American Psychologist
2017, Vol. 72, No. 7, 644–654

0003-066X/17/$12.00 http://dx.doi.org/10.1037/amp0000147

Voice-Only Communication Enhances Empathic Accuracy

Michael W. Kraus
Yale University, School of Management

This research tests the prediction that voice-only communication increases empathic accuracy over communication across senses. We theorized that people often intentionally communicate their feelings and internal states through the voice, and as such, voice-only communication allows perceivers to focus their attention on the channel of communication most active and accurate in conveying emotions to others. We used 5 experiments to test this hypothesis ($N =$ 1,772), finding that voice-only communication elicits higher rates of empathic accuracy relative to vision-only and multisense communication both while engaging in interactions and perceiving emotions in recorded interactions of strangers. Experiments 4 and 5 reveal that voice-only communication is particularly likely to enhance empathic accuracy through increasing focused attention on the linguistic and paralinguistic vocal cues that accompany speech. Overall, the studies question the primary role of the face in communication of emotion, and offer new insights for improving emotion recognition accuracy in social interactions.

Keywords: emotion, empathic accuracy, person perception, organizational behavior

Supplemental materials: http://dx.doi.org/10.1037/amp0000147.supp

Social mammals have a profound capacity to connect with others: Young Rhesus monkeys will cling to a cloth surrogate that provides the simulated tactile warmth of a caregiver, rather than a wire one that provides nutrients (Harlow, 1958), and infants have the ability to mimic simple facial expressions soon after birth (Meltzoff, 1985). Social connections are critical for managing the survival-related threats that individuals experience (Bowlby, 1988). One way that individuals develop and maintain social connections is through empathic accuracy—the ability to judge the emotions, thoughts, and feelings of other individuals (Côté & Miners, 2006; Ickes, Stinson, Bissonnette, & Garcia, 1990; Stinson & Ickes, 1992). With enhanced empathic accuracy, individuals can respond more appropriately to conflicts at work (Côté & Miners, 2006) and to support-seeking romantic partners (Richards, Butler, & Gross, 2003). Enhanced empathic accuracy also allows individuals to more easily navigate complex political organizations and social networks (Mayer, Salovey, & Caruso, 2008). In contrast, a dearth of empathic accuracy is a common symptom of many psychological disorders (American Psychiatric Association, 2013).

Despite powerful motivations to connect with others many people experience failures in social connection and understanding (Hawkley & Cacioppo, 2010). In the present research, we suggest that one potent barrier to empathic accuracy is the ways in which emotion expressions across modalities divide our attention between more and less relevant channels of communication. Humans have an impressive array of tools for expressing and perceiving the emotions of others (Zaki, Bolger, & Ochsner, 2009). Research on emotion recognition began with studies testing the hypothesis that people can recognize facial expressions of emotion cross-culturally (Ekman, 1989; Russell, 1994). More recent research reveals the power of other senses to accurately communicate emotions: Touches on the body and forearms of a stranger communicate an array of emotions (Hertenstein, Keltner, App, Bulleit, & Jaskolka, 2006) as do nonword vocal bursts played back to strangers (Gendron, Roberson, van der Vyver, & Barrett, 2014; Simon-Thomas, Keltner, Sauter, Sinicropi-Yao, & Abramson, 2009). In particular, we contend that the voice, including both speech content and the linguistic and paralinguistic vocal cues (e.g., pitch, cadence, speed, and volume) that accompany it, is a particularly powerful channel for perceiving the emotions of others. This assertion supports the central prediction tested in these studies—that voice-only

I thank Noam Segal for his contributions to early versions of this article and Zoë Kraus for inspiration. I also thank Jessica Halten, the members of the Champaign Social Interaction Laboratory, and the Yale University School of Management Behavioral Laboratory for assistance with data collection.

Correspondence concerning this article should be addressed to Michael W. Kraus, Organizational Behavior, Yale University School of Management, 165 Whitney Avenue, 5477 Evans Hall, New Haven, CT 06511. E-mail: michael.kraus@yale.edu

communication enhances empathic accuracy relative to communication across senses.

Our prediction is supported by three lines of argument and evidence. First, the voice is a dynamic and accurate channel of emotion communication when compared with other sense modalities. The voice conveys emotion both through speech content (i.e., what is said), and the linguistic and paralinguistic cues that accompany it (i.e., how it is said). Both sources of information contained in the voice provide access to the internal emotional states of others. With respect to speech content, language remains a potent way in which people intentionally relay their internal states, including emotions, to others (e.g., Pennebaker & King, 1999). As well, the core affect model of emotion suggest that people have basic low-level affective states that then become complex emotions when meaning is ascribed to them via language and conceptual learning (Russell, 2003). That people understand their own emotional states through language is a critical reason why we contend voice-only communication is likely to enhance empathic accuracy (e.g., Lindquist & Barrett, 2008).

