuated naturally throughout their conversations (people's minds wandered away from the conversation 24% of the time), they were able to adjust their listening in line with instructions and incentives—by either listening attentively, inattentively, or dividing their attention—and their conversation partners struggled to detect these differences. Specifically, speakers consistently *overestimated* their conversation partners' attentiveness—often believing their partners were listening when they were not. Our results suggest this overestimation is (at least partly) due to the largely indistinguishable behavior of inattentive and attentive listeners. It appears that people can (and do) divide their attention during conversation and successfully feign attentiveness. Overestimating others' attentiveness extended to third-party observers who were not immersed in the conversation, listeners who looked back on their own listening, and people interacting with partners who couldn't hear their words (but were incentivized to act like they could). Our work calls for a re-examination of a fundamental social behavior—listening—and underscores the distinction between *feeling* heard and *being* heard during live conversation.

Key words: listening; conversation; perceptual accuracy; mind perception; interpersonal relations.

Public Significance Statement

This research reveals that there is a difference between *being* and *feeling* heard. People's listening fluctuated naturally throughout their conversations (with mind wandering reported 24% of the time)—and their conversation partners were often unable to detect the rise and fall of their partner's attentiveness (whether via natural fluctuation or via our experimental interventions). Across a diverse set of studies, we find support for three key results: (1) perceptions of conversational listening often do not align with listeners' cognitive attentiveness; (2) perceptions of listening are often inaccurate due to a lack of diagnostic behavioral cues displayed by listeners—inattentive listeners behave similarly to attentive listeners; and consequently (3) perceivers primarily *overestimate* the extent to which their conversation partners are listening to them. These results emphasize the importance of recognizing that moments of inattentiveness happen in conversation—encouraging conversants to acknowledge, forgive, and repair these moments to achieve relational and informational success.

Conveying and Detecting Listening During Live Conversation

"The single biggest problem with communication is the illusion that it has taken place."

–George Bernard Shaw

You're in a coffee shop, meeting with a new acquaintance over steaming lattes. While you're talking, your conversation partner seems engaged: they hold your gaze, smile at the funny parts, and nod warmly. You think they're a good listener, and you're excited to see them again soon. But were they *really* listening? If you probed their mind during the conversation, what were they actually attending to and thinking about? Were they really listening attentively, or just creating the impression of good listening?

From business to medicine to romance, being seen as a "good listener" is widely advised and highly desired. And for good reason—perceptions of listening are associated with many beneficial outcomes. In the workplace, employees who feel that their supervisor listens to them report lower emotional exhaustion, lower turnover intentions, greater internal motivation (Lloyd et al., 2015) and seek more feedback (Qian et al., 2019). In romantic relationships, signals of listening are associated with an improved ability to respond to and cope with stressors as well as overall relationship satisfaction (Bodenmann, 2005; Kuhn et al., 2018). In healthcare, patients who feel that their healthcare provider listens to them show higher levels of medication adherence (Shafran-Tikva & Kluger, 2016) and are more satisfied with their care during a hospital stay (Wanzer et al., 2004). Perceptions of listening even play a critical role during first encounters among strangers: perceptions of call center employees as good (or bad) listeners drive customer satisfaction ratings (De Ruyter & Wetzels, 2000; Min et al., 2021), and responsive strangers are better liked in get-to-know-you conversations—they even receive more second-date offers on first dates (Huang et al., 2017). However, though perceptions of listening are consequential across a wide range of domains, research has not examined the extent to which perceptions of listening are *accurate*.

Previous Research on Listening – A Conundrum for Determining Accuracy

The psychological process of conversational listening is much more complex than simply hearing sounds (i.e., auditory processing). To capture the cognitive experience of listening, and account for the temporally unfolding nature of conversation, which requires participants to listen and respond dynamically and recursively, we define listening as: *attending to and processing another person's verbal, nonverbal, and paralinguistic cues amidst conversation.* This definition reflects the separable cognitive processes of *attending to* and *processing* content from a conversation partner (Collins, 2022), including their verbal cues (e.g., words, grammar, syntax), nonverbal cues perceived visually (e.g., facial expressions, body language, hand gesticulation), and paralinguistic (i.e., prosodic) cues perceived auditorily (e.g., pauses, interruptions, back-channel utterances like "mmhm," "yea," laughter, tone, accent, and volume of voice; Yeomans et al., 2021).

In conversation, the private cognition involved in attending to and processing a partners' cues is happening in a profoundly interpersonal, co-constructed system—in tandem with at least one other human mind. Seminal work in psycholinguistics and conversation analysis suggests that "language use is fundamentally a joint activity" (p. 244, Clark 1994; Clark & Schaefer, 1989; Clak & Wilkes-Gibbs, 1986; Schegloff et al., 2977): while speaking and listening are individual actions, they contribute toward joint "problems"—errors in achieving shared understanding—and dealing with these "problems" requires joint management, strategies that can help both prevent and repair glitches in shared understanding (Clark, 1994). Taken together, individuals use different verbal, nonverbal, and paralinguistic cues to coordinate their shared understanding on a turn-by-turn basis, and doing so is a uniquely human ability (Dingemanse et al., 2021).

Importantly, though, this coordination process is fraught. While the private act of listening may be *signaled* through responsive verbal cues (e.g., words of affirmation, paraphrasing, follow-up questions), nonverbal cues (e.g., nodding, eye gaze, facial expression), and paralinguistic cues (e.g., back-channels like “uh huh,” silence, laughter), those cues may (or may not) represent the underlying cognitive process of listening (Collins, 2022). Furthermore, those cues may (or may not) be accurately perceived. Janusik (2007) describes this conundrum aptly: “Listening research is a challenge, as listening is performed cognitively and perceived behaviorally, but listening cognitions and behaviors are not always congruent (Witkin, 1990).” Accordingly, prior psychological research on listening has focused on the *intrapersonal* (cognitive) experience of listening (and its consequences) and, separately, on the *interpersonal* perception of listening (and its consequences).

Early listening scholars attempted to measure listening using hearing, comprehension, and recall measures. However, more recent scientific consensus has concluded that these measures of listening cannot meaningfully distinguish listening from memory capacity (Thomas & Levine, 1994)—if one is attentively listening in the moment, they may not necessarily remember the content later. Thus, researchers to date have not uncovered a robust measure to capture listening as a unique cognitive activity, distinct from other related cognitive processes.

Separately, prior interpersonal models of listening describe how people *perceive* others’ listening. This work has largely relied on two types of measures: 1) the listener’s behavioral cues, such as eye contact and nodding, which researchers have assumed are indicative of whether or not someone is paying attention (based on lay beliefs about what good listening looks like), and 2) their partners' self-reported perceptions of listening (e.g., "How well do you think your partner listened?"; Itzchakov et al., 2016; Wanzer et al., 2004).

Much of the attention in this area has focused on "active listening," a construct initially advocated by humanistic psychologist Carl Rogers (e.g., 1954, 1957) and considered desirable by many fields (Cheon & Grant, 2009; Hafen & Crane, 2003; Kubota et al., 2004; McNaughton et al., 2008; Mishima et al., 2000; Rautalinko & Lisper, 2004). Active listening embraces the benefits of listening, as well as the importance of communicating that one is doing so (i.e., ensuring that listening is *perceived*). According to this literature, by engaging in behaviors that people associate with listening (see Bodie et al., 2012), the listener can signal to their partner that they are, which will improve the interaction and, ultimately, the relationship. Indeed, active listeners are better liked, and people find interacting with them more satisfying (e.g., Weger et al., 2010; Weger et al., 2014). However, research on active listening has focused on *perceptions* of listening, usually by surveying the perceiver. Thus, it remains unclear whether high-performing active listeners also perform the cognitive work of listening well, or whether they are simply better at conveying this impression.

The current research seeks to address this puzzle by manipulating people's ability to listen during live conversation, as well as employing measures of actual listening, to triangulate on an understanding of the congruence between people's *cognitive* experience of listening and their partners' *perceptions* of it. Given the mismatch between how listening is performed (cognitively), conveyed (behaviorally), and perceived (interpersonally), we predict that perceptions of listening may not always align with listeners' private cognitions:

Hypothesis 1: Perceptions of conversational listening are often inaccurate (i.e., perceiver ratings of listening do not align with target self-reported listening).

A Lens Model Approach

Why might people struggle to perceive others' conversational listening? The classic Brunswik lens model offers a framework to understand the process (and potential failures) of interpersonal perception (Brunswik, 1956; Karelaia & Hogarth, 2008; Nestler & Back, 2013). According to the lens model, perceptual accuracy requires three stages: (1) a latent trait (in our case, "listening") is expressed through observable cues by a target, (2) perceivers must attend to these observable cues, and then (3) perceivers must use this information to inform their perceptions of the target (in our case, their listening). If there is a breakdown at any of these stages—target expression, perceiver attention, or perceiver judgment—then perceptual accuracy will suffer.

The lens framework has been applied to many forms of interpersonal perception. For example, research on lie detection finds that people are, on average, no better than chance at detecting when a person is lying versus telling the truth (see reviews by DePaulo et al., 1985; Kraut, 1980; Zuckerman et al., 1981). One meta-analysis found an average accuracy rate of 54%—and found no difference in performance between lay people and experts such as law enforcement personnel (Bond & DePaulo, 2006). Following from the lens model, two explanations have been offered for this finding: (1) a lack of valid observable cues that reveal deception, and/or (2) the idea that people rely on the wrong behavioral cues to form their judgments. Research suggests that the former explanation plays a larger role—liars behave similarly to truth tellers, resulting in a dearth of valid cues that reveal deception (Hartwig & Bond, 2011; Wiseman et al., 2012).

Importantly, "valid cues" have a particular meaning when people are motivated to mislead their counterparts. In order to accurately diagnose deception, observers must seek cues that cannot be faked. We expect that the same is true for perceptions of listening. That is, the social value of being perceived as a “good listener” means that people are highly motivated to feign listening even in instances of inattentiveness—and are likely well practiced in doing so. Just as liars seek to conceal cues of their deception, listeners are likely to conceal cues of their inattentive listening. Thus, it may be the case that just as with lying, there are few *valid* cues of good listening that exist—cues that cannot be feigned when one is motivated to do so (Collins, 2022), leading to our second prediction:

> *Hypothesis 2:* Inaccuracies in perceptions of conversational listening are (at least partly) due to a lack of diagnostic behavioral cues conveyed by the listener (i.e., the cues that perceivers focus on are successfully feigned by motivated actors).

Overly Optimistic Perceptions of Listening

We consider two types of perceptual errors in listening detection, sometimes referred to as Type I and Type II errors. First, what happens when people believe that their conversation partners are listening when they are not (Type I error)? For example, imagine you believe a work colleague is listening to you in a meeting, but they are in fact, mentally preparing their grocery list or ruminating about a different project. While you might leave with the impression that you effectively communicated important information, your colleague may be just as uninformed as before the meeting began.

However, the opposite perceptual error might also occur. What happens when people believe that their conversation partners aren’t listening when they actually are (Type II error)? From marriage counselors to customer service employees to international mediators, conflict resolution professionals are frequently tasked with repairing the relational damage resulting from one party accusing the other of failing to listen (e.g., Agne, 2018).

Type I errors (someone perceiving that you are listening when you aren't) may prevent attainment of information exchange goals. However, these errors do allow people to "feel heard," an emotional state known to be extremely positive. Meanwhile, Type II errors (someone perceiving that you aren't listening when you are) likely result in relational costs such as negative emotional reactions and accusations of inattentiveness—though they might make information transmission more likely. Overall, it is hard to know which of these errors is more costly. The costs and benefits of each type is likely to vary across contexts. To make specific predictions about which error is more common in the case of conversational listening, we turn to the theory underlying "want-should conflicts," common situations in which people make choices between behaviors that are pleasant in the moment (e.g., eating ice cream) versus beneficial in the long-run (e.g., eating broccoli; see Bitterly et al., 2015 for a review).

In the case of conversation (as in other want-should conflicts), the "want-self" is myopic and prioritizes instant gratification (i.e., being perceived as a good listener in this moment) while the "should-self" prioritizes long-term benefits (i.e., ensuring that one leaves a conversation well-informed, even at the cost of some immediate awkwardness). Amid conversation—as in battles between ice cream and broccoli—we hypothesize that the "want-self" will win out. In moments of inattentiveness, individuals will mask their wandering mind to maintain the impression that they are listening. This will be especially true when relational goals are or seem more important than informational goals (Yeomans et al., 2021). Just as liars want to be perceived as truthful, listeners may be motivated to cover up moments of inattention to make a good impression—because, as an abundance of research has demonstrated, being perceived as a good listener confers benefits. Of course, some conversations require a greater emphasis on informational goals (e.g., a doctor explaining how to properly take a medication), which may encourage the "should-self" to allow more Type II errors—or to use more preventative and repair strategies to reveal glitches in shared understanding (Clark, 1994).

This interpersonal want-should conflict as experienced by the target, is largely in line with the one being experienced by a perceiver. Calling out a counterpart for inattentiveness when they are in fact listening—a false accusation—carries an immediate relational penalty. In contrast, giving the counterpart the benefit of the doubt when they are not attending to your words may or may not lead to costs down the road. Thus, prior theorizing suggests that both the target and the perceiver have reason to downplay the occurrence of cues that suggest attentional lapses. This analysis also aligns with error management theory (EMT; Haselton & Buss, 2000), which argues that people are systematically more likely to make a particular error when the costs and benefits of Type I versus Type II errors are asymmetric.

Taken together, we predict that the social desirability of being perceived as a good listener will drive listeners to mask inattentiveness in conversations to serve immediate relational goals (such as those highly salient in our studies). As a result, conversationalists may often be left with the impression that information exchange and social connection occurred, even when those impressions are erroneous (Yeomans et al., 2021). This leads to our third prediction:

Hypothesis 3: Perceptions of conversational listening are over-optimistic (i.e., perceivers believe targets are listening more than they are).

The Current Work

Across five studies, we investigate the extent to which people convey and detect listening in conversation. Our core prediction is that speakers frequently make errors in their perceptions of whether their counterpart is listening to them. We theorize that these errors occur because 1) people's minds wander more often than we realize amidst live conversation, and yet 2) people are motivated to come across as good listeners, and 3) there are few behavioral cues that allow attentive versus inattentive listening to be detected.

Previous investigations of interpersonal listening have focused on a sense of feeling heard or understood—which incorporate concepts like empathy, care, and support into their measures, for example, by surveying the perceiver with scale items like "X understands how I feel" (Bodie, 2011) or "X cares about me" (Lloyd et al., 2015). We advance this work by simultaneously examining listeners' cognitive engagement during live conversation (via recalled and contemporaneous self-report, as well as a variety of experimental manipulations), allowing us to measure whether they are actually listening, and comparing these measures to whether the perceiver *feels* heard, understood, or validated (during, and after, the conversation).

In an exploratory study, we investigate this phenomenon by assessing the relationship between ratings of self-reported and perceived listening during a live conversation. Then, in a series of experiments (Experiments 1-4), we ask dyads to engage in a live conversation, with individuals later asked to assess the extent to which a target was listening attentively. Using a variety of paradigms, we manipulate actual listening between experimental conditions by introducing distractions, adjusting incentives, or using technology to limit participants' physical ability to hear their partner—and then measure interpersonal perceptions of listening.

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study (as suggested by Simmons, Nelson & Simonsohn,

2011). All data and materials are available here:
https://osf.io/w4nf9/?view_only=db83cc05213c4964951ab88b1f9e5de0.

Exploratory Study

In an exploratory study, we sought to investigate our phenomenon of interest: the association between perceived and actual listening during live conversation. In our study design and analysis, we follow in the tradition of interpersonal perception research, which captures perceptual accuracy as the correlation of perceivers' inferences and targets' self-ratings (e.g., Back & Nestler, 2016; Brunswik, 1952, 1956; Zaki et al., 2009). We sought to investigate this phenomenon as directly as possible during live conversation. Thus, rather than asking participants to recall their own and/or their partner's listening after the conversation, we collected a series of moment-to-moment assessments during conversation, following an approach used in mind-wandering research (Smallwood & Schooler, 2006). All procedures and analyses were pre-registered: https://aspredicted.org/blind.php?x=LB5_3GQ.

Exploratory Study Method

Participants

We recruited 141 dyads (made up of two people who were previously unacquainted) from the participant pool at a university in the northeastern United States to take part in a 45-minute study about conversation for which they were paid $18. Our sample consisted of university students, staff, and members of the local community. As per our pre-registration, data from 41 dyads were excluded from analyses.[1] We analyzed data from the remaining 100 dyads (N = 200 participants). Participants reported their gender ("What is your gender? Male, Female,

[1] 27 dyads had technological problems (e.g., poor audio/video quality or dropping off the call completely); 13 dyads failed to follow instructions (e.g., not responding to the private chat messages from the experimenter, or not engaging in the full 25-minute conversation); 1 dyad did not complete the post-conversation survey

Non-binary/Other"; 32% Male, 65% Female, 3% Non-Binary/Other) and their age ("What is your age?" [Open-ended numeric response]; M_{age} = 28 years, SD_{age} = 12 years).

Procedure

This study took place over the Zoom video conferencing platform. All participants were randomly assigned to role of "target" or "perceiver" and paired with a partner assigned to the opposite role for a 25 min conversation. When participants logged onto the online platform, they were randomly assigned by the experimenter to take on the role of 'target' or 'perceiver,' and read detailed instructions about the study procedures. Participants were told they would have a 25-minute "get-to-know-you" conversation. To make the conversation easier, we provided participants with a list of five topics they could potentially discuss (i.e., favorite food, hobby, book or TV show, place to visit, and animal).

We told participants that at various times throughout the conversation, the researcher would send them a private message asking them to report their own (in the 'target' condition) or their partner's (in the 'perceiver' condition) listening at that moment. Specifically, every 5 minutes, the participants were asked to respond to the question: "Think about the last time [your partner was/you were] talking, right before you received this message. In that moment, [were you/was your partner] attentively listening to [your partner/you]?" (1: *Yes, [I/my partner] was fully attentive*, 2: *[My/my partner's] mind was wandering*). Finally, to encourage honest responding, and following the mind-wandering literature (Smallwood & Schooler, 2006), we told all participants that natural fluctuations in attentive listening are normal and commonly occur during conversation. Thus, for each dyad, we obtained five yoked measurements of one participant's self-reported listening and their conversation partner's contemporaneous perception of their listening.

Exploratory Study Results

Targets (who were asked about their own listening) reported listening attentively during 76% of all measurement occurrences, and reported mind wandering during 24% of all measurement occurrences. To the best of our knowledge, this is the first-ever measure of the extent to which listeners' minds wander during live conversation. Our primary measure of interest was the perceivers' detection of their partner's attentiveness versus inattentiveness. To that end, 69% of perceiver guesses of (attentive versus inattentive) listening matched targets' self-reports of listening (attentive versus inattentive). Perceptions aligned 64% of the time when the partner reported listening, and 5% of the time when the partner reported *not* listening. On the other hand, approximately one third of perceivers' guesses (31%) did not align with targets' self-reports. Perceptions diverged from self-reports 19% of the time when the partner reported attentive listening, and 12% of the time when the partner reported *not* listening.

Signal Detection of Attentive Listening

To investigate the direction of this misalignment, we drew on signal detection theory, treating target self-reports of listening as the signal (see Table 1; Stanislaw & Todorov, 1999). Here we see that when the signal was present (i.e., the target reported attentive listening), perceivers recognized it as such in 84% of occurrences. However, when the stimulus was absent (i.e., the target reported not listening), perceivers incorrectly believed their partner was listening in 78% of occurrences (Type I error). These results suggest a strong bias toward believing that the target was listening, regardless of whether they actually were (or were not). We summarize these signal detection results in Table 1 and depict Type I errors across the 5-min measurement increments in Figure 1.

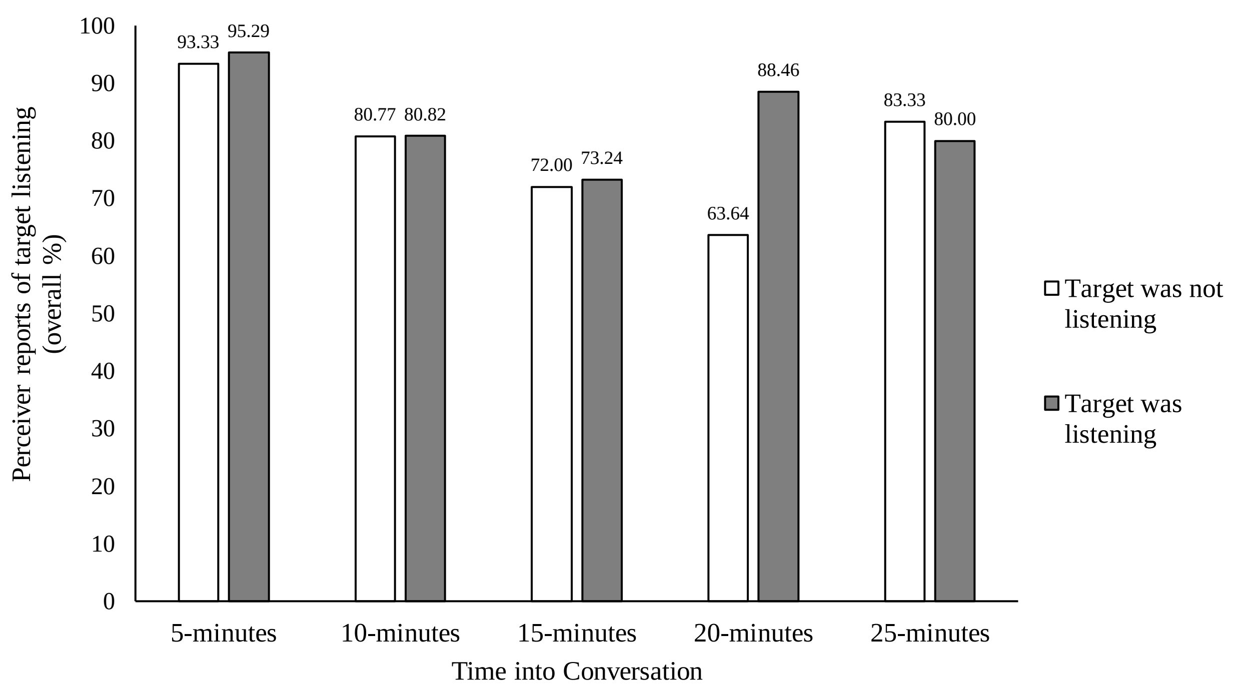
Table 1

Signal Detection Analysis of Exploratory Study Results.

	Perceiver thought target was not listening	**Perceiver thought target was listening**
Target reported not listening	22% CORRECT	78% TYPE I ERROR
Target reported listening	16% TYPE II ERROR	84% CORRECT

Figure 1

Comparison of Perceived Listening and Target Experienced Listening Throughout Conversation.



Note: This figure depicts the percentage of perceivers who reported that their partner was attentively listening at each timepoint. Perfect accuracy would be depicted by all white bars at 100% and all gray bars at 0%. Instead, with an exception at the 20-minute mark, the percentage of perceivers who reported that their partner was listening is nearly identical when their partners self-reported listening attentively (gray bars) and when their partners self-reported not listening attentively (white bars).

Perceptions vs. Experience of Listening During Conversation

Moving beyond signal detection, per our pre-registered analysis plan, we analyzed the extent to which targets' self-reported listening during the conversation corresponded to perceivers' ratings of their listening on average. When we compared the average number of times participants self-reported listening throughout the conversation (i.e., the mean for *targets*) to the average number of times participants perceived their partner to be listening throughout the conversation (i.e., the mean for *perceivers*) using a paired samples t-test, we found that perceived listening (M = 4.14 out of 5, SD = 1.02) was significantly higher than self-reported listening (M = 3.82 out of 5, SD = 1.06), $t(95) = -2.05$, $p = .04$, *Cohen's d* = 0.30, *95% CI* [0.01, 0.59].[2]

Exploratory Study Discussion

Our initial study allowed us to explore our phenomenon of interest—the extent to which perceptions of listening align with listeners' self-reports during a live conversation. Results provide preliminary evidence for two insights: (1) in approximately one-third of instances, perceptions of conversational listening did not match target self-reports (Hypothesis 1), and (2) perceivers largely over-attributed attentive listening, often believing their conversation partners were listening to them when they were not (Hypothesis 3).

The results of this exploratory study reveal that perceptions of listening in natural conversation do not always match reality. Instead, people's minds seem to naturally wander

[2] Five dyads were dropped due to a single missing observation. We repeated this analysis predicting listening rating from a fixed-effect for condition (target vs. perceiver), and a random-effect for dyad to account for repeated observations (Bates et al., 2015). This allowed us to drop only missing responses instead of entire dyads. Results confirmed that perceiver rated listening was higher than target self-reported listening, $b = 0.07$, $p = .009$, *95%CI*[0.02, 0.11]. Additional analyses are included in the Supplemental Material.

away from the conversation without their counterparts noticing. However, this study was correlational and utilized a binary response scale that may have biased our results—there are many attentive states in between "fully attentive" and "mind wandering" that are not captured here. Furthermore, it is impossible to establish "ground truth": were perceivers over-attributing listening or were targets under-reporting it? Thus, in Experiment 1, we sought to examine this phenomenon further, by experimentally inducing variation in participants' motivation to listen (by using distraction and financial incentives) at various levels of attentiveness.

Experiment 1

In Experiment 1, we investigate people's tendency to convey and detect different levels of listening in live conversation by experimentally manipulating the listener's motivation to attend to the conversation. In order to induce variation in listening, we instructed one participant in each dyad to pay careful attention to their partner (*Listening* condition), direct their attention elsewhere (*Distracted* condition), or direct their attention elsewhere while *pretending* to listen to their partner (*Feigned Listening* condition), all during live conversation.

Our three-condition design allowed us to investigate whether participants who are explicitly incentivized to feign listening are perceived differently than those who are simply asked to direct their attention elsewhere. Though a large body of research establishes listening as a desirable social behavior, our exploratory study suggests that conversation partners frequently suffer from moments of inattention, even in short conversations with relatively few distractions. Yet, social desirability may motivate people to maintain the appearance of consistent listening. We thus predict that perceivers will struggle to detect listening differences across these three conditions—attentive and inattentive listeners alike will be given credit for listening attentively.

Experiment 1 Method

Participants

We recruited 162 pairs of strangers from a participant pool at a university in the northeastern United States consisting of university students, staff, and members of the local community. Dyads came into the lab to participate in a 10-minute study on everyday conversations for which they were paid a flat rate of $15, with potential to earn up to an additional $5 in bonus payments. Data from 11 dyads were excluded from analysis,[3] and thus our analyses are based on the remaining 151 dyads (N = 302). Participants reported their gender ("What is your gender? Male, Female, Non-Binary"; 34% Male, 55% Female, 1% Non-Binary, 10% non-response) and their age ("What is your age" [Open-ended numeric response]; M_{age} = 23 years, SD_{age} = 5 years).

Procedure

When they arrived at the lab, participants were paired with a partner they had not met before and were told that they would spend five minutes in conversation. Specifically, we instructed dyad members to get to know each other and determine whether or not they would make good roommates. We also told participants that there would be a series of videos playing on a screen in the room. The instructions regarding these videos varied by experimental condition as follows.

Listening Manipulation. Within each dyad, one participant sat with their back to the video screen and was instructed to ignore the videos playing behind them. This was the "unmanipulated partner." The other participant (the "manipulated partner") was seated in full view of the video screen and was randomly assigned to one of three conditions: (1) *Listening* condition (n = 50), (2) *Distracted* condition (n = 49), or (3) *Feigned Listening* condition (n = 52).

---

[3] 2 dyads knew each other; 2 dyads did not complete the questionnaires; 4 dyads experienced a technical malfunction; 3 dyads received a questionnaire that did not match their condition assignment.

The unmanipulated partner did not know the manipulated partner's private listening instructions. This approach follows the methods of recent conversation research, whereby one conversationalist enacts an experimental intervention unbeknownst to their partner, and researchers observe their partner's natural reactions (e.g., Huang et al., 2017; Yeomans et al., 2020). The manipulated participants were video recorded throughout the interaction.

We instructed participants in the *Listening* condition to ignore the videos and listen attentively to their partner. Participants learned that they could earn a bonus based on how well they remembered what their partner said.

Participants in the *Distracted* condition were instructed to pay attention to the video playing on the screen behind their partner. They learned that the video consisted of a series of muted commercials and that they could earn a bonus for each commercial that they recalled.

Participants in the *Feigned Listening* condition were instructed to pay attention to the commercials playing on the screen while *pretending* to listen attentively to their partner. These participants were offered a bonus for each commercial they recalled *only if* their partner reported thinking they had been listening throughout the conversation (we did not provide specific criteria for how the partner would make this assessment).

After both participants in a dyad read their instructions independently, they were brought into the conversation room and seated face-to-face.

Measures

Our primary measure of interest was whether the listening manipulation influenced the impressions that the unmanipulated participants formed of their manipulated partners. To this end, we asked the unmanipulated participants to evaluate the quality of their partner's listening, the conversation overall, as well as their partner on several dimensions.

Perceptions of Listening. Unmanipulated participants reported the extent to which they thought their partner was "a good listener," was "interested in what I had to say," and was "engaged in this conversation" on a scale from 1: "Extremely disagree" to 6: "Extremely agree." Responses to these three items were averaged to create an overall measure of the extent to which participants thought their partner was listening ($\alpha = 0.91$).

Enjoyment. Unmanipulated participants also reported the extent to which they agreed (1: "Extremely disagree" to 6: "Extremely agree") with five statements assessing their enjoyment of their partner as well as their conversation (e.g., "I liked my partner" and "I found the conversation with my partner interesting"; $\alpha = 0.91$).

Warmth, Competence, and Status. Unmanipulated participants also rated their partner's warmth (4-items; e.g., "I think my partner is tolerant"; $\alpha = 0.78$; Fiske et al., 2002), competence (5-items; e.g., "I think my partner is independent"; $\alpha = 0.64$; Fiske et al., 2002), and power (3-items; e.g., "I think my partner is dominant"; $\alpha = 0.78$; Smith et al., 2008) on a scale from 1: "Extremely disagree" to 6: "Extremely agree." Warmth, competence, and status are dimensions on which individuals readily evaluate each other and which have important effects for subsequent interactions (Fiske et al., 2007). We were interested in testing whether perceptions of listening, even inaccurate ones, would affect such interpersonal inferences, further supporting our argument that good listening is a highly desired, but often misperceived, behavior. Importantly, both participants were told that their interpersonal ratings would remain private to mitigate potential impression management concerns.

Manipulation Check. To assess whether participants followed our instructions, manipulated participants were asked to recall as many of the commercials as they could. We tallied the number of commercials participants correctly recalled (out of nine). Manipulated participants also predicted the extent to which their partner would report that they were a good listener on a scale from 1: "Extremely disagree" to 6: "Extremely agree."

In an attempt to more directly measure the extent to which participants listened to their partner using a recall measure, unmanipulated participants completed the Activities Preferences Questionnaire (APQ; Surra & Longstreth, 1990; Swann & Gill, 1997) prior to the conversation, and manipulated participants predicted their responses to the questionnaire after the conversation. In reviewing the conversation transcripts, the vast majority of the conversation pairs did not discuss the activities included in the APQ. Because of this, we do not include this measure in our main analyses.