With respect to linguistic and paralinguistic content, several studies indicate the importance of such vocal cues for perceiving the internal states of others. Vocal bursts without language are sufficient to accurately communicate emotions (Simon-Thomas et al., 2009). In thin-slicing studies of social status, pitch modulation and other vocal cues play a prominent role in accurate judgments (Hall, Coates, & LaBeau, 2005). For instance, the social status of a sample of speakers from across the United States was judged accurately by a panel of judges based only on hearing the same seven words (i.e., “yellow” “thought”) uttered by each of the speakers out of context (Kraus, Park, & Tan, in press). In particular, vocal cues tend to signal accurate information about the internal states of targets because these cues require extensive attention and control to mask—making leakage of internal states more likely (Ambady & Rosenthal, 1992; DePaulo & Rosenthal, 1979).

Three studies directly support this assertion that the voice is essential for empathic accuracy. In the studies, researchers compared empathic accuracy in perceivers when exposed to targets engaging in interactions, or discussing an emotional event, while different channels of information were available. Across the studies, voice-only communication enhanced empathic accuracy relative to visual-only information (Zaki et al., 2009), relative to visual information with audio where meaning was filtered out (Gesn & Ickes, 1999), and relative to both a silent video or a transcript (Hall & Schmid Mast, 2007). Though voice-only communication did not significantly improve empathic accuracy relative to communication across senses in these studies, a departure from our current predictions, these data are supportive of our general assertion that voice-only communication is critical for empathic accuracy.

Second, facial and nonverbal expressions are a less reliable source of accurate emotion expression than the voice. Several studies indicate that people use facial and nonverbal expressions of emotion to mask their internal states. For instance, one set of studies found that people instructed to intentionally mask their facial expressions when exposed to pleasant or unpleasant stimuli tended to do so effectively, leading to only slightly above chance lie detection in a separate observer panel (Porter & ten Brinke, 2008). Together, this research raises the possibility that nonverbal expressions detract from empathic accuracy unless one has extensive training in facial muscle recognition (e.g., Ekman & Rosenberg, 1997). As well, though people rely on facial expressions of emotion to infer traits, visual information is also a source of several known perceptual biases: For instance, people make reliable inferences about personality traits from static features of faces despite little evidence for their accuracy (Olivola, Funk, & Todorov, 2014; Zebrowitz & Montepare, 2008). In one illustrative study, stereotypes about targets (i.e., new mothers), and not emotion expressions, were used by perceivers to infer emotional states (Lewis, Hodges, Laurent, Srivastava, & Biancarosa, 2012).

Third and finally, multiple channels of emotion perception may elicit cognitive costs. As more available modalities of emotion expression provide perceivers with more information, one could argue that emotion perception across multiple senses offers the greatest chance for accurate emotion recognition. However, this logic goes against decades of research on multitasking which finds that switching between cognitive tasks—likened here to switching between modes of expression when perceiving emotions—reduces speed and accuracy of task completion relative to repeating the same task (e.g., Meyer & Kieras, 1997; Rubinstein, Meyer, & Evans, 2001). Applied to the present research, judging emotions through a single channel of communication, such as the voice, is already a complex cognitive perceptual process that becomes even more cognitively taxing when sense modalities are added. Thus, having more modalities for emotion recognition might paradoxically impede empathic accuracy.

If this analysis is sound and the voice communicates emotions at higher rates of accuracy than do visual or combined channels of communication, then there are several potential theoretical and practical implications. Much of the research on emotion and emotion perception accuracy has relied on visual cues from facial and nonverbal expressions (e.g., Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Ekman & Rosenberg, 1997). For instance, assessments of emotional intelligence of adults (Mayer et al., 2008) and children with clinical diagnoses (Baron-Cohen et al., 2001) rely primarily on assessing people’s capacity to read emotions in microexpressions surrounding the facial muscles and eyes. As well, decades of research on the universality versus cultural specificity of emotion expressions has been primarily contested in the realm of facial expressions (e.g., Ekman & Friesen, 1971) with only a few notable exceptions (e.g., Gendron et al., 2014). See Figure 1 for two graphical depictions of this imbalance of scholarship focused on facial relative to vocal expressions of emotion. In short, emotion research might benefit from added focus on the channel of communication most active in emotion recognition processes.

The above analysis sets the stage for our central hypothesis. We predict that voice-only communication enhances empathic accuracy relative to vision-only and multisensory communication. We tested this hypothesis across five experiments. In the first three experiments we examine empathic accuracy of perceivers passively watching social interactions (Experiments 1 and 3) or actively engaged in interactions (Experiment 2). In the final two experiments, in addition to examining empathic accuracy in the voice versus in combined modes of communication, we assess relative contributions of speech content and vocal cues in this process (Experiments 4 and 5).

Experiment 1

In Experiment 1, we examined our empathic accuracy hypothesis in the context of passive perceptions of interactions between friends teasing each other with voice-only, visual-only, or combined voice and visual communication enabled. We expected that participants would more accurately perceive emotions communicated through the voice-only relative to visual-only and multisense communication.