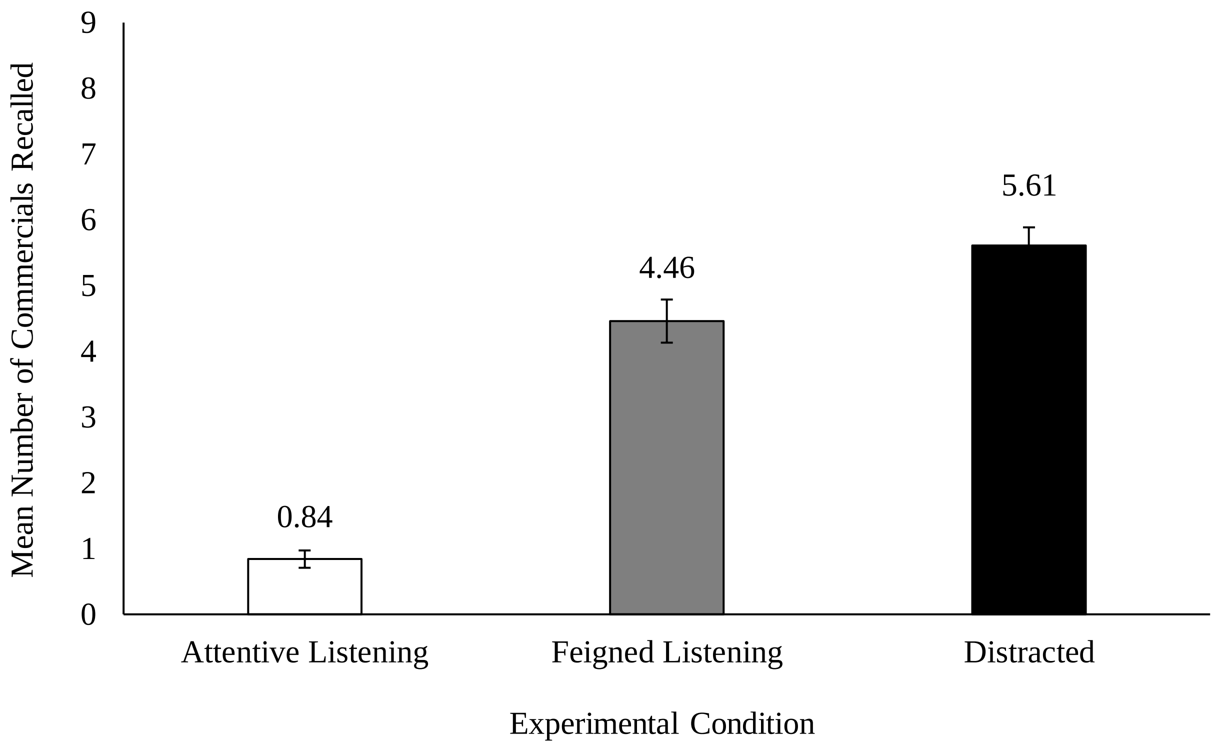
Experiment 1 Results

Self-Perceptions

Our manipulation successfully shifted participants' attention toward the videos playing in the room, as evidenced by the number of commercials the manipulated participants recalled, $F(2, 148) = 90.11$, $p < .001$. Participants in the *Distracted* condition correctly recalled the greatest number of commercials ($M = 5.61$, $SD = 1.93$). Conducting Tukey's test for post-hoc comparisons, participants in the *Feigned Listening* condition recalled significantly fewer commercials than those in the *Distracted* condition ($M = 4.46$, $SD = 2.36$; $b = -1.15$, *95% CI* [-2.02, -0.28], $SE = 0.37$, $p = .006$), but recalled significantly more than those in the *Listening* condition ($M = 0.84$, $SD = 0.93$; $b = 3.62$, *95% CI* [2.75, 4.49], $SE = 0.37$, $p < .001$; see Figure 2). These results suggest that participants in the *Distracted* and *Feigned Listening* conditions were indeed devoting a substantial amount of attention to a stimulus other than their partner.

Figure 2

Mean Number of Commercials Recalled Across Conditions.



Note: Mean number of commercials correctly recalled by participants across conditions (Experiment 1), with error bars representing standard errors.

Furthermore, manipulated participants believed that the quality of their listening would be easily detected, $F(2, 148) = 5.79$, $p = .004$. Again, using Tukey's test for post-hoc comparisons, participants in the *Distracted* condition thought that their partner would rate them as a significantly worse listener ($M = 4.78$, $SD = 1.18$) than participants in the *Listening* ($M = 5.32$, $SD = 0.77$; $b = -0.54$, *95%CI* [-0.96, -0.13], $SE = 0.17$, $p = 0.006$) and *Feigned Listening* ($M = 5.25$, $SD = 0.56$; $b = -0.47$, *95% CI*[-0.88, -0.07], $SE = 0.17$, $p = 0.02$) conditions. The two latter conditions did not differ from each other ($b = -0.07$, *95%CI* [-0.48, 0.34], $SE = 0.17$, $p =$

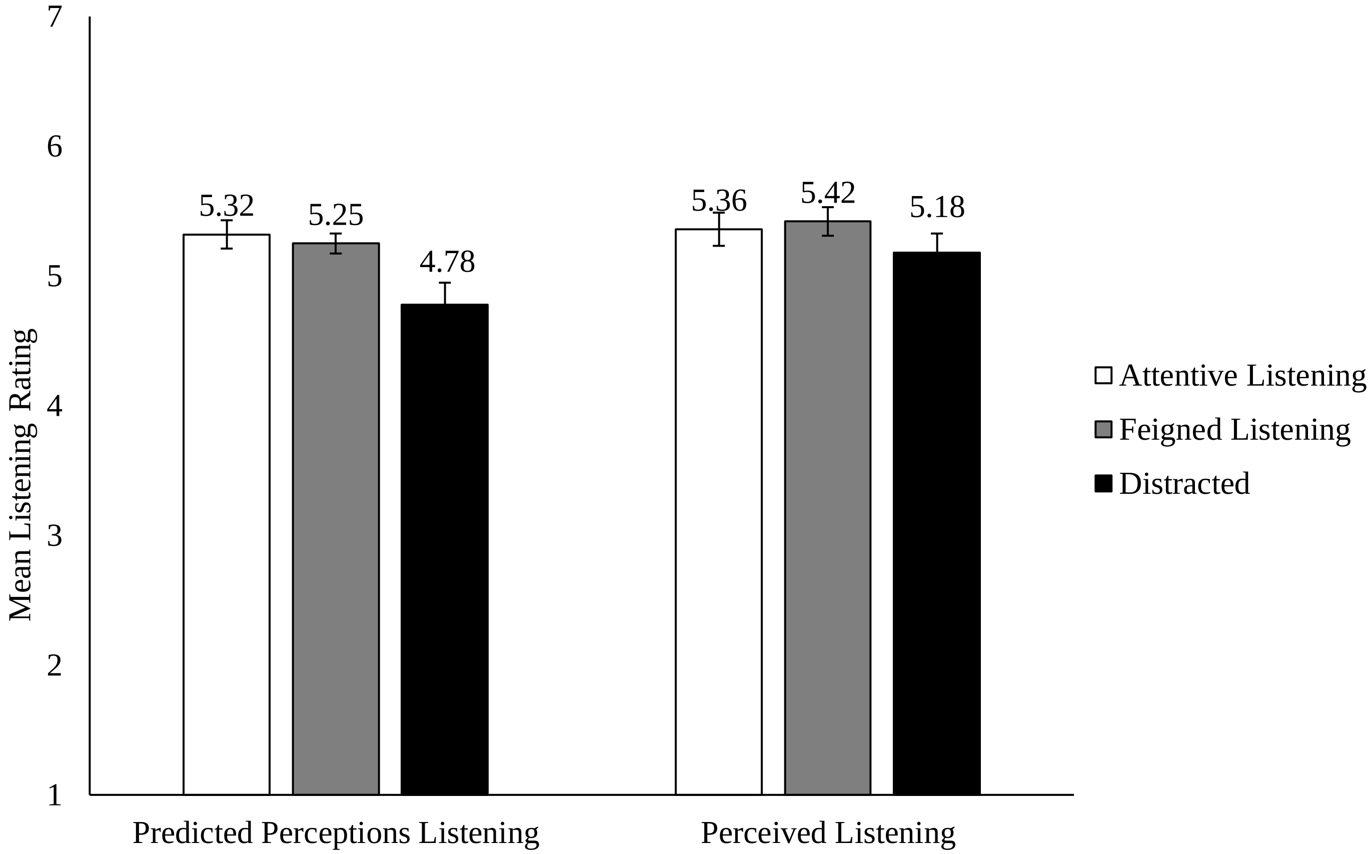
0.91), suggesting that participants in the *Feigned Listening* condition felt confident in their ability to convince their partner that they were listening attentively (see Figure 3).

Partner Perceptions

We find no differences between the three conditions in the partners' evaluations of listening quality ($M_{Listening}$ = 5.36, $SD_{Listening}$ = 0.91; $M_{Distracted}$ = 5.18, $SD_{Distracted}$ = 1.01; $M_{Feigned\ Listening}$ = 5.42, $SD_{Feigned\ Listening}$ = 0.82; $F(2, 147) = 0.84$, $p = .43$; Listening vs. Distracted: b = -0.18, *95% CI* [-0.61, 0.26], *SE* = 0.19, p = 0.61; Listening vs. Feigned Listening: b = 0.06, *95% CI* [-0.37, 0.48], *SE* = 0.18, p = 0.95; Distracted vs. Feigned Listening: b = 0.23, *95% CI* [-0.20, 0.67], *SE* = 0.18, p = 0.42). Participants who conversed with a partner who was secretly memorizing commercials rated their partner's listening quality similarly to those interacting with a fully attentive partner (see Figure 3).

Figure 3

Perceived Listening Ratings (Predicted and Actual) Across Conditions.



Note: "Perceived Listening" bars show ratings of listening by the unmanipulated partners, while "Predicted Perceptions of Listening" bars show how manipulated partners believed they would be rated by their naïve partners (Experiment 1). Though participants in the distracted condition predicted that their partner would rate them as a poorer listener, there were no differences in perceived listening ratings across conditions. Error bars represent standard errors.

Similarly, there were no significant differences between the three conditions in ratings of warmth, competence, or status (see Table 2). However, conversation partners in the *Feigned Listening* condition reported marginally greater enjoyment of the conversation than partners in the *Listening* condition (with no differences between the *Distracted* condition and the *Listening* or *Feigned Listening* conditions; see Table 2). One possible explanation for this result is that listeners incentivized to "fake it" may have been more emphatic in their attempts to convey their listening.

Table 2

Evaluations of Listener Warmth, Competence, Power, and Conversational Enjoyment Across Conditions.

	Omnibus F-Value	***p***	**Attentive Listening *M (SE)***	**Feigned Listening *M (SE)***	**Distracted *M (SE)***
Warmth	0.37	.69	5.22_a (0.11)	5.20_a (0.11)	5.10_a (0.11)
Competence	0.12	.89	4.83_a (0.09)	4.88_a (0.09)	4.88_a (0.09)
Power	0.69	.50	3.50_a (0.14)	3.54_a (0.14)	3.71_a (0.14)
Enjoyment	2.56	.09	4.91_a (0.12)	5.28_b (0.12)	5.08_{ab} (0.12)

Note: Means in each row with different subscripts were significantly different at the $p < .01$ level. For example, a mean with subscript 'a' differs from a mean with subscript 'b'.

Experiment 1 Discussion

The results of Experiment 1 suggest that people's ability to detect others' listening is limited, providing further evidence in support of Hypothesis 1. While some partners listened attentively, others pretended to listen, and still others focused on an external stimulus, their counterparts did not discern these differences. These results also offer additional evidence in support of Hypothesis 2: the error appeared to be driven by the over-attribution of attentive listening—distracted and feigned listeners were rated similarly to those who were attentively listening (with all condition-level means above 5 on a 1-7 scale). We believe that the lack of difference in perceptions between distracted and feigned listeners was driven by the social desirability of appearing like a good listener, leading participants in the *Distracted* condition to feign listening to hide their inattentiveness, even without explicit experimental instructions to do so. While it may also be possible that manipulated partners were able to effectively divide their attention between multiple stimuli, in which case their attentiveness may not have differed across the three conditions, previous research shows that directing cognitive attention toward an additional stimulus interferes with the attention paid to the original stimulus (e.g., Cohen & Gordon-Salant, 2017; McKnight & McKnight, 1993; Patten et al., 2014). Thus, we do not believe that this explains our results. However, we address this alternative explanation further in the design of Experiment 4.

Applying a lens model approach (Brunswik, 1952), we propose that the inaccuracies in interpersonal perception found in our Exploratory Study and Experiment 1 could stem from two possible explanations: (1) people rely on invalid cues when judging listening, and (2) listeners convey limited or invalid cues of their listening. The former explanation would suggest that people who are listening behave differently than those who are not, but perceivers are relying on the wrong cues when making their judgments. Conversely, the latter explanation suggests that observable behavioral differences between those who are listening and those who are not are limited (or absent). We examine these two explanations in Experiment 2.

Experiment 2

In Experiment 2, we investigated behavioral differences between people who are attentively listening and those who are not. In doing so, we aim to clarify why perceptions of conversational listening do not align with actual listening: do targets fail to convey cues of inattentive and/or attentive listening? Or are perceivers simply missing diagnostic, observable cues? In Experiment 2, we ask third-party observers to code the behavior of the listeners (i.e., targets) from Experiment 1 and assess their listening throughout the conversation (by guessing their assigned condition, and separately, by coding potential behavioral cues of attentive and inattentive listening). This design also allowed us to test whether removing the cognitive demands of live conversation might improve peoples' ability to accurately detect listening quality.

Experiment 2 Method

Participants

A total of 650 participants were recruited through Amazon's Mechanical Turk (mTurk) to complete a 15-minute survey about conversations in exchange for $1.20. After data exclusions due to technical difficulties and failed attention checks,[4] we obtained data from 398 participants. Participants reported their gender ("What is your gender? Male, Female, Non-binary/Other; 62% Male, 38% Female) and their age ("How old are you (in years)?" [Open-ended numeric entry]; M_{age} = 35 years, SD_{age} = 10 years).

Procedure

When we recorded the interaction in Experiment 1, we positioned the camera facing the "manipulated" partner in each conversation, offering the viewer an unobstructed frontal view of this participant. In Experiment 2, participants ("observers") watched and evaluated the videos of the conversations we collected in Experiment 1. Specifically, they watched a recording of one of the interactions and were asked to report the extent to which the participant in the video engaged in several behaviors. To gather fine-grained coding of the videos, we asked observers to watch each video divided into 1-minute segments. After each 1-minute segment, the video paused, and

[4] We excluded data from 252 participants (19 reported they would be unable to watch a video as part of the study; 61 participants failed an attention check at the start of the survey; 61 participants reported they would be unable to listen to audio during the study; 104 participants failed our second attention check which occurred in the middle of the video coding task; 7 participants failed to complete the full survey).

the observer reported how frequently the individual in the video engaged in various behaviors. Then the video would resume playing and pause again after another minute. This process repeated until the video ended. Overall, due to random assignment and our exclusions, 10 videos were rated by a single observer, 24 were rated by two observers, 129 were rated by three observers, and 224 videos were rated by four observers.

Measures

Based on prior listening theory (e.g., Bodie et al., 2012) and the open-ended responses of participants from a pilot study[5], we generated a list of verbal and non-verbal behaviors that social scientists and laypeople have suggested might represent valid cues of good and bad listening. For each video segment, observers rated the extent to which the manipulated conversation partner engaged in each of those behaviors (detailed below) on a scale from 1: "Not at all" to 4: "A lot."

Specifically, observers were asked to report the extent to which the manipulated conversation partner deployed verbal cues—*verbal interruptions* (cutting the partner off, talking over the partner, etc.) and *verbal affirmations* (defined as 'mhmm', 'uh-huh', 'yes', etc.)—and nonverbal cues: *nodding*, *eye contact*, *looking away* (behind the partner's head, up, down, or to the side), *smiling*, *fidgeting*, *leaning forward*, and *leaning backward*.

Predicting Listener Condition. After watching a full video, observers received an explanation of the lab paradigm that we used to collect the video—including an explanation of the experimental manipulation. We asked observers to guess which condition the manipulated conversation partner had been assigned to: attentive, distracted, or feigned listening.

[5] In this pilot study, a sample of participants (N = 829) were asked what they look for when they are trying to detect whether someone is "listening attentively to you, pretending to listen, or tuning out altogether." Participants listed "3 qualities, behaviors, and/or tendencies" using an open-ended response format. These responses were read by the authors to derive a list of behaviors commonly thought to signal listening.

Own and Partner Perceptions. Observers also reported whether they thought the individual was a good listener, and whether they were likable using the same items as in Experiment 1. Further, we asked observers to guess the evaluations that the Experiment 1 conversation partners provided of the manipulated participant in the video.

Experiment 2 Results

In the analyses that follow, we assess differences in observers' evaluations of the target individuals across conditions. To do so, we used mixed effects models with a fixed-effect for the condition and a random-effect for the dyad number to account for multiple ratings of each conversation.

Behavioral Coding

Conducting Tukey's test for post-hoc comparisons, we find that target participants (our manipulated participants from Experiment 1) were more likely to interrupt their partner when they were in the *Feigned Listening* compared to the *Listening* condition ($b = 0.2$, *95% CI* [0.04, 0.49], $p = .02$), and marginally more than in the *Distracted* condition ($b = 0.20$, *95% CI* [-0.02, 0.43], $p = .09$; see Table 3). However, beyond this behavioral difference, observers did not detect any significant differences across the three conditions in terms of the frequency of other verbal or non-verbal behaviors that both lay people and psychologists consider to be pervasive, observable, and diagnostic cues of attentive versus inattentive listening (see Table 3). In other words, conversation partners who were instructed to listen attentively were no more likely to nod, affirm, lean forward, lean backward, maintain eye contact, or look away than those explicitly assigned to direct their attention elsewhere.

Table 3

Frequency of Various Behavioral Listening Cues Across Conditions.

	Omnibus F-Value	Attentive Listening *M (SE)*	Feigned Listening *M (SE)*	Distracted *M (SE)*
Behavior				
Verbal interruptions	4.33*	1.72_{a} (0.07)	1.99_{b} (0.07)	1.78_{ab} (0.07)
Verbal affirmations	1.12	3.00_{a} (0.06)	2.87_{a} (0.06)	2.91_{a} (0.07)
Nodding	2.17	3.05_{a} (0.07)	2.85_{a} (0.07)	2.92_{a} (0.07)
Eye Contact	0.03	3.43_{a} (0.06)	3.45_{a} (0.06)	3.45_{a} (0.06)
Looking Away	0.76	2.01_{a} (0.07)	2.09_{a} (0.08)	2.14_{a} (0.08)
Smiling	1.94	3.05_{a} (0.08)	3.01_{a} (0.08)	3.21_{a} (0.08)
Fidgeting	1.54	2.58_{a} (0.08)	2.72_{a} (0.09)	2.51_{a} (0.09)
Leaning Forward	0.65	1.89_{a} (0.09)	2.03_{a} (0.09)	1.91_{a} (0.10)
Leaning backwards	1.42	1.60_{a} (0.08)	1.77_{a} (0.08)	1.61_{a} (0.08)
Predicted Partner Perceptions				
Listening	0.23	5.12_{a} (0.08)	5.06_{a} (0.08)	5.12_{a} (0.08)
Liking	0.09	4.82_{a} (0.08)	4.78_{a} (0.08)	4.80_{a} (0.08)
Enjoyment	0.27	4.96_{a} (0.08)	4.88_{a} (0.08)	4.95_{a} (0.08)
Own Perceptions				
Listening	0.76	4.91_{a} (0.09)	4.75_{a} (0.09)	4.81_{a} (0.09)
Liking	1.43	4.83_{a} (0.08)	4.77_{a} (0.08)	4.64_{a} (0.08)

Note: Omnibus F-values for condition effects, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Means in each row with different subscripts were significantly different at the $p < .05$ level. For example, a mean with subscript ‘a’ differs from a mean with subscript ‘b’.

Condition Guess

Next, we parsed the data based on methods from signal detection theory, where the signal was the target’s listening (signal is considered present in the *Listening* condition; signal is considered absent in the *Distracted* and *Feigned Listening* conditions; see Table 4). When the signal was present (i.e., the target was in the *Listening* condition), observers recognized this to be the case 80% of the time. However, when the signal was absent (i.e., the target was in either the *Distracted* or *Feigned Listening* conditions), observers still incorrectly believed they were in the

Listening condition 74% of the time. As in Experiment 1, these results show a strong bias toward reporting that the target was listening, regardless of whether this was actually the case. In other words, observers were insensitive to actual variations in listening, such that they believed *distracted* or *feigned listening* to be attentive listening in three-quarters of cases.

Table 4

Signal Detection Analysis of Experiment 2 Results

	Observer guessed target *not listening*	**Observer guessed target *listening***
Target *not listening*	26% CORRECT	74% TYPE I ERROR
Target *listening*	20% TYPE II ERROR	80% CORRECT

Our inferential analyses are in line with the descriptive pattern. Overall, 37% of observers correctly guessed the listener's condition assignment, but accuracy differed by video condition. Using a binary logistic regression mixed-effects model, with Tukey-method for pairwise comparisons, we found that observers were significantly less likely to correctly guess the condition (0/1) when the target was in the *Distracted* (14% correct; $b = -3.16$, *95% CI* [-3.96, -2.36], $SE = 0.34$, $p < .001$) or *Feigned Listening* conditions (15% correct; $b = -3.13$, 95% *CI* [-3.92, -2.34], $SE = 0.34$, $p < .001$) compared to the *Listening* condition (80% correct). We found no differences between the *Distracted* and *Feigned Listening* conditions ($b = 0.03$, 95% *CI* [-0.83, 0.89], $p = .997$). This difference was driven by an omnibus tendency for observers to over-ascribe listening irrespective of its actual presence: 76% of observers guessed that the target individual in the video was in the *Listening* condition.

Own and Partner Perceptions

Observer evaluations of the target's listening and likability did not differ significantly between conditions. Similarly, there were no differences in observers' predictions of how the manipulated conversationalist was perceived by their partner during the live conversation (see Table 4).

Finally, we linked the perceptions of third-party observers to those made by targets' original conversation partner from Experiment 1. There was no relationship between perceptions of listening as rated by third-party observers and by conversation partners from Experiment 1, $b = 0.17$, *95% CI* [-0.13, 0.47], $SE = 0.15$, $p = .26$. Thus, there was little to no consensus on the targets' level of listening between conversation partners and observers. Combined with the overall lack of behavioral differences between attentive and inattentive listeners, this result suggests that there was no clear signal of listening to be detected.

Experiment 2 Discussion

The results of Experiment 2 were consistent with those of our Exploratory Study and Experiment 1: even when people did not bear the cognitive load of active conversation (but simply observed others conversing), they were unable to detect whether conversation partners were listening or merely pretending to. As before, third-party observers showed something like the "truth bias" (Bond & DePaulo, 2006; Levine et al., 1999; Vrij, 2008) such that they believed a target was attentively listening even when that individual was incentivized to direct their attention elsewhere.

Behavioral coding revealed one observable behavioral difference between conditions: participants in the *Feigned Listening* condition were more likely to verbally interrupt their partner. It may be the case that because these participants were attending to the videos, and *pretending* to attentively listen to their partner, their spoken participation in the conversation was poorly timed. Prior work suggests that even minor disruptions or distractions can disrupt conversational flow and turn-taking (e.g., Boland et al., 2021; Truong et al., 2020), so it is no wonder that attending to a simultaneous distractor stimulus changed interruptive conversational behavior. What may be more surprising is that targets' observable behavior did not show differences on any other verbal or non-verbal cues.

These results clarify our understanding of Experiment 1. Namely, it does not appear that the unmanipulated partners simply missed obvious cues of listening because they were too cognitively taxed with participating in the conversation. Rather, drawing on the lens model (Brunswik, 1952), it appears that inaccuracies in listening perception are, at least in part, due to a dearth of diagnostic observable cues of attentive listening (Hypothesis 2).

Thus far, our results suggest that people have a strong baseline assumption that counterparts are listening, and that inaccuracies in listening detection are largely driven by the over-attribution of listening when it is not taking place. In Experiment 3, we investigate whether correcting this baseline assumption improves accuracy in listening detection. Namely, we examine whether people can accurately detect listening even when they have perfect insight into the mind of the listener—when they know *exactly* how much time the speaker spent listening (and not).

Experiment 3

We have found that people commonly believe a conversation partner is listening when they aren't—this was true for individuals actively engaged in a conversation as well as dispassionate, cognitively unencumbered third-party observers. Across these studies, however, participants assessed someone else's listening—someone whose mind they could not access. Thus, these results may represent a strong baseline assumption about attentiveness (on average) which may lead participants to err purely because they don't have insight into the *true* amount of attentive listening that has occurred. If participants were aware of the amount of inattentive listening their partner engaged in overall, would they be able to detect the rise and fall of it with greater accuracy? In Experiment 3, we test whether people can accurately detect listening when they are perfectly aware of the underlying base rate of attentiveness, by asking them to diagnose *their own* listening after the fact.

Method

Participants

130 individuals were recruited from the participant pool at a university in the northeastern United States consisting of university students, staff, and members of the local community. Participants came to the behavioral lab to participate in a 30-minute study about conversation for which they were paid a flat rate of $10, with the potential to earn bonus payments of up to $21. Data from a total of 40 participants were excluded,[6] leaving a final sample of data from 90 participants. Participants reported their gender ("What is your gender?"; 51% Male, 49% Female), their age ("What is your age (in years)? [Open-ended numeric entry]; M_{age} = 37 years, SD_{age} = 16 years), their ethnicity ("Which choice most accurately describes your ethnicity?" 52% White, 9% Black or African American, 1% American Indian or Alaska Native, 21% Asian, 17% Other), and their employment status ("What statement best describes your current employment situation?" 40% Working as paid employee, 17% Self-employed, 32% Student, 11% Other).

[6] 16 participants experienced technological issues (e.g., the videos didn't record, or the song did not play properly through the speakers); 14 participants did not follow the instructions (e.g., telling the researcher they didn't realize they had to listen to the story or song when they were explicitly instructed to do so); 10 participants compromised the experimental procedure so that they could more easily win the bonus (e.g., wearing a hat or sunglasses during one part of the procedure but not during the other).

Procedure

In this study, we used a within-subjects design. Participants were seated in a room with an experimenter, who read two stories out loud to each participant, with story order counterbalanced, while music was playing in the room. We selected two songs ("I Am My Own Grandpa," and "Big Rock Candy Mountain") that were obscure enough that participants were unlikely to be familiar with them, but used plain language that would be easy to understand and remember. Participants received different instructions about how to direct their attention and behave with respect to the story and the background music while each story was read, according to the experimental condition to which they were assigned. Specifically, all participants were told to listen attentively to one story and were randomly assigned to one of two conditions for the other story (described below; order randomized). Participants were video recorded as they listened. To ensure this procedure was similar to the social interaction procedures in our other studies, we told the participant that the experimenter (who was reading the story to them while sitting a few feet away in the same room) would be evaluating their listening quality. Additionally, while reading the story, we instructed the experimenter to look up and make eye contact with the participant at regular intervals to ensure that the experience felt interactive. These elements of the design were a purposeful effort to introduce the real-world social pressures of being perceived as a good listener.

Attentive Listening Instructions. During one of the two stories that the experimenter read to the participant during the study, participants were asked to "listen as attentively as possible to the story." They were told that they would answer comprehension questions about the story and would receive a $1.00 bonus for each question they answered correctly.

Inattentive-Listening Instructions. During the other story (counterbalanced), all participants were incentivized to experience one of two levels of cognitive distraction, which we manipulated between participants. In the *Inattentive-Listening* condition (n = 45), participants were asked to "listen as attentively as possible to the song playing in the room" and told that they would receive a $1.00 bonus for each line of lyrics that they correctly recalled. In the *Semi-Attentive Listening* condition (n = 45), participants were asked to "listen as attentively as possible to the story AND the song playing in the room" and that they would receive a $1.00 bonus for each comprehension question they answered about the story *and* each line of song lyrics they correctly recalled.

Manipulation Check

Participants completed seven comprehension questions about the story their partner read to them, as well as seven fill-in-the-blank questions about the lyrics of the song playing in the room. Further, after completing each listening task, participants reported how "attentively" they listened to the story, and how "distracted" they were while their partner was reading the story (1: "*Not at all"*; 5: "*Extremely"*). This new measure addresses a limitation in the design of Experiment 2: we did not have a measure of how much of the conversation participants could recall of the conversation, but instead focused on their recall of the commercials (i.e., the distractor task). Thus, in our Experiment 3 design, we include recall measures for both the interpersonal listening content (story) as well as the distractor task (song).

Detecting their Own Listening

After participants finished the listening task, they were shown ten separate 5-second, muted video clips of themselves listening (five clips from each story), a methodology commonly used in "thin-slice" research (e.g., Ambady et al., 2000; Ambady & Rosenthal, 1992). For each video clip, participants were asked to guess which set of instructions they were following during the recorded time. We then asked participants "How confident are you in your answer?" (1 = Not at all confident; 7 = Extremely confident). Finally, participants guessed the number of clips (of the ten) for which they thought they had correctly identified the condition assignment.

Results

Manipulation Check

Our listening manipulation appears to have been successful: Participants in both the *Inattentive* and *Semi-Attentive Listening* conditions reported listening more attentively to the story when they were instructed to do so ($M_{Inattentive}$ = 5.69, $SD_{Inattentive}$ = 1.12; $M_{Semi\_Attentive}$ = 5.91, $SD_{Semi\_Attentive}$ = 1.00) than when they were instructed to listen to the song ($M_{Inattentive}$ = 3.18, $SD_{Inattentive}$ = 1.85; $M_{Semi\_Attentive}$ = 5.16, $SD_{Semi\_Attentive}$ = 1.28; *Inattentive*: $t(44) = 9.03$, $p < .001$, *Cohen's d* = 1.60, 95% *CI* [1.07, 2.14]; *Semi-Attentive*: $t(44) = 4.19$, $p < .001$, *Cohen's d* = 0.65, 95% *CI* [0.31, 0.99]). Additionally, participants in the *Inattentive* and *Semi-Attentive Listening* conditions reported feeling more distracted from the story when they were instructed to listen to the song ($M_{Inattentive}$ = 4.73, $SD_{Inattentive}$ = 1.74; $M_{Semi\_Attentive}$ = 4.87, $SD_{Semi\_Attentive}$ = 1.70) than when they were instructed to listen to the story ($M_{Inattentive}$ = 3.67, $SD_{Inattentive}$ = 1.49; $M_{Semi\_Attentive}$ = 3.91, $SD_{Semi\_Attentive}$ = 1.87; *Inattentive*: $t(44) = 4.28$, $p < .001$, *Cohen's d* = 0.66, 95% *CI* [0.32, 0.99]; *Semi-Attentive*: $t(44) = 4.17$, $p < .001$), *Cohen's d* = 0.53, 95% *CI* [0.26, 0.80]).

When we examine the responses to the story comprehension questions, we see that participants in the *Inattentive Listening* condition answered more questions correctly when they were asked to listen to the story (M = 3.78, SD = 1.61) than when they were asked to listen to the song (M = 2.60, SD = 1.37, $t(44) = 3.70$, $p < .001$, *Cohen's d* = 0.79, 95% *CI* [0.30, 1.27]). However, those in the *Semi-Attentive Listening* condition answered a similar number of story comprehension questions correctly when they were asked to listen to the story (M = 3.56, SD = 1.37) as when they were asked to listen to both the story *and* the song (M = 3.04, SD = 1.43, $t(44) = 1.70$, $p = 0.10$, *Cohen's d* = 0.36, 95% *CI* [-0.08, 0.80]).