Method

All our studies were approved by the institutional review boards of the University of Illinois, Urbana–Champaign and Yale University. All participants were from the United States and were at least 18 years old. For a full list of study materials and measures please go to (https://osf.io/ux9wa/). Prior to conducting these studies we had no clear estimate for the size of the hypothesized effect of voice-only communication enhancements of empathic accuracy. Thus, we set a general sample size threshold of at least 100 participants per experimental condition. This sample size provides us 80% statistical power to detect a difference between independent means of Cohen's $d = .40$—roughly equivalent to the average effect size in all of social psychology (Richards, Bond, & Stokes-Zoota, 2003)—and a dependent means difference of $d_{rm} = .20$. For some studies we were able to collect larger samples for greater precision.

In Experiment 1, we analyzed data from 300 perceivers recruited online through Amazon Mechanical Turk. Participants were paid $1.50 for taking part in the study. The majority of the sample was female ($n = 174$), the average age was 33.38 years ($SD = 9.80$). The majority of the sample was White ($n = 251$), followed by Asian American ($n = 22$), Latino/a ($n = 16$), Black ($n = 21$), and participants who listed themselves as an other ethnic category ($n = 3$). Ethnic identity numbers exceed total sample in all experiments because participants could nominate more than one ethnic category. For exploratory analyses on gender and empathic accuracy, see the online supplementary materials.

Perceivers were exposed to a series of three social interactions where two female friends took turns teasing each other using a nickname created from two sets of initials (i.e., A. D. or L. I.). For example, one target used the nickname "Attention Deficit" to tease her friend about the friend's lack of listening skills. Following exposure to each interaction, which lasted between one and five minutes, perceivers were instructed to estimate the emotions experienced by each of the targets, starting with the first teaser, and then the second teaser. Perceivers were instructed to rate emotions

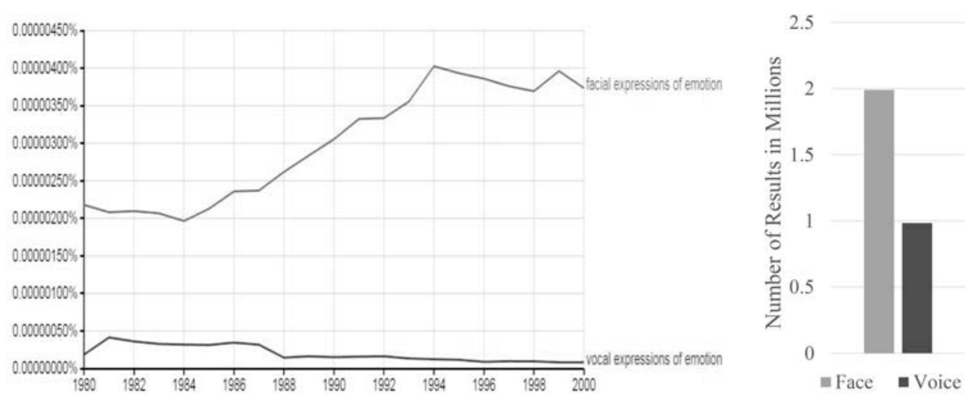

Figure 1. The left panel shows a Google Ngram search result of the corpus of English language books from 1980 to 2000 for the terms *facial expressions of emotion* and *vocal expressions of emotion* controlling for number of words in the corpus. The right panel shows the number of returned searches for the terms *face* and *emotion* relative to the terms *voice* and *emotion* on Google Scholar. See the online article for the color version of this figure.

while taking into account the entire interaction, including both the teasing portion and the portion where the target was being teased—consistent with the target ratings of their own emotions in the original study (Kraus, Horberg, Goetz, & Keltner, 2011). We examined teasing because it activates a broad range of positive and negative emotions (Keltner, Capps, Kring, Young, & Heerey, 2001), making it a particularly relevant domain for testing empathic accuracy (e.g., Kraus et al., 2011). Perceivers estimated the same 23 discrete emotion words filled out by targets at the time of the teasing interaction using 9-point Likert scales (0 = *not at all*, 8 = *a great deal*). The emotions were amusement, anger, anxiety, appreciation, compassion, contempt, curiosity, discomfort, disgust, embarrassment, empathy, enthusiasm, fear, guilt, gratitude, happy, hope, irritation, joy, love, pride, sad, and shame. Perceivers viewed the emotion rating scale after the first teasing exchange. The order of teasing interactions was randomized and we observed no order effects (p = .664).

Empathic accuracy scores were created by computing the absolute value of the difference between participant estimates of emotion and partner self-reports during each of the conversations, where lower scores indicate greater accuracy and a score of zero indicates complete agreement (Fletcher & Kerr, 2010). Across all five experiments, these scores were averaged across all emotion items for an overall metric of empathic accuracy, as in prior research (e.g., Kraus, Côté, & Keltner, 2010). See supplementary materials for exploratory analyses on empathic accuracy for positive and negative affect.