Further, participants in both the *Inattentive* and *Semi-Attentive Listening* conditions recalled more song lyrics correctly when they were asked to listen to the song ($M_{Inattentive} = 2.11$, $SD_{Inattentive} = 1.70$; $M_{Semi\_Attentive} = 1.31$, $SD_{Semi\_Attentive} = 1.41$) than when they were asked to listen to the story ($M_{Inattentive} = 0.33$, $SD_{Inattentive} = 0.56$; $M_{Semi\_Attentive} = 0.69$, $SD_{Semi\_Attentive} = 1.00$; *Inattentive:* $t(44) = 6.74$, $p < .001$, *Cohen's d* = 1.40, 95% *CI* [0.82, 1.98]; *Semi-Attentive:* $t(44) = 2.85$, $p = .007$, *Cohen's d* = 0.50, 95% *CI* [0.13, 0.88]).

In sum, these results suggest that our manipulations were effective. Both self-report and behavioral measures show that participants paid less attention to the story when instructed to do so.

Signal Detection of Attentive Listening

To test our key hypotheses, we first examined participants' accuracy in detecting their own level of listening, leveraging both their perception of their own behavioral cues and their recall of how much they were actually listening during the experimental task.

Overall, participants correctly guessed their listening on 64% of trials (31% when listening attentively to the story; 33% when listening inattentively to the story). Thus, in over one third of trials (36%), participants did not correctly guess their own listening in a task they completed minutes prior (19% when listening attentively to the story; 17% when inattentively listening to the story).[7] As before, we draw from signal detection theory to investigate the

[7] Accuracy rates were slightly higher for participants in the *Inattentive* (69% correct overall; 34% when listening attentively to the story; 35% when inattentively listening to the story) than the *Semi-Attentive Listening* (58% correct overall; 28% when listening attentively to the story; 30% when inattentively listening to the story) condition.

direction of this error—treating attentive listening (whether the participant had been instructed to attend to the story) as the signal (see Table 5; Stanislaw & Todorov, 1999). Here we see that when the signal was present (i.e., the participant had been instructed to listen to the story), participants judged it to be so in 61% of trials (66% *Inattentive* condition; 56% *Semi-Attentive* condition). However, when the signal was absent (i.e., the participant had been instructed to listen to the song), participants incorrectly believed they were listening attentively in 49% of trials (49% *Inattentive* condition; 48% *Semi-Attentive* condition). Unlike in Studies 1 and 2, we did not observe a bias for participants to report that they were listening—having completed the listening task just minutes prior, participants appeared to know exactly how much they were (and were not) listening throughout the task. However, this did not increase their accuracy in detecting their own listening—in fact, hit rates decreased in this study compared to the earlier studies, with participants correctly guessing they were listening in less than two-thirds of trials. Thus, even when we corrected participants' baseline over-assumption of listening, they were still relatively unimpressive in their ability to decipher the rise and fall of attentive listening at specific times during the interaction.

Table 5

Signal Detection Analysis of Experiment 3 Results

	Clip identified as *inattentive or semi-attentive listening*	**Clip identified as *attentive listening***
***Inattentive Listening* Condition**		
Inattentive listening trial	52% CORRECT	48% TYPE I ERROR

Attentive-listening trial	34% TYPE II ERROR	66% CORRECT
***Semi-Attentive Listening* Condition**		
Semi-attentive listening trial	51% CORRECT	49% TYPE I ERROR
Attentive-listening trial	44% TYPE II ERROR	56% CORRECT

Perceptual Accuracy in Listening Detection

Moving beyond descriptive measures of accuracy, we investigate whether participants were more likely to guess they were listening on trials in which they were instructed to listen attentively to the story. Thus, we conducted a binary logistic regression predicting participant guesses of listening (values recoded such that 0 = *Guess not listening*, 1 = *Guess listening*) from a fixed-effect for trial type (whether the participant had been instructed to listen attentively to the story) and trial number (one out of ten guesses), and a random effect for participant to account for repeated observations (Bates et al., 2015). We found a positive association between trial type and participant guess of listening, $b = 0.57$, *95% CI* [0.19, 0.95], $p = .003$, and no effect of trial number, $b = 0.008$, *95% CI* [-0.06, 0.07], $p = .80$.[8] Translating this result into an odds ratio, participants were 1.77 times more likely to guess they were listening attentively when they had received instructions to do so than when they had not. Although participants were significantly more likely to guess that they were listening during a given clip when they were actually listening, the results are not encouraging regarding the extent of participants' discernment—they correctly guessed their listening on only 61% of trials.

Confidence in Listening Perceptions

[8] We find no interaction between trial type and condition (*Inattentive* vs. *Semi-Attentive*), $b = -0.44$, *95% CI* [-0.99, 0.10], $p = .11$.

Next, we assess participants' self-reported confidence in their guesses. We conducted a mixed-effects regression predicting participant confidence ratings in their guesses from a fixed-effect for whether they correctly guessed their listening (0 = incorrect guess; 1 = correct guess) and trial number, and a random effect for participant to account for repeated observations (Bates et al., 2015). We found a positive relationship between guess correctness and confidence, $b = 0.30$, *95% CI* [0.14, 0.46], $p < .001$, and a positive relationship between trial number and guess confidence, $b = 0.05$, *95% CI* [0.03, 0.08], $p < .001$. On average, participants reported that they thought they had guessed approximately six clips correctly ($SD = 2.10$). Further, there was a significant positive relationship between the number of clips that participants thought they had correctly identified and the number of clips they had actually correctly identified, $b = 0.46$, *95% CI* [0.25, 0.67], $p < .001$. Overall, 39% of participants thought they correctly identified more clips than was the case (44% underestimated; 17% correctly estimated). These results suggest that participants had some insight into the limitations of their listening detection accuracy, perhaps because they knew their base rate of inattention and recognized that they couldn't tell by watching when those moments occurred.

Experiment 3 Discussion

The results of Experiment 3 suggest that people do not detect listening with high accuracy, even when they observe their *own* nonverbal cues immediately after engaging in attentive, inattentive, or semi-attentive listening. In particular, mirroring the results of Experiments 1-2, people overestimated (their own) attentive listening, even just minutes after experiencing the rise and fall of their attentiveness firsthand during an interaction.

One interpretation of the results of Experiment 3 is that when people have an accurate baseline for the amount of attentive vs. inattentive listening—in this case, because they have just engaged in the task themselves—they still do not achieve anywhere near perfect accuracy in listening detection. Participants guessed incorrectly on 36% of the trials. Thus, even with insight into the mind of the listener, listening detection is still poor, likely because listeners do not give off high-fidelity, observable cues of their attentive versus inattentive minds (as was found in Experiment 2).

As at a dinner party or work meeting, our participants were under pressure to seem attentive to the experimenter sharing a story. The pressure to appear attentive and interested even when other tasks draw one's attention away is shared by most psychology experiments and real social settings alike.

Still, it remains unclear whether people were able to effectively divide their attention (i.e., multi-task): perhaps distracted listeners were still listening to their partners to a substantial extent. In our final study, we disentangle multi-tasking from feigned attentive listening by limiting listeners' physical ability to hear (rendering multi-tasking impossible).

Experiment 4

In Experiments 1 and 3, we guided participants' listening behavior with instructions and incentives (to listen attentively, inattentively, or semi-attentively). However, it's possible that the human mind is highly capable of dividing its attention between multiple stimuli, including live conversation, and that all participants in Experiments 1 and 3—even those whose attention was divided—were able to process their partner's words and respond accordingly (even while simultaneously attending to video advertisements or music lyrics).

In Experiment 4, we tested this explanation by strictly limiting participants' *ability to hear* their conversation partner's words—by garbling portions of the conversation. At the same time, we incentivized "listeners" to act as if the conversation was proceeding without disruption.

Specifically, we told listeners that their payment depended on maintaining their partner's ignorance about the sporadically garbled content. Although individuals in the real world are not explicitly financially incentivized to appear like good listeners, we hoped to emulate the social incentives that often lead people to feign attentive listening.

While we define listening as attending to and processing another person's verbal, nonverbal, and prosodic cues during conversation, we chose to limit targets' access to their partner's verbal cues in this study, rather than nonverbal or prosodic cues. The exchange of verbal content (words) between two or more people is what defines conversation (Yeomans et al., 2021). By limiting verbal content, this study provides a stringent test of listening perception, while still allowing participants to carry on a responsive, live interaction. Though people are well-practiced in conversing with limited or no access to their partner's prosodic or nonverbal cues (e.g., via text-based media like email or text messaging), conversing without verbal content presents a more stringent test of the ability to feign attentive listening while maintaining a responsive interaction.

Between participants, we varied how much of the conversation was garbled. This design helped us answer two important questions. First, to what extent are our earlier results due to the fact that people are excellent at dividing their attention between listening and a distractor task? Secondly, how pervasive can lapses in hearing and listening become before conversation partners begin to notice them? As previously noted, recall (which was used in the previous studies) is an imperfect measure of listening (Thomas & Levine, 1994). We circumvent this challenge in Experiment 4 by limiting auditory input itself. All procedures and analyses were pre-registered: https://aspredicted.org/blind.php?x=sf5nz6.

Experiment 4 Method

Participants

A total of 242 participants[9] were recruited from the participant pool at a university in the northeastern United States, consisting of university students, staff, and members of the local community. Participants came to our behavioral lab in groups of 4-6 to participate in a 60-minute study about conversation in which they would engage in a series of one-on-one conversations with the other participants in the session. Participants were paid a flat rate of $20 with the potential to earn up to $20 in additional bonus payments. After excluding data from conversations in which a participant expressed confusion about the instructions or disclosed the listening manipulation to their partner (n = 9 conversations), we analyzed data from 305 conversations (N = 235 unique participants; 66% engaged in three conversations, 27% engaged in two conversations, 7% engaged in a single conversation).

Participants reported their gender ("What is your gender? Male, Female, Non-binary/Other"; 44% Male, 52% Female, 1% Non-binary/Other, 3% Non-response) and their age ("What is your age?" [Open-ended numeric entry]; M_{age} = 32 years, SD_{age} = 14 years).

Protocol

Participants completed ten minute, one-on-one, round-robin video chats with 2-3 different partners. Conducting these conversations over video chat (instead of face-to-face) allowed us to asymmetrically manipulate audio for one participant in each dyad. The number of participants that arrived for each experimental session determined the number of conversations each participant completed (we maximized the number of unique round-robin dyads possible in each session). We instructed participants to "get to know their conversation partners," and provided them with a list of seven conversation topics that they could (but did not have to) use

[9] We were unable to reach our pre-registered sample size due to the onset of the COVID-19 pandemic. This study was actively running in the lab when local shelter-in-place restrictions required data collection to stop.

(e.g., Where did you grow up? What did you do last summer? Do you have any pets? Have you watched any good TV shows lately?). Participants also received private instructions: half of the participants in each session were randomly assigned to the role of *target*, the other half to the role of *perceiver*.

Target. Those participants assigned to the role of *target* learned that there may be times during the conversation when they might not understand their partner—instead of hearing their partner's words, they would hear garbled sounds, and this garbled sound was an *intentional* part of the study, not a technical glitch. By design, we used a computer program to obscure what the targets (but not the perceivers) could hear during the conversation. Unbeknownst to the targets, we randomly assigned them to one of four conditions corresponding to the amount of time the ten-minute conversation would be garbled from their perspective only: 0% (0 seconds garbled), 25% (150 seconds garbled), 50% (300 seconds garbled), or 75% (450 seconds garbled). We configured the computer program to intermittently turn on and off a voice filter that obscured what the targets heard from their partner at specified intervals (25% garbled condition: 30 seconds filter off, 10 seconds filter on; 50% garbled condition: 10 seconds filter off, 10 seconds filter on; 75% garbled condition: 3 seconds filter off; 10 seconds filter on), which repeated throughout the ten-minute conversation. Random assignment to the level of garbling (0%, 25%, 50%, 75%) was performed at the session level (n_0 = 80 conversations; n_{25} = 87 conversations; n_{50} = 70 conversations; n_{75} = 68 conversations).

Our manipulation allowed us to test whether people can give the impression that they are listening, even when they cannot actually hear the words that their partner is saying. The results of our earlier studies might be explained by the idea that targets who were rated as good listeners despite distractions were simply able to effectively divide their attention. However, this explanation cannot apply to the current study where participants could not hear parts of the conversation. If participants exposed to largely garbled speech manage to create the impression of good listening, we can conclude with greater confidence that feigned listening is very difficult to detect. Importantly, we incentivized targets to act as if nothing was amiss:

> *"Your primary goal is to be (or at least appear to be) a GREAT LISTENER. After each conversation, your partner will rate how well they think you listened to them. If your partner reports that you were a good listener (i.e., one of the 40 best listeners in the study)…then you will earn a $20 bonus…In order to earn the "good listener" bonus, your partner should think the conversation has played out naturally and smoothly."*

Perceivers. Those assigned as *perceivers* were completely ignorant of the targets' instructions and the conversational garbling. But for fairness in payment potential, the perceivers were also incentivized:

> *"Your primary goal is to be likeable. After each conversation, your partner will rate how much they liked you. If your partner reports that you were highly likeable (i.e., one of the 40 most likeable people in the study)…then you will earn a $20 bonus."*

In this manner, both targets and perceivers were financially incentivized to make a positive impression on each other.

After each ten-minute conversation, participants completed a post-conversation survey, which included self-reported items about the conversation and their partner. At the end of the lab session, participants completed a final demographic survey and received payment.

Measures completed by perceivers

After each conversation, the perceivers (who were blind to the manipulation) reported their perceptions of their (manipulated) counterpart's listening ("My partner was a good listener," "My partner was engaged in the conversation") and responsiveness ("My partner made me feel heard," "My partner made me feel validated," "I felt that my partner cared about me") on a scale from 1: "*Strongly Disagree"* to 7: "*Strongly Agree*." They also reported the extent to which they agreed or disagreed that their partner "worked hard to listen to me," "was attentive to what I was saying," and "understood what I was saying" (1: "*Strongly Disagree"*; 7: "*Strongly Agree"*). These evaluations represent our key dependent variables in this study.

Perceivers also estimated the percentage of the things they said during each conversation that they believe the target heard (1-100%), and their general assessments of their (manipulated) counterpart, including likeability ("My partner is likable," "I liked my partner," "I would enjoy spending time with my partner," "I disliked my partner (R)"), intelligence ("My partner is smart"), and interestingness ("My partner is interesting") on a scale from 1: "*Strongly Disagree"* to 7: "*Strongly Agree*."

Measures completed by targets

Targets made predictions about their partners' perceptions of them. Each target predicted their partner's perceptions of their listening ("My partner thought I was a good listener," and "My partner thought I was engaged in the conversation") and responsiveness ("My partner felt heard," "My partner felt validated," and "My partner felt that I cared about them") on a scale from 1: "*Strongly Disagree"* to 7: "*Strongly Agree*." Targets also reported whether they "worked hard to listen to my partner," "was attentive to what my partner was saying," and "understood what my partner was saying" on a scale from 1: "*Strongly Disagree*" to 7: "*Strongly Agree*." Targets also reported whether they thought their partner would say that they could hear them (yes/no) and guessed "My partner would say that I could hear ___% of what they said" (1-100%). These measures served as manipulation checks.

Finally, targets predicted how likeable their partner would rate them ("My partner thinks I'm likable," "My partner liked me," "My partner would enjoy spending time with me," "My partner dislikes me (R)," 1: "*Strongly Disagree"* to 7: "*Strongly Agree"*).

We told all targets and perceivers that their ratings of each other would remain private to assuage impression management concerns.

Experiment 4 Results

Since each participant engaged in several conversations, we conducted mixed-model regressions, clustering at the participant level and controlling for order effects (Bates et al., 2015). We report specific results for each dependent variable, specifying the results of each pairwise comparison (conducting Tukey-tests for multiple comparisons), as well as the overall omnibus F-test. All results are presented in aggregate in Table 6.