We assigned participants to one of three communication modalities for the experiment: voice-only, vision-only, and combined voice and vision communication. We accomplished this by editing the video of the interactions so that audio, video, or both were displayed for participants. Of note, though the sample was selected from the same source, participants in the vision-only condition were added to the experiment after the voice-only and combined conditions were collected.

Results

We analyzed empathic accuracy scores using a one-way analysis of variance (ANOVA) and then probed conditions differences using Fisher's least significant difference (LSD) Test. The ANOVA yielded a significant omnibus effect, $F(2, 297) = 3.22$, $p = .041$. Probing between condition differences revealed that perceivers in the voice-only condition scored lower on our empathic accuracy index, indicating greater accuracy ($M = 2.57$, $SD = 0.41$) than both the vision-only condition ($M = 2.74$, $SD = 0.56$; Fisher's LSD, $p = .019$; $d = .35$) and the multisense condition ($M = 2.71$, $SD = 0.49$; LSD, $p = .045$; $d = .31$). The vision-only and multisense conditions were not different from each other. These results provide evidence in support of our central hypothesis, and indicate that perceivers showed greater empathic accuracy for voice-only communication than for vision-only and multisense modalities.

Experiment 2

In the first experiment we found some initial support for our hypothesis that voice-only communication enhances empathic accuracy relative to communication across senses and vision-only communication. We sought to extend these results in a second study moving beyond passive viewing of emotions to estimating emotions in the context of a live interaction with a stranger. In Experiment 2, we manipulated the context of two interactions so that they occurred with voice-only communication versus communication across all senses using a novel paradigm where lighting was switched on or off in the interaction room.

Method

In Experiment 2, 266 strangers were paired into 133 dyads at the University of Illinois, Urbana–Champaign. The sample size exceeded our goal of 200 because we planned to collect data throughout the fall semester of the 2012–2013 academic year. The sample contained 171 women with the majority of the sample identifying as White ($n = 157$), followed by Asian American ($n = 67$), Latino/a ($n = 24$), Black ($n = 23$), and as other ethnic categories ($n = 16$). The mean age of the sample was 19.5 years ($SD = 2.30$). Dyads included 20 male–male dyads, 116 female–female dyads, and 55 mixed-gender dyads.

Participants were seated across from each other in the same laboratory room for a set of open social interactions on scripted topics. To describe the conditions, the experimenter instructed participants that the study was "a study focused on understanding how people communicate with each other about various topics and in unique social situations." Participants then engaged in a social interaction in two contexts in counterbalanced order: In the lighted room condition, participants engaged in a normal interaction with the room lights switched on. In the darkened room condition, participants engaged in the same interaction, except the lights were switched off and the door to the outside lab room was shut. Just prior to shutting the lights off for the darkened room interaction, the experimenter mentioned to participants, "For this interaction, we will be completely shutting off the lights." Participants were permitted to refuse the darkened room interaction if they wished although none of the participants declined this portion of the experiment. Two Foscam FI8910W (Foscam, Shenzhen, PRC) night-vision cameras positioned to the right of each participant captured the interactions. We chose this particular paradigm, rather than a dividing wall, because humans have many real-world opportunities to communicate emotions in darkness, and we were concerned that the dividing wall would create a sound barrier that might reduce voice communication as well as visual communication.

For the two social interaction tasks, participants discussed their choices and preferences in two domains: (a) film and TV and (b) food and beverages. These topics were chosen because they elicit a variety of emotions in undergraduate participants. The film and TV conversation always occurred first. Participants engaged in the interactions for three minutes each. Following each interaction, participants rated their own and their partner's emotions.

We indexed emotions using six emotions blends, which were assessed following each interaction: pride/accomplishment ($M_{dark} = 5.03$, $SD_{dark} = 1.01$; $M_{light} = 4.95$, $SD_{light} = 1.09$), shame/devaluation/guilt ($M_{dark} = 1.77$, $SD_{dark} = 1.18$; $M_{light} = 1.89$, $SD_{light} = 1.22$), anger/hostility/disgust ($M_{dark} = 1.59$, $SD_{dark} = 1.17$; $M_{light} = 1.59$, $SD_{light} = 1.11$), embarrassment/anxiety ($M_{dark} = 2.76$, $SD_{dark} = 1.54$; $M_{light} = 2.83$, $SD_{light} = 1.66$), happiness/joyfulness ($M_{dark} = 5.30$, $SD_{dark} = 1.06$; $M_{light} = 5.29$, $SD_{light} = 1.05$), and compassion/loving ($M_{dark} = 4.83$, $SD_{dark} = 1.13$; $M_{light} = 4.79$, $SD_{light} = 1.11$). Participants rated these emotions by completing the phrase, "Right now I feel" on 8-point Likert scales (0 = *disagree strongly*, 7 = *strongly agree*). Following these interactions and emotion ratings for self and partner, participants were debriefed about the hypotheses of the study.

Results

To test our central prediction that voice-only communication enhances empathic accuracy relative to communication across senses, we compared mean empathic accuracy scores of participants in the darkened and lighted interactions using a paired-samples *t* test. The analysis found that participants showed significantly greater mean empathic accuracy in estimating their partner's emotions in total darkness, where the voice was the primary mode of communication ($M = 1.05$, $SD = 0.60$) than they did in normal lighting, where other sensory modalities were available ($M = 1.15$, $SD = 0.64$), $t(265) = -2.69$, $p = .008$, Cohen's $d_{rm} = .16$.

We also tested our central hypothesis using a repeated-measures ANOVA at the dyad level. This analysis accounts for nonindependence in perceiver and partner empathic accuracy scores within each dyad (Kenny, Kashy, & Cook, 2006) as well as interaction order effects. Participant, identified based on room assignment (interaction room vs. adjacent room), and room condition (darkened vs. lighted) were the within subjects factors, the empathic accuracy score was the dependent variable, and order of interaction was the between subjects factor. The analysis revealed the predicted main effect of room condition, $F(1, 131) = 4.65$, $p = .033$, such that participants were more accurate recognizing emotions in the darkened ($M = 1.05$) than the lighted room condition ($M = 1.15$). The analysis yielded no significant effect of person, $F(1, 131) = 0.004$, $p = .950$, no effect of order, $F(1, 131) = 0.084$, $p = .773$, and no significant interactions, $Fs(1, 131) < 1.63$, $p > .20$. Importantly, these results could not be accounted for by condition differences in emotion expressivity as rated by a panel of objective coders or mean level differences in emotion self-reports (see supplementary analyses).

Experiment 3

Experiment 2 showed evidence consistent with our prediction that voice-only communication enhances empathic accuracy relative to communication across modalities, but the experiment could not completely address confounds due to condition differences in motivation or arousal engendered by the darkened relative to the lighted room interaction. In Experiment 3, we sought to rule out these alternatives using the interactions from Experiment 2 as perceptual stimuli.

Method

We analyzed data from 600 perceivers recruited online through Amazon Mechanical Turk. All participants were adults from the United States, and were paid $1 for taking part in the study. Two participants were excluded from analyses due to substantial missing data on emotion perception measures. All analyses are reported on the remaining sample of 598 participants. The majority of the sample was female ($n = 315$), the average age was 35.93 years ($SD = 11.9$). The majority of the sample was White ($n = 474$), followed by Asian American ($n = 41$), Latino/a ($n = 39$), Black ($n = 42$), and participants who listed themselves as an other ethnic category ($n = 13$). In this experiment we collected a sample much larger than our target sample size to more precisely estimate the effect size of the hypothesized voice-only enhancement of empathic accuracy effect. We did not analyze data until the full sample of participants was collected.

Perceivers were exposed to a series of social interactions and were instructed to estimate the emotions experienced by each target. Perceivers were then randomly assigned to one of three conditions: In the lighted room voice and visual condition, participants watched three separate videos showing female targets from Experiment 2 engage in the lighted interaction with a partner off camera. In the darkened room voice and visual condition, participants watched these same three female targets engage in the darkened room interaction from Experiment 2 with the same partner, with their nonverbal behavior visible via our night vision camera. In the voice-only condition, perceivers listened to the targets engage in the darkened room interaction from Experiment 2 with the same partner, without the ability to view their nonverbal behavior.

Experiment 2 targets were chosen as stimuli if they had large variability in their discrete emotion ratings across the darkened and lighted interactions. We chose high variance targets because we did not want perceiver accuracy to be artificially inflated by scale use artifacts related to entering similar responses across all positive or negative affect experiences (e.g., entering the same scale score for all positive emotions). Perceivers always viewed a woman interacting with a male partner so that the target of the emotion estimates could be easily identified. Order of targets was randomized and did not influence empathic accuracy (p = .882).

Perceivers estimated the emotions of targets using six discrete emotion dimensions: pride/accomplishment, shame/devaluation/guilt, anger/hostility/disgust, embarrassment/anxiety, happiness/joyfulness, and compassion/loving. Because the emotion composites in Experiment 2 might have forced participants to combine emotion terms they would not normally combine (e.g., anxiety and embarrassment), we also included separate indices of each emotion term in Experiment 3. Thus, Experiment 3 used three or four items, assessed on 8-point Likert scales (0 = *disagree strongly*, 7 = *strongly agree*) attached to the prompt, "Right now the person in the video feels" to assess perceptions of the experience of each discrete emotion composite. Pride/accomplished (M = 2.68, SD = 1.31) was indexed using three items (i.e., pride, accomplishment, and pride/accomplishment), shame/devaluation/guilt (M = 0.74, SD = 0.96) was indexed using four items (i.e., shame, devalued, guilty, and shame/devaluation/guilt), anger/hostility/disgust (M = 0.49, SD = 0.83) was indexed using four items (i.e., anger, hostile, disgust, and anger/hostility/disgust), embarrassment/anxiety (M = 2.28, SD = 1.33) was indexed using three items (i.e., embarrassment, anxiety, and embarrassment/anxiety), happiness/joyfulness (M = 3.07, SD = 1.34) was indexed using three items (i.e., happy, joyful, and happiness/joyfulness), and compassion/loving (M = 2.30, SD = 1.46) was indexed using three items (i.e., compassion, loving, and compassion/loving). All emotion composite scales showed high internal consistency across the three experimental conditions (αs = .81 to .99).