Table 6

Perceiver and Target Ratings Across Conditions.

		Experimental Condition: Percentage of Garbled Conversational Content			
	Omnibus F-Value	**0%** ***M (SE)***	**25%** ***M (SE)***	**50%** ***M (SE)***	**75%** ***M (SE)***
Perceiver Ratings					
Listening	0.36	6.24_{a} (0.16)	6.17_{a} (0.15)	6.19_{a} (0.17)	6.00_{a} (0.17)
Responsiveness	1.95	6.00_{a} (0.18)	5.72_{a} (0.17)	5.99_{a} (0.19)	5.44_{a} (0.19)
Worked Hard	1.13	5.79_{a} (0.19)	5.96_{a} (0.17)	6.13_{a} (0.19)	5.66_{a} (0.20)
Attentive	0.90	6.18_{a} (0.17)	6.13_{a} (0.16)	6.24_{a} (0.18)	5.85_{a} (0.18)
Understood	3.88*	6.18_{a} (0.19)	6.12_{a} (0.17)	6.13_{a} (0.19)	5.39_{b} (0.19)
Guess % Heard	5.77**	88.40_{a} (2.49)	85.70_{a} (2.35)	81.90_{ab} (2.61)	74.20_{b} (2.63)
Liking	0.81	6.03_{a} (0.17)	6.06_{a} (0.16)	5.99_{a} (0.18)	5.71_{a} (0.18)
Intelligence	1.54	6.35_{a} (0.17)	6.16_{a} (0.16)	6.08_{a} (0.18)	5.83_{a} (0.18)
Interestingness	1.05	6.00_{a} (0.20)	5.94_{a} (0.19)	5.93_{a} (0.21)	5.54_{a} (0.21)
Target Ratings					
Predicted Listening	3.37*	6.22_{a} (0.16)	6.23_{a} (0.16)	5.94_{ab} (0.17)	5.57_{b} (0.17)

Predicted Responsiveness	3.02*	6.04_a (0.16)	6.02_a (0.15)	5.88_{ab} (0.17)	5.42_b (0.17)
Predicted Liking	4.65**	5.56_a (0.14)	6.18_b (0.14)	5.98_{ab} (0.15)	5.56_a (0.15)
Worked Hard	3.74*	5.94_a (0.19)	6.61_{ab} (0.18)	6.77_b (0.20)	6.63_{ab} (0.20)
Attentive	0.39	6.54_a (0.15)	6.64_a (0.14)	6.59_a (0.16)	6.42_a (0.16)
Understood	31.51***	6.59_a (0.24)	5.52_b (0.23)	4.64_c (0.25)	3.34_d (0.25)
Percent Heard	109.10***	94.20_a (2.64)	68.90_b (2.52)	50.40_c (2.80)	27.00_d (2.82)
Guess Partner Rated % Heard	4.90**	78.70_a (9.57)	78.00_{ab} (5.66)	67.20_a (6.31)	50.50_c (5.35)

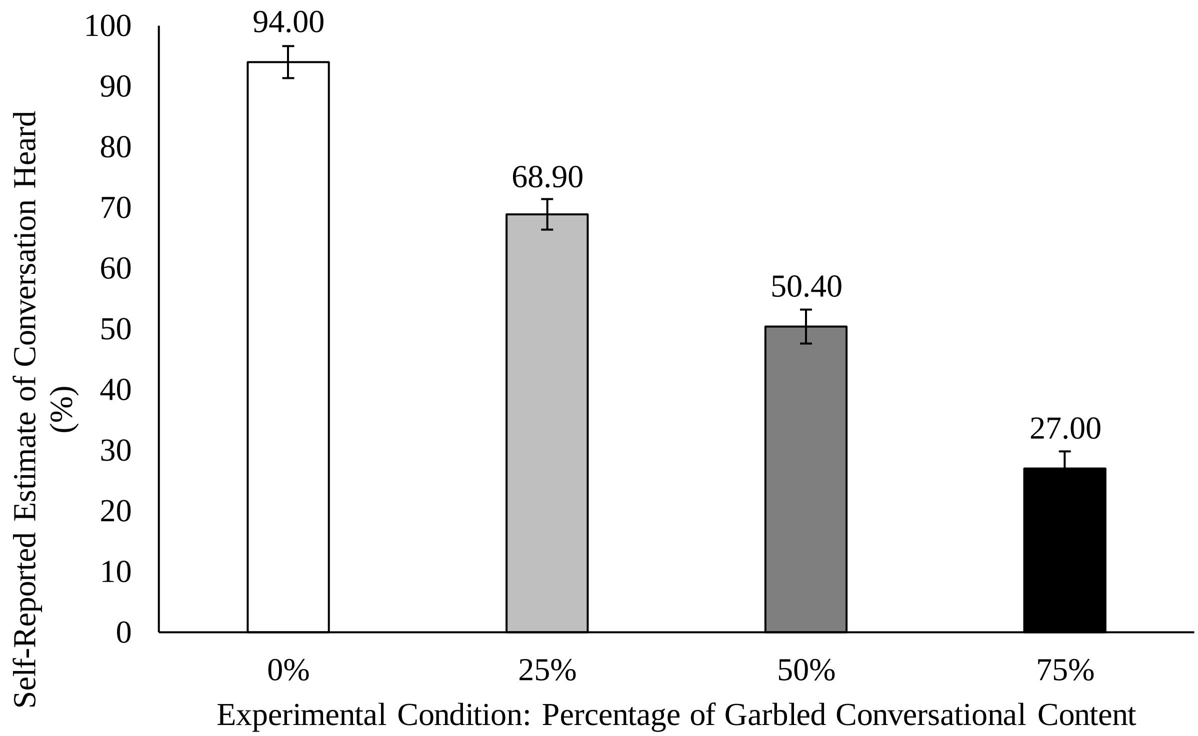
Note: Omnibus F-values for condition effects, *p < .05, **p < .01, ***p < .001. Means in each row with different subscripts were significantly different at the p < .05 level. For example, a mean with subscript 'a' differs from a mean with subscript 'b'.

Was the listening manipulation successful?

Several results suggest that our listening manipulation was successful. First, targets' reports of the percentage of the conversation that they heard decreased significantly with each increase in garbling (see Figure 4; 0 vs. 25%: b = -25.20, *95% CI* [-34.70, -15.73], p < .001; 25 vs. 50%: b = -18.60, *95% CI* [-28.40, -8.74], p < .001; 50 vs. 75%: b = -23.30, *95% CI* [-33.70, -12.99], p < .001). Further, targets' self-reported understanding during the conversation diminished with each increase in garbling (see Figure 5; 0 vs. 25%: b = -1.07, *95% CI* [-1.92, -0.21], p = .008; 25 vs. 50%: b = -0.88, *95% CI* [-1.76, 0.001], p = .05; 50 vs. 75%: b = -1.30, *95% CI* [-2.23, -0.37], p = .002).

Figure 4

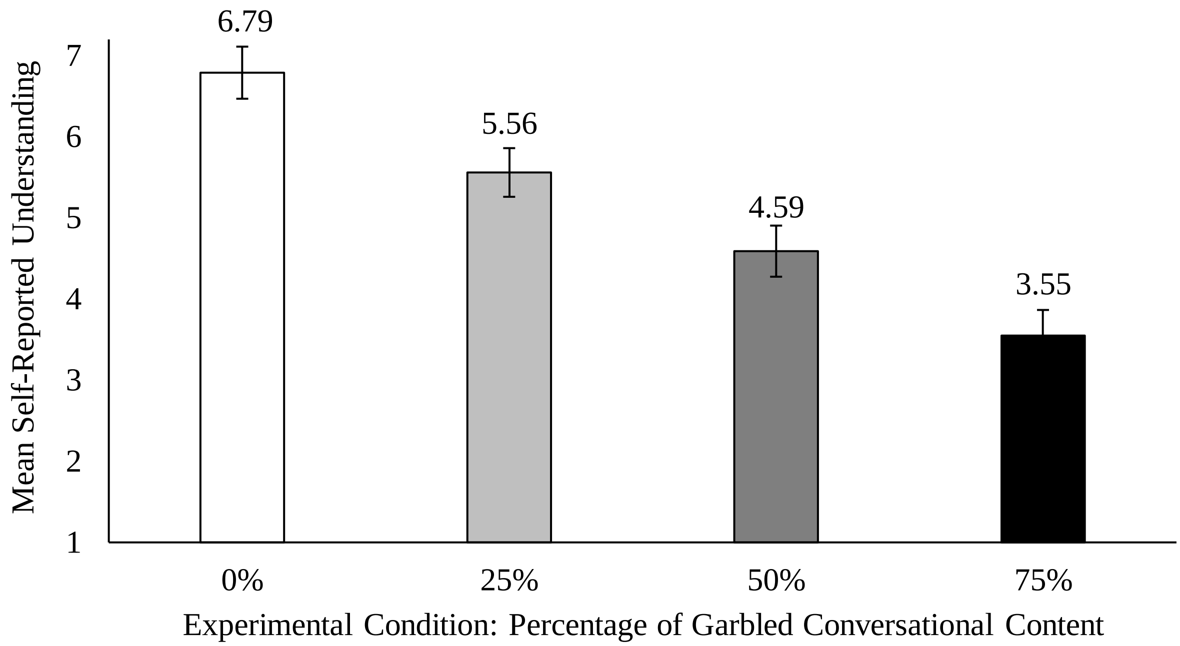
Target Self-Reports of Amount of Garbled Content Across Conditions.



Note: Targets estimated the percentage of conversational content they heard, with error bars representing standard errors (Experiment 4). Participants' self-reported estimates were quite accurate. Those who heard 100% of the content estimated they heard 94%, those who heard 75% estimated 69%, those who heard 50% estimated 50%, and those who heard 25% estimated 27%. Each condition significantly differed from all others.

Figure 5

Target Self-Reported Understanding Across Conditions.



Note: Targets' mean self-reported understanding during the conversation, with error bars representing standard errors (Experiment 4). Each condition significantly differed from all others.

Additionally, targets reported working harder to listen to their partner in the 50% compared to the 0% condition (0 vs. 50%: *b* = 0.83, *95% CI* [0.11, 1.55], *p* = .003)–and marginally harder in the 25% and 75% compared to the 0% condition (0 vs. 25%: *b* = 0.67, *95% CI* [-0.01, 1.35], *p* = .06; 0 vs. 75%: *b* = 0.69, *95% CI* [-0.03, 1.41], *p* = .07). Interestingly, there were no significant differences on this measure between the obscured conditions (25 vs. 50%: *b* = 0.16, *95% CI* [-0.55, 0.86], *p* = .94; 25 vs. 75%: *b* = 0.02, *95% CI* [-0.68, 0.73], *p* = .99; 50 vs. 75%: *b* = -0.13, *95% CI* [-0.88, 0.60], *p* = .96).

Importantly, targets reported no differences in their attentiveness during the conversation across the conditions, suggesting they were indeed incentivized to appear as though they were listening to their partner (0 vs. 25%: $b = 0.11$, *95% CI* [-0.42, 0.64], $p = .95$; 25 vs. 50%: $b = -0.05$, *95% CI* [-0.60, 0.50], $p = .99$; 50 vs. 75%: $b = -0.17$, *95% CI* [-0.75, 0.41], $p = .87$; all other *p*s > .72).

Did targets think they could feign listening?

Comparing targets' predictions of their partners' perceptions of their listening and responsiveness across conditions, we find that participants in the 75% garbled condition predicted that their partner would rate them as being poorer listeners than those in the 0% ($b = -0.65$, *95% CI*[-1.27, -0.02], $p = .04$) and 25% conditions ($b = -0.66$, *95% CI*[-1.27, -0.05], $p = .03$), and as less responsive than in the 0% ($b = -0.62$, *95% CI* [-1.23, -0.01], $p = .04$) and 25% conditions ($b = -0.60$, *95% CI* [-1.20, -0.005], $p = .047$), but show no differences from the 50% condition (listening: $b = -0.37$, *95% CI* [-1.01, 0.27], $p = .44$; responsiveness: $b = -0.45$, *95% CI* [-1.08, 0.17], $p = .24$). Additionally, when asked to predict how likable their partner would find them, targets in the 0% condition actually predicted their partner would find them less likable than those in the 25% condition ($b = 0.62$, *95% CI* [0.10, 1.13], $p = .01$), would be similarly likable to those in the 50% ($b = 0.42$, *95% CI* [-0.13, 0.96], $p = 0.20$) and 75% conditions ($b = 0.00007$, *95% CI* [-0.55, 0.55], $p > .99$). Further, those in the 75% condition also felt they would be seen as less likable than those in the 25% condition ($b = -0.62$, *95% CI* [-1.15, -0.08], $p = .02$), but no different than those in the 50% ($b = -0.42$, *95% CI* [-0.98, 0.15], $p = .22$) condition. Thus, targets did feel that the manipulation would affect their partners' impression of them.

Finally, when comparing targets' predictions of how much of the conversation their partner thought they heard, targets in the 75% condition reported that their partner would think they heard significantly less than in 25% condition ($b = -27.52$, *95% CI* [-48.70, -6.36], $p = .006$), and marginally less than in the 0% and 50% conditions (0 vs. 75%: $b = -28.20$, *95% CI* [-

57.80, 1.41], *p* = .07; 50 vs. 75%: *b* = -16.71, *95% CI* [-39.10, 5.65], *p* = .20), with no other between-condition differences reaching significance (*p*s > .58). These results suggest that, except for those whose conversations were 75% garbled, most targets felt they were able to convince their partner of their listening.

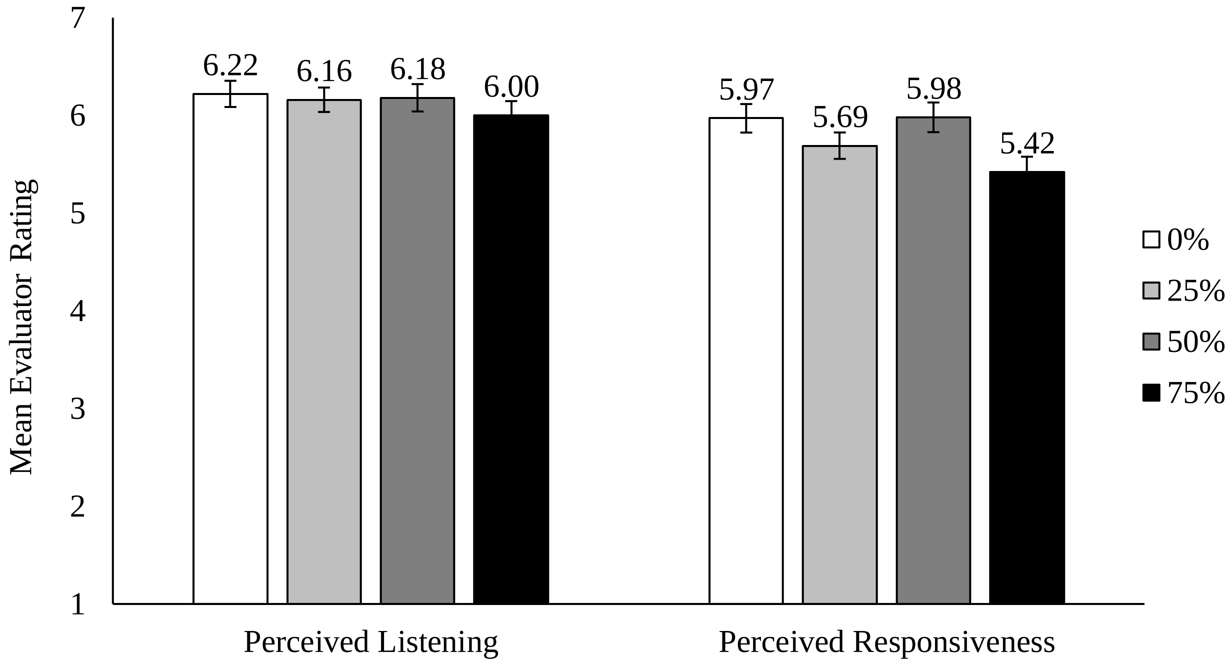
Did perceivers notice?

As in Experiments 1-3, we find no differences across conditions in (unmanipulated) perceivers' perceptions of the targets' (manipulated) listening (see Figure 6; 0 vs. 25%: *b* = -0.07, *95% CI* [-0.65, 0.52], *p* = .99; 0 vs. 50%: *b* = -0.04, *95% CI* [-0.66, 0.57], *p* = .99; 0 vs. 75%: *b* = -0.23, *95% CI* [-0.85, 0.39], *p* = .78; all other *p*s > .86).

Further, when reporting how responsive their partner was in the conversation, perceivers rated targets in the 0%, 25%, 50%, and 75% conditions as similarly responsive (0 vs. 25%: *b* = -0.28, *95% CI* [-0.94, 0.38], *p* = .69; 0 vs. 50%: *b* = -0.005, *95% CI* [-0.70, 0.69], *p* = .99; 0 vs. 75%: *b* = -0.56, *95% CI* [-1.27, 0.16], *p* = .18; all other *p*s >.69; see Figure 6).

Figure 6

Partner Perceptions of Target Listening and Responsiveness Across Conditions.



Note: Perceivers' mean ratings of target listening and responsiveness across conditions, with error bars representing standard errors. There were no significant differences in perceived listening across conditions. Perceived responsiveness was significantly lower in the 75% garbled condition, compared to the 0% and 50% (but not 25% garbled) conditions.

Perceivers' ratings of targets' effort during the conversation also did not differ across conditions (0 vs. 25%: *b* = 0.17, *95% CI* [-0.49, 0.84], *p* = .91; 0 vs. 50%: *b* = 0.35, *95% CI* [-0.35, 1.05], *p* = .57; 0 vs. 75%: *b* = -0.13, *95% CI* [-0.83, 0.58], *p* = .97; all other *p*s > .09) and attentiveness (0 vs. 25%: *b* = -0.05, *95% CI* [-0.66, 0.56], *p* = .99; 0 vs. 50%: *b* = 0.06, *95% CI* [-0.58, 0.71], *p* = .99; 0 vs. 75%: *b* = -0.32, *95% CI* [-0.97, 0.33], *p* = .56; all other *p*s > .43). Like ratings of responsiveness, we found no differences between the 0%, 25% and 50% conditions on perceptions of the targets' understanding (0 vs. 25%: *b* = -0.07, *95% CI* [-0.72, 0.59], *p* = .99; 0 vs. 50%: *b* = -0.06, *95% CI* [-0.75, 0.63], *p* = .99; 25 vs. 50%: *b* = 0.008, *95% CI* [-0.66, 0.68], *p*

= .99). Only when targets heard 75% of the conversation garbled did perceivers notice that the targets' understanding suffered (0 vs. 75%: *b* = -0.79, *95% CI* [-1.48, -0.10], *p* = .02; 25 vs. 75%: *b* = -0.72, *95% CI* [-1.40, -0.05], *p* = .03; 50 vs. 75%: *b* = -0.73, 95%*CI* [-1.44, -0.03], *p* = .04).

When directly asked to estimate the percentage of the conversation their partner heard, perceivers made statistically equivalent estimates across the 0%, 25% and 50% garbled conditions (0 vs. 25%: *b* = -2.73, *95% CI* [-11.70, 6.21], *p* = .86; 0 vs. 50%: *b* = -6.49, *95% CI* [-15.90, 2.93], *p* = .28; 25 vs. 50%: *b* = -3.76, *95% CI* [12.90, 5.40], *p* = .71). Though perceivers guessed that targets heard less in the 75% condition compared to the 0% and 25% conditions (0 vs. 75%: *b* = -14.18, *95% CI* [-23.60, -4.74], *p* = .001; 25 vs. 75%: *b* = -11.45, *95% CI* [-20.70, -2.26], *p* = .008), they reported no differences between the 50% and 75% conditions (*b* = -7.70, *95% CI* [-17.40, 1.96], *p* = .17). Even in that condition, perceivers overestimated targets' ability to hear, guessing that they *could* hear 75% of the conversation, when they could only hear 25%.

Finally, when asked to make person-level judgments about the targets, the perceivers reported no differences in liking (0 vs. 25%: *b* = -0.03, *95% CI* [-0.58, 0.64], *p* = .99; 0 vs. 50%: *b* = -0.04, *95% CI* [-0.68, 0.61], *p* = .99; 0 vs. 75%: *b* = -0.31, *95% CI* [-0.96, 0.33], *p* = .59; all other *p*s > .69), judgments of interestingness (0 vs. 25%: *b* = -0.06, *95% CI* [-0.77, 0.66], *p* = .99; 0 vs. 50%: *b* = -0.07, *95% CI* [-0.83, 0.69], *p* = .99; 0 vs. 75%: *b* = -0.47, *95% CI* [-1.23, 0.30], *p* = .39 all other *p*s > .47), or judgments of intelligence (0 vs. 25%: *b* = -0.20, *95% CI* [-0.81, 0.42], *p* = .84; 0 vs. 50%: *b* = -0.28, *95% CI* [-0.92, 0.37], *p* = .68; 0 vs. 75%: *b* = -0.53, *95% CI* [-1.17, 0.12], *p* = .15; all other *p*s > .52).

These results present a fairly extreme example of listening perception inaccuracy. Even when targets could only hear 25% of their partner's spoken words in the conversation, they were rated as high-quality listeners, appearing engaged and responsive to their partner. Interestingly, at this extreme, evaluations of understanding suffered—suggesting that the inability to hear one's partner did have some effect on perceivers' impressions. In support of results from Studies 1-3, these findings reveal the insensitivity of listening perceptions to the actual cognitive experience of listeners.

Experiment 4 Discussion

The results of Experiment 4 suggest that perceptions of listening are largely impervious to a listener's inability to actually hear their partner's words. It was only when 75% of the verbal content was garbled that perceivers noticed that their conversation partners understood less. But even in this extreme case, perceiver ratings of listening, responsiveness, effort, attentiveness, interestingness, likeability, and intelligence did not differ (both compared to targets who could hear more and those who could hear everything).

Even though these results are contingent on targets being incentivized to feign their listening, it is striking that people can portray attentive listening even when they cannot hear most of what is being said. This suggests that perceptions of listening are largely determined by behaviors that are surprisingly untethered from the informational meaning of the conversation. While one would hope such extreme situations are uncommon outside the laboratory, these results reveal that listening perceptions can be divergent from reality—especially when incentives to deceive are present. These incensitves are likely common in many naturalistic conversations given the social desirability of being perceived as a "good listener."

While these these findings do not rule out the possibility of listener multi-tasking in Experiments 1 and 3, they do rule *in* the highly skilled ability of listeners to convey attentiveness and understanding, even in the absence of it. Overall, these results reinforce the findings from our previous studies that perceptions of listening are often inaccurate: there seems to be a substantial gap between *feeling* heard and actually *being* heard.

General Discussion

Conversational listening is a key building block of human social functioning. Information transmission, interpersonal connection, conflict management, happiness—the key foundations of human flourishing—hinge critically on our ability to hear, understand, and respond to others (e.g., Schiller, 1996; Yeomans et al., 2021). A large body of work finds that speakers and listeners alike experience myriad benefits when people are perceived as "good listeners" (e.g., Bodenmann, 2005; Huang et al., 2017; Kuhn et al., 2018; Lloyd et al., 2015; Qian et al., 2019; Shafran-Tikva & Kluger, 2018; Wanzer et al., 2004; Yeomans et al., 2020). At the same time, a rich literature on the failings of mind perception (Epley, 2008; Eyal et al., 2018) and the inability to detect lies from truth (Bond & DePaulo, 2006) calls into question whether perceptions of listening accurately reflect the internal experience of listening (i.e., *being heard*) or merely reflect an illusory subjective experience in the mind of the speaker (i.e., *feeling heard*). Although the subjective experience matters immensely, it may not represent the construct it is understood to represent: *attending to and processing of another person's verbal, nonverbal, and paralinguistic cues during conversation*.

Across five studies, we find that there is a difference between *being* and *feeling* heard. In live conversation, people's perceptions of their conversation partners' listening were only moderately related to the partners' internal cognitive experiences of listening (Exploratory Study). Though people's listening fluctuated naturally throughout their conversations (with mind wandering reported 24% of the time), they were also able to nimbly adjust their listening in line with instructions—by either listening attentively, inattentively, or dividing their attention when they were told to do so (Experiment 1)—and their conversation partners were scarcely able to detect the rise and fall of their partner's attentiveness, whether via natural fluctuation or via our experimental intervention. This phenomenon extended to third-party observers who were not immersed in the conversation (Experiment 2), listeners who looked back on their own listening (Experiment 3), and people interacting with listeners who could not physically hear what their partner was saying (Experiment 4). Thus, across a diverse set of studies, we find support for our three primary hypotheses: that (1) perceptions of conversational listening often do not align with listeners' internal cognitive experiences; (2) they are often inaccurate due to a lack of diagnostic behavioral cues displayed by the listener—attentive listeners behave similarly to inattentive listeners; and consequently (3) perceivers primarily *overestimate* the extent to which their conversation partners are listening to them.

Theoretical Contributions

Our work makes several contributions that advance our understanding of interpersonal perception, listening, and the psychology of conversation more broadly. Across studies, we find a consistent pattern of *overestimation*: perceivers were biased towards over-attributing listening, frequently believing their conversation partners were listening attentively to them when they were not. These results are similar to the truth bias, in which people assume others are telling the truth more often than they are (Bond & DePaulo, 2006; Levine et al., 1999; Vrij, 2008). In fact, the direction of this error may be socially adaptive for perceivers—perhaps it is less costly to mistakenly assume someone is listening when they aren't than to erroneously accuse them of inattentiveness. This is in line with prior work on "want-should conflict" (Bitterly et al., 2015) because an accusation of inattentiveness carries immediate social costs, whereas the risks of being misunderstood are probabilistic and temporally distant.

Inaccurate perceptions of listening make sense, as we find a dearth of behavioral differences between attentive listeners and those pretending to listen attentively—with both engaging in a range of verbal and non-verbal behaviors to the same extent. A similar pattern has been found to contribute to inaccuracy in lie detection, such that truth tellers and liars behave very similarly across a range of observable behaviors (Hartwig & Bond, 2011; Wiseman et al., 2012). It may be that conversationalists engage in a specific *type* of deception—people may feign their listening when their attention is drawn away from the speaker, perhaps due to the social desirability of appearing as a "good listener." In fact, it may be helpful for conversationalists to think of feigned cues of listening as a specific type of deception, even if these moments of deception are often prosocial (e.g., Levine & Schweitzer, 2015).

Our findings raise important questions: When are communicators better or worse off erring in the direction of over-optimism about their partners' attentiveness? And when are listeners better or worse off covering up moments of inattentiveness? Though people seem well practiced in these maneuvers, their optimality depends on interlocutors' informational and relational goals, such as whether they care more about accurate information exchange, politeness, enjoyment, or the smoothness of the conversation (Yeomans et al., 2021).

Further, these findings contribute to an emerging literature shedding light on people's inattentiveness to others during conversation. Recent work has shown that people are insensitive to conversational coherence and perspective-taking—in both their lack of reaction to nonsensical turns of phrase (e.g., "colorless green ideas sleep furiously"), and even to whom their conversational partners are (e.g., Galantucci & Roberts, 2014; Galantucci et al., 2018; Roberts et al., 2016; Yeomans & Brooks, 2021). Together with previous work, our findings build on a growing literature that highlights the risks of miscoordination inherent in live conversation (see Yeomans et al., 2021 for a review).

Recent research on the importance of establishing and sustaining shared reality—the perceived commonality of internal states with other people—for the development and maintenance of relationships suggests that our findings may be particularly consequential (Higgins et al., 2021; Rossignac-Milon et al., 2021; Rossignac-Milon & Higgins, 2018). Inaccuracy in listening perception is almost by definition a roadblock to shared reality—it may feel polite to feign listening in the moment, but this deception, if left unnoticed or unrepaired, will likely erode shared understanding and may jeopardize relationships over time.

Limitations and Future Directions

Our methods and findings are qualified by several limitations that offer fruitful avenues for future research. First, we primarily observed conversations between strangers. Future work should explore the listening behaviors between people who know each other. Perhaps close pairs are better able to detect idiosyncratic cues of poor listening (or perhaps they believe they're better, but aren't); perhaps people are particularly good at fooling close others that they're listening; perhaps close pairs are less likely to devote the effort needed to feign good listening; and so on—with meaningful consequences for relational harmony and information exchange beyond one-time conversations.

Additional work is also required to understand how these processes play out across different *types* of conversations (Yeomans et al., 2021). Every conversation is wildly different—in fact, every aspect of conversational context can change between conversations, and within them, from one turn to the next (who, what, where, when, why, and how they're occurring). We focus here on interactions primarily driven by relational goals, however, perceptual errors about others' attentiveness—and the ways in which conversationalists manage them—may be different in contexts where high-informational goals (e.g. learning, brainstorming, making decisions, persuading, exchanging accurate information) loom large.

Additionally, by design, we conducted our studies in a controlled lab setting. Future research should investigate more naturalistic contexts, especially contexts in which it may be easier (or harder) to detect feigned or inattentive listening—for example by considering the role of communication medium or modality (e.g., Boland et al., 2022; Schroeder et al., 2017; Liberman et al., 2022). Different types of communication media often constrain the cues that are available to listeners (e.g., video-conferences might only show people from the shoulders up; phone calls provide no nonverbal cues at all); change the reliability of certain cues of attention (e.g., one cannot be sure of directed eye contact over video-conferencing, Abi-Esber et al., forthcoming); and therefore alter people's overall ability to accurately detect listening in an increasingly digital world.

Second, future work should examine how misperceptions of listening influence important downstream consequences, such as learning, productivity, decision-making, trust, liking, and other indicators of relationship quality. As suggested above, this work should specifically investigate the role of repeated interaction—while feigned listening may provide benefits within a single conversation, the costs of such deceptions may reveal themselves over time and may influence both relational and informational outcomes across many conversations in a relationship (Yeomans et al., 2021).

Third, additional research is required to understand the relative role of perceived attention and perceived information processing. Prior research has identified multiple stages in attitude formation, distinguishing between attention to information versus information processing

(Kunda, 1990; McGuire, 1968; Minson et al., 2020). Whereas attention has some external markers (eye gaze, absence of interruption), information processing is internal. Thus, it may be the case that individuals use cues of attention (which are easy to feign) to infer others' internal processing. Future work could investigate whether reminding people that signals of *attention* are not reliable signals of *processing* might improve the accuracy of perceived listening overall.

Finally, additional research is required to understand the relative roles of the speaker and listener in promoting these misperceptions. Are speakers insensitive to their partners'—and perhaps their own—lapses in listening? Do we all underestimate the extent to which the human mind wanders? Are listeners skilled at feigning attention when their attention is divided—and when their minds predictably wander? All of the above? We find preliminary evidence that listeners feign attentiveness in live conversation through various non-verbal behaviors—nodding and smiling when their attention is elsewhere. But we suspect speakers may also be to blame by relying too heavily on low-fidelity signals of their attentiveness. In fact, it's likely that these effects are recursively reinforcing: speakers hold overly optimistic beliefs about partner attentiveness, which creates unrealistic norms and expectations of attentiveness, which puts pressure on people to feign attentive listening to cover frequent moments of inattentiveness, which, unnoticed and unrepaired, reinforce overly optimistic beliefs.