Results

Experiment 3 revealed a pattern of results aligning with our central hypotheses: A one-way ANOVA revealed significant group differences, $F(2, 594) = 6.34$, $p = .002$. An examination of means revealed that the voice-only condition showed lower scores reflective of greater empathic accuracy in comparison to the other two conditions employing both voice and visual modalities, LSD_{Dark} p = .004, d_{Dark} =.30; LSD_{Light} p = .001; d_{Light} = .30 (see Figure 2).

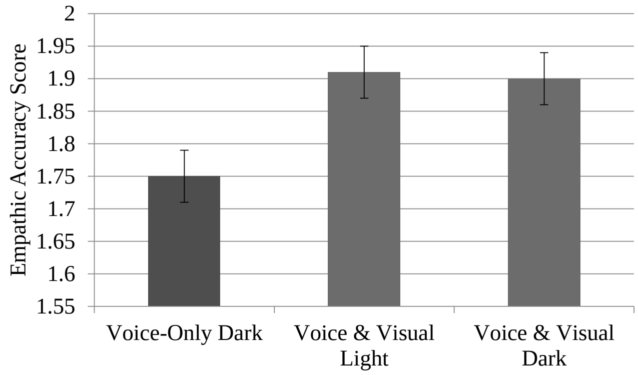

Figure 2. Empathic accuracy scores, with lower numbers indicating higher empathic accuracy, for participants in the voice-only darkened, voice and visual lighted, and voice and visual darkened room conditions, respectively. Lower scores indicate greater empathic accuracy and a score of zero indicates perfect accuracy. Error bars indicate 95% confidence intervals. See the online article for the color version of this figure.

Importantly, that the voice-only condition was significantly different from the darkened room voice and visual combined condition provides the strongest support for our central hypothesis that voice-only communication enhances empathic accuracy because in this comparison the content of the interactions was held completely constant.

Experiment 4

In Experiment 4, our goal was to generalize our empathic accuracy findings to real world contexts where voice-only modes of communication are common (i.e., the workplace). Experiment 4 also allowed us to assess the proposed central mediating mechanism—that voice-only communication allows individuals to focus more of their attention on the speech content of interactions rather than on peripheral nonverbal emotion cues that divide attention as well as distract from emotion perception. A final goal of Experiment 4 was to conduct confirmatory analyses of our central hypothesis using a preregistered research design.

Method

We analyzed data from 208 strangers paired into dyads. Participants were all recruited through the Behavioral Laboratory at the Yale University School of Management. The majority of the sample was female (n = 126), with 18 male–male dyads, 43 female–female dyads, and 40 mixed-gender dyads. The majority of the sample was White (n = 108), followed by Asian American (n = 50), Latino/a (n = 14), Black (n = 17), and participants who listed themselves as an other ethnic category (n = 25). All participants who took part in the study were included in analyses except in the case of three dyads (n = 6). Two of the dyads were removed because they were not strangers at the time of the interaction, and one dyad was removed because of a catastrophic failure to follow instructions that resulted in substantial missing data. The reported sample size ($n = 202$) excludes these six participants. All study materials, sample size, and data analytic strategy were preregistered at the Open Science Framework prior to data collection (https://osf.io/7qepz/).

Participants arrived at the experiment and were seated in separate rooms for consent procedures. Following consent and baseline measures used in a larger study, participants were provided a tablet (i.e., an Apple iPad) where they would have two 3-min interactions with a stranger over the Zoom conferencing platform (www.zoom.us). Each interaction was about a workplace challenge. The first conversation asked participants to talk about their last experience in a workplace team and the emotions that this experience engendered. The second conversation examined the challenges of forming close relationships at work. These two conversations occurred with (a) voice-only communication or (b) with both voice and video communication enabled (counterbalanced). Following the conversations, participants rated their own and their partner's emotions during the interaction. We used 10 emotions in the study rated on an 8-point Likert scale (0 = *not at all*, 8 = *a great deal*) with respect to self and partner: joy, fear, sadness, shame, embarrassment, fear, contempt, disgust, pride, and compassion.

After each conversation, participants rated what percentage of their attention they used for each of 12 specific emotion cues during the conversation with their partner. The emotion cues were presented randomly and included facial expressions, vocal pitch, vocal volume, word choice, eye contact/movements, head tilts, vocal speed, vocal cadence (stopping and starting of speech), eyebrow movements, and speech content. The sum total across all emotion cue items had to equal 100 across the 12 items, and participants could choose to report attending to as few as one cue or as many as all 12 during the interaction. We were interested in participant reports of how much they attended to "facial expressions" ($M = 14.63$, $SD = 9.60$), "speech content" ($M = 16.79$, $SD = 13.96$), and the average composite of vocal speed, cadence, pitch, and volume—our index of attention to paralinguistic and linguistic vocal cues ($M = 10.45$, $SD = 3.54$).

Results

For our preregistered analysis, we used the same dyad-level repeated-measures ANOVA accounting for nonindependence in perceiver and partner empathic accuracy scores and the order of voice-only and voice and visual conferencing conditions as in Experiment 2. The results from this analysis aligned with our predictions: A significant effect of communication condition emerged such that voice-only communication ($M = 1.10$) elicited heightened empathic accuracy relative to voice and visual communication ($M = 1.24$), $F(1, 99) = 5.10$, $p_{\text{one-tailed}} = .013$. The analysis yielded no significant effect of person, $F(1, 99) = 0.95$, $p = .332$, no effect of order, $F(1, 99) = 0.114$, $p = .736$, and no significant interactions, $Fs(1, 99) < 1$, $ps > .30$, save for one: between condition and order, $F(1, 99) = 12.94$, $p = .001$. An examination of means revealed that the voice-only condition ($M_{\text{voice}} = 1.02$) elicited greater empathic accuracy than the voice and visual condition ($M_{\text{voice/visual}} = 1.36$) particularly when the voice-only condition came second. When the voice-only condition came first, the two conditions showed equivalent levels of empathic accuracy ($M_{\text{voice}} = 1.19$; $M_{\text{voice/visual}} = 1.11$). We caution interpretation of this interaction because it was unanticipated and did not emerge in the identical analysis in Experiment 2. As in Experiment 2, we also used a paired-samples *t* test to examine our central hypothesis and this analysis observed a similar pattern as the prior studies. Participants in the voice-only condition ($M = 1.10$, $SD = 0.66$) were more accurate in reading emotions than were participants in the combined voice and visual condition ($M = 1.24$, $SD = 0.72$), $t(201) = -2.66$, $p = .008$, $d_{\text{rm}} = .20$.

In follow-up exploratory analyses we examined the extent that reported attentional focus on facial expressions, speech content, and linguistic and paralinguistic vocal cues explained why voice-only communication enhances empathic accuracy relative to communication across senses. To test this prediction we subjected our attention focus questions to a 2 (Communication Modality) × 3 (Speech Content/Facial Expression/Vocal Cues) repeated-measures ANOVA. The analysis revealed a significant effect of modality, $F(1, 201) = 40.63$, $p < .001$, and no effect of attention focus, $F(2, 402) = 2.56$, $p = .078$. Both of these results were qualified by our predicted interaction effect, $F(2, 402) = 241.28$, $p < .001$. Examination of means revealed, not surprisingly, that participants paid significantly more attention to the speech content and vocal cues of speech in the voice-only modality than they did to facial expressions (see Figure 3). Supporting our perspective on divided attention, having the extra visual information available in the voice and visual combined condition relative to the voice-only condition actually reduced participant attentiveness to the speech content of the interactions by more than six percentage points, $t(201) = 6.11$, $p < .001$, $d_{\text{rm}} = .44$. Also in line with our predictions, the combined voice and visual condition reduced attention to vocal cues by an even larger margin (7.91%), $t(201) = 20.13$, $p < .001$, $d_{\text{rm}} = 1.46$.

We next explored the potential mediating role that attention to specific aspects of the voice plays in enhancing empathic accuracy in voice-only communication relative to communication across senses. For this analysis we used the MeMore macro developed to assess statistical mediation in within subjects designs (e.g., Montoya & Hayes, 2016). We calculated a parallel mediation analysis with the difference

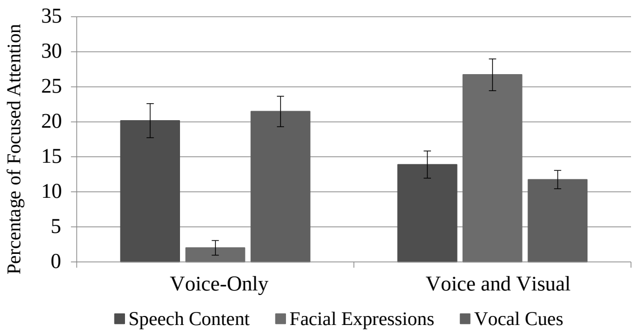

Figure 3. Participant self-reports of focused attention during the interactions across chat modality. Participants reported focusing more on the speech content and vocal cues of interactions than on facial expressions when audio was the only mode of communication. In contrast, participants reported focusing more on facial expressions than speech or vocal content when both audio and video were available. All focused attention domains differed between conditions. Speech content and vocal cue attention did not differ from each other. Error bars indicate 95% confidence intervals surrounding the mean. See the online article for the color version of this figure.