On the other hand, recognizing this perceptual error may present an opportunity for growth. We call for future work to investigate explicit conversational strategies that may disrupt this fallacious listening loop. For example, with minimal intervention, listeners may be nudged to more explicitly admit their lapses in listening, which would allow for more immediate repair of glitches in information exchange, or learn to use more verbal signals of attentive listening that cannot be faked, such as making call-backs to earlier topics, paraphrasing, or asking follow-up questions—behaviors recently described as powerful signals of "conversational uptake" (Collins, 2022; Demszky et al., 2021; Huang et al., 2017; McQuaid et al., 2015). At the same time, speakers, too, may be able to adjust their expectations to anticipate the cognitive demands of attentive listening, to become more forgiving of their partners' lapses in listening, and to remain open to repair strategies and attuned to high-fidelity signals of attentive listening, like verbal uptake.

Conclusion

Recent work suggests that people are blind to major disruptions in logical coherence during interpersonal encounters (Galantucci et al., 2018; Galantucci & Roberts, 2014; Roberts et al., 2016). On the one hand, this is unsurprising: conversation is a complex, overwhelming decision environment that requires relentless thinking, perceiving, monitoring, and deciding—the human mind is bound to make mistakes. On the other hand, the misalignment between perceptions of listening and listeners' underlying cognitive experiences that we document here is a radically different approach than one that has been taken by the prior literature which has largely focused on developing, conveying, and identifying "good" and "active" listening skills. Taken together, our findings suggest that people often misjudge whether their partners are listening (or not) and call for a re-examination of this important and pervasive behavior for which cognitive and social experiences seem misaligned. Though both are important, it seems there is a noticeable (but navigable) gap between *feeling* heard and *being* heard in conversation.

Constraints on Generality

Participants throughout our studies had short get-to-know-you conversations with strangers. This is something people experience in many contexts—at the grocery store, dinner parties, work conferences, and on. Indeed, conversations in the lab follow the same mechanics and norms as they do in the real world—people take turns speaking and listening, and providing verbal, non-verbal, and paralinguistic cues to their partners in pursuit of a vast array of relational and informational goals. And though our manipulations (e.g., nudging people to listen attentively or inattentively) were somewhat artificial, we aimed to simulate real-world contexts. For example, Experiment 1 was modeled after the experience of chatting with someone at a bar with TVs playing sports in the background. Experiment 4 was modeled after virtual calls in which people experience audio disruptions but for the sake of continuity and naturalness, forge ahead. Thus, we expect that these results would generalize to conversations between strangers outside of the lab, especially when the goal of the conversation is relational (as in our studies). Nevertheless, our studies were all studied in a lab context, and therefore future research is needed in the "real world," where environments are more complex, social (and monetary) stakes may be higher, and conversants don't feel like they're being watched. Specifically, additional research is required to understand these effects in group conversations, conversations between close partners, and conversations where informational goals are at the forefront.

Context of Research

Prior work on conversational listening has focused largely on perceptions of listening and the consequences of such perceptions. The current research is part of a broader stream of work in which we seek to identify the extent to which people's perceptions of conversational listening are accurate. This paper builds on two important themes in our ongoing research programs. First, we are curious about the accuracy of normative claims often made by lay people and researchers alike. In this case, are people who are seen as "good listeners" actually good? Do they deserve the credit they are getting? Or are they good at multi-tasking and/or faking it? In what ways is pretending to listen good or bad? Secondly, in our work, we seek to study interpersonal phenomena—including cooperation and conflict—by studying *conversation*, the ubiquitous context where much of human social life unfolds. By analyzing transcripts from real conversations, we can capture observable behavior and unobservable psychology on a continuous or turn-by-turn basis—and connect those behaviors and cognitions to real consequences (embracing all the challenges and joy that this approach brings).

Author Contributions

Conceptualization: HKC, JAM, AK, AB
Data curation: HKC
Formal analysis: HKC
Investigation: HKC
Methodology: HKC, JAM, AK, AB
Project administration: HKC
Resources: HKC
Software: HKC
Supervision: JAM, AB
Validation: HKC, JAM, AK, AB
Visualization: HKC
Writing – original draft: HKC, JAM, AB
Writing – review & editing: HKC, JAM, AK, AB

References

Agne, R. R. (2018). “You’re not listening”: Negotiating and problematizing ideal interaction through metatalk about listening in the Waco Standoff. *Western Journal of Communication*, *82*(3), 356-375.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In *Advances in experimental social psychology* (Vol. 32, pp. 201-271). Academic Press.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological bulletin*, *111*(2), 256.

Back, M. D., & Nestler, S. (2016). 5 Accuracy of judging personality. *The Social Psychology of Perceiving Others Accurately*, 98.

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

Bitterly, T. B., Mislavsky, R., Dai, H., & Milkman, K. L. (2015). Want-should conflict: A synthesis of past research. In W. Hofmann & L. F. Nordgren (Eds.), *The psychology of desire* (pp. 244–264). The Guilford Press.

Bodenmann, G. (2005). Dyadic coping and its significance for marital functioning. In T. A. Revenson, K. Kayser, & G. Bodenmann (Eds.), Couples coping with stress: Emerging perspectives on dyadic coping (pp. 33–49). Washington, DC, US: American Psychological Association.

Bodie, G. D. (2011). The Active-Empathic Listening Scale (AELS): Conceptualization and evidence of validity within the interpersonal domain. *Communication Quarterly*, *59*(3), 277-295.

Bodie, G., St. Cyr, K., Pence, M., Rold, M., & Honeycutt, J. (2012). Listening Competence in Initial Interactions I: Distinguishing Between What Listening Is and What Listeners Do. *International Journal of Listening, 26*(1), 1-28.

Boland, J. E., Fonseca, P., Mermelstein, I., & Williamson, M. (2021). Zoom disrupts the rhythm of conversation. *Journal of Experimental Psychology: General*.

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

Boothby, E.J., Cooney, G., Sandstrom, G.M., & Clark, M.S. (2018). The liking gap in conversations: Do people like us more than we think? *Psychological Science, 29(11)*.

Brunswik, E. (1952). The conceptual framework of psychology. (International Encyclopedia of Unified Science, Volume 1, Number 10.) Chicago: The University of Chicago Press.

Brunswik, E. (1956). Historical and thematic relations of psychology to other sciences. *The Scientific Monthly, 83*(3), 151-161.

Cheon, J., & Grant, M. (2009). Web-based assessment tool for communication and active listening skill development. *Tech trends, 6*(53), 24-34.

Collins, H. K. (2022). When listening is spoken. *Current Opinion in Psychology,* 101402.

Demszky, D., Liu, J., Mancenido, Z., Cohen, J., Hill, H., Jurafsky, D., & Hashimoto, T. (2021). Measuring conversational uptake: A case study on student-teacher interactions. *arXiv preprint arXiv:2106.03873*.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Telling ingratiating lies: Effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology*, *48*(5), 1191.

De Ruyter, K., & Wetzels, M. G. (2000). The impact of perceived listening behavior in voice-to-voice service encounters. *Journal of Service Research*, *2*(3), 276-284.

Epley, N. (2008). Solving the (real) other minds problem. *Social and Personality Psychology Compass*, *2*(3), 1455-1474.

Eyal, T., Steffel, M., Epley, N. (2018). Perspective mistaking: Accurately understanding the mind of another requires getting perspective, not taking perspective. *Journal of Personality and Social Psychology*.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, *82*(6), 878.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences*, *11*(2), 77-83.

Galantucci, B., & Roberts, G. (2014). Do we notice when communication goes awry? An investigation of people's sensitivity to coherence in spontaneous conversation. *PloS one*, *9*(7), e103182.

Galantucci, B., Roberts, G., & Langstein, B. (2018). Content deafness: When coherent talk just doesn't matter. *Language and Communication*, *61*, 29–34.

Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, *150*(4), 700.

Hafen Jr, M., & Crane, D. R. (2003). When marital interaction and intervention researchers arrive at different points of view: The active listening controversy. *Journal of Family Therapy*, *25*(1), 04-14.

Hartwig, M., & Bond Jr, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological bulletin*, *137*(4), 643.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *Journal of personality and social psychology*, *78*(1), 81.

Higgins, E. T., Rossignac-Milon, M., & Echterhoff, G. (2021). Shared reality: From sharing-is-believing to merging minds. *Current Directions in Psychological Science*, *30*(2), 103-110.

Huang, K., Yeomans, M., Brooks, A. W., Minson, J., & Gino, F. (2017). It doesn't hurt to ask: Question-asking increases liking. *Journal of personality and social psychology*, *113*(3), 430.

Itzchakov, G., Castro, D. R., & Kluger, A. N. (2016). If you want people to listen to you, tell a story. *International Journal of Listening, 30*(3), 120-133. doi:10.1080/10904018.2015.103744

Janusik, L. A. (2007). Building listening theory: The validation of the conversational listening span. *Communication Studies*, *58*(2), 139-156.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological bulletin*, *134*(3), 404.

Kraut, R. (1980). Humans as lie detectors. *Journal of communication*, *30*(4), 209-218.

Kubota, S., Mishima, N., & Nagata, S. (2004). A study of the effects of active listening on listening attitudes of middle managers. *Journal of Occupational Health*, *46*(1), 60-67.

Kuhn, R., Bradbury, T. N., Nussbeck, F. W., & Bodenmann, G. (2018). The power of listening: Lending an ear to the partner during dyadic coping conversations. *Journal of Family Psychology, 32*(6), 762.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communications Monographs*, *66*(2), 125-144.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, *126*, 88-106.

Lieberman, A., Schroeder, J., & Amir, O. (2022). A voice inside my head: The psychological and behavioral consequences of auditory technologies. *Organizational Behavior and Human Decision Processes*, *170*, 104133.

Lloyd, K. J., Boer, D., Keller, J. W., & Voelpel, S. (2015). Is my boss really listening to me? The impact of perceived supervisor listening on emotional exhaustion, turnover intention, and organizational citizenship behavior. *Journal of Business Ethics*, *130*(3), 509-524.

McGuire, W. J. (1968). Personality and attitude change: An information-processing theory. *Psychological foundations of attitudes*, *171*, 196.

McNaughton, D., Hamlin, D., McCarthy, J., Head-Reeves, D., & Schreiner, M. (2008). Learning to listen: Teaching an active listening strategy to preservice education professionals. *Topics in Early Childhood Special Education*, *27*(4), 223-231.

McQuaid, S. M., Woodworth, M., Hutton, E. L., Porter, S., & ten Brinke, L. (2015). Automated insights: Verbal cues to deception in real-life high-stakes lies. *Psychology, Crime & Law*, *21*(7), 617-631.

Min, K. S., Jung, J. M., & Ryu, K. (2021). Listen to their heart: Why does active listening enhance customer satisfaction after a service failure?. *International Journal of Hospitality Management*, *96*, 102956.

Minson, J. A., Chen, F. S., & Tinsley, C. H. (2020). Why won't you listen to me? Measuring receptiveness to opposing views. *Management Science*, *66*(7), 3069-3094.

Mishima, N., Kubota, S., & Nagata, S. (2000). The development of a questionnaire to assess the attitude of active listening. *Journal of Occupational Health*, *42*(3), 111-118.

Nestler, S., & Back, M. D. (2013). Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science*, *22*(5), 374-379.

Qian, J., Wang, B., Song, B., Li, X., Wu, L., & Fang, Y. (2019). It takes two to tango: The impact of leaders' listening behavior on employees' feedback seeking. *Current Psychology*, *38*(3), 803-810.

Rautalinko, E., & Lisper, H. O. (2004). Effects of training reflective listening in a corporate setting. *Journal of Business and Psychology*, *18*(3), 281-299.

Roberts, G., Langstein, B., & Galantucci, B. (2016). (In)sensitivity to incoherence in human communication. *Language and Communication*, *47*, 15–22.

Rogers, C. R., & Dymond, R. F. (1954). Psychotherapy and personality change. *Chicago, IL*.

Rogers, C. R., & Farson, R. E. (1957). Active listening. *Chicago, IL*.

Rossignac-Milon, M., Bolger, N., Zee, K. S., Boothby, E. J., & Higgins, E. T. (2021). Merged minds: Generalized shared reality in dyadic relationships. *Journal of Personality and Social Psychology, 120*(4), 882.

Rossignac-Milon, M., & Higgins, E. T. (2018). Epistemic companions: Shared reality development in close relationships. *Current opinion in psychology*, *23*, 66-71.

Schiller, D. (1996). *Theorizing communication: A history*. Oxford University Press.

Schroeder, J., Kardas, M., & Epley, N. (2017). The humanizing voice: Speech reveals, and text conceals, a more thoughtful mind in the midst of disagreement. *Psychological science*, *28*(12), 1745-1762.

Shafran-Tikva, S., & Kluger, A. N. (2018). Physician's listening and adherence to medical recommendations among persons with diabetes. *International Journal of Listening*, *32*(3), 140-149.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological bulletin*, *132*(6), 946.

Smith, P. K., Wigboldus, D. H., & Dijksterhuis, A. P. (2008). Abstract thinking increases one's sense of power. *Journal of Experimental Social Psychology*, *44*(2), 378-385.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, *31*(1), 137-149.

Surra, C. A., & Longstreth, M. (1990). Similarity of outcomes, interdependence, and conflict in dating relationships. *Journal of Personality and Social Psychology*, 59, 501–516. http://dx.doi.org/10.1037/0022- 3514.59.3.501

Swann, W. B., Jr., & Gill, M. J. (1997). Confidence and accuracy in person perception: Do we know what we think we know about our relationship partners? *Journal of Personality and Social Psychology*, 73, 747–757. http://dx.doi.org/10.1037/0022-3514.73.4.747

Thomas, L. T., & Levine, T. R. (1994). Disentangling listening and verbal recall: Related but separate constructs?. *Human Communication Research, 21*(1), 103-127.

Truong, M., Fast, N. J., & Kim, J. (2020). It's not what you say, it's how you say it: Conversational flow as a predictor of networking success. *Organizational Behavior and Human Decision Processes*, *158*, 1-10.

Vonesh, E.F., Chinchilli, V.P. & Pu, K.W. (1996) Goodness-of-fit in generalized nonlinear mixed-effects models. *Biometrics*, *52,* 572–587.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley & Sons.

Wanzer, M. B., Booth-Butterfield, M., & Gruber, K. (2004). Perceptions of health care providers' communication: relationships between patient-centered communication and satisfaction. *Health communication*, *16*(3), 363-384.

Weger Jr, H., Castle, G. R., & Emmett, M. C. (2010). Active listening in peer interviews: The influence of message paraphrasing on perceptions of listening skill. *The International Journal of Listening*, *24*(1), 34-49.

Weger, H., Castle Bell, G., Minei, E., & Robinson, M. (2014). The relative effectiveness of active listening in initial interactions. *International Journal of Listening*, *28*(1), 13-31.

Wiseman, R., Watt, C., ten Brinke, L., Porter, S., Couper, S. L., & Rankin, C. (2012). The eyes don't have it: Lie detection and neuro-linguistic programming. *PloS one*, *7*(7), e40259.

Witkin, B. R. (1990). Listening theory and research: The state of the art. *Journal of the International Listening Association*, *4*(1), 7–32.

Yeomans, M., & Brooks, A. W., (2021). Topic Preference Detection: A novel approach to understand perspective-taking in conversation. *Working Paper*, https://www.mikeyeomans.info/papers/topic_selection.pdf

Yeomans, M., Minson, J., Collins, H., Chen, F., & Gino, F. (2020). Conversational receptiveness: Improving engagement with opposing views. *Organizational Behavior and Human Decision Processes*, *160*, 131-148.

Yeomans, M., Schweitzer, M. E., & Brooks, A. W. (2021). The Conversational Circumplex: Identifying, Prioritizing, and Pursuing Informational and Relational Motives in Conversation. *Current Opinion in Psychology*, *44*:1-10

Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion*, *9*(4), 478.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In *Advances in experimental social psychology* (Vol. 14, pp. 1-59). Academic Press.