between empathic accuracy in the voice-only versus combined senses conditions as the outcome and condition differences in attention to speech content and vocal cues as parallel mediators. We explored whether improvements in empathic accuracy in voice-only communication relative to communication across senses were accounted for by condition differences in attention to speech content, to vocal cues, or to both, using a boostrapping procedure with 5,000 resamples. The analysis yielded a significant indirect effect of attention to vocal cues on condition differences in empathic accuracy, $b = .21$, $t(197) = 2.41$, $p = .017$; 95% confidence interval [.01, .40]. Interestingly, attention to speech content did not yield a significant indirect effect on condition differences in empathic accuracy, $b = .001$, $t(197) = 0.27$, $p = .79$; 95% confidence interval [–.05, .05]. Although mediation analyses such as these are preliminary and correlational (e.g., Smith, 2012), they provide some initial evidence that voice-only communication enhances empathic accuracy relative to communication across senses because it increases attention to vocal cues in speech that convey emotion. This finding fits with prior work indicating that vocal cues play a crucial role in empathic accuracy (Hall & Schmid Mast, 2007; Zaki et al., 2009) because they are particularly likely to leak authentic internal emotional states during interactions (Ambady & Rosenthal, 1992; DePaulo & Rosenthal, 1979).

Experiment 5

In our final experiment we sought to build on the initial mediation evidence in Experiment 4, suggesting that attention to vocal cues explains why voice-only communication enhances empathic accuracy relative to communication across senses. To accomplish this we used the same passive emotion perceiver paradigm and stimuli from Experiment 1, only in this instance we added a fourth experimental condition—in addition to all senses, voice-only, and visual-only conditions, we added a second digital voice-only condition. For this digital voice condition participants listened to a monotonic digital voice (e.g., Siri from the Apple iPhone) read a transcript of the teasing interaction from the voice-only condition. This condition contains all the speech content of the original voice-only condition but removes all linguistic and paralinguistic vocal cues. We expected voice-only communication to enhance empathic accuracy relative to all other conditions. Aligning with our mediation account, we expected that the absence of linguistic and paralinguistic vocal cues would reduce empathic accuracy relative to all other conditions.

Method

In Experiment 5, we analyzed data from 427 perceivers recruited online through Amazon Mechanical Turk. Participants were paid $1.50 for taking part in the study. To enhance data quality we used a new audio attention check item that allowed us to remove 21 participants from this initial sample who responded as if they did not have audio capability, yielding a final sample of 406 participants. The majority of the sample was female ($n = 219$), the average age was 37.60 years. The majority of the sample was White ($n = 307$), followed by Asian American ($n = 30$), Latino/a ($n = 30$), Black ($n = 40$), and participants who listed themselves as an other ethnic category ($n = 23$).

Perceivers estimated emotions for the six targets from Experiment 1, all within three separate friendship dyads engaging in teasing. Perceivers estimated the same 23 discrete emotion words filled out by targets at the time of the teasing interaction using 9-point Likert scales (0 = *not at all*, 8 = *a great deal*). The emotions were amusement, anger, anxiety, appreciation, compassion, contempt, curiosity, discomfort, disgust, embarrassment, empathy, enthusiasm, fear, guilt, gratitude, happy, hope, irritation, joy, love, pride, sad, and shame. Perceivers viewed the emotion rating scale after the first teasing exchange. The order of teasing interactions was randomized and we observed no order effects ($p = .413$).

In Experiment 5, we used a between subjects manipulation of communication modality. We assigned a fourth of participants to perceive the emotions of the interacting friends with voice-only communication enabled, vision-only communication enabled, with multiple sensory modalities enabled, and with monotonic digital voice-only enabled (for sample digital voice-only stimuli, go to https://osf.io/eqvz6/) created with text to digital voice translation software called Natural Reader.

Results

We analyzed empathic accuracy scores using a one-way ANOVA and then probed conditions differences using Fisher's LSD test. The ANOVA yielded a significant omnibus effect, $F(3, 402) = 44.07$, $p < .001$. First, we examined the replication of our original Experiment 1 findings, that voice-only communication would enhance empathic accuracy relative to visual-only and combined modalities. Indeed, voice-only communication elicited heightened accuracy ($M = 2.41$, $SD = 0.35$) than both the vision-only condition ($M = 2.73$, $SD = 0.34$; LSD, $p < .001$; $d = .90$) and the multisense condition ($M = 2.54$, $SD = 0.37$; LSD, $p = .014$; $d = .36$), a finding consistent with the results from Experiment 1.

Next we examined the results from the digital voice-only condition in comparison to the other three conditions. As predicted, digital voice-only communication showed the lowest levels of empathic accuracy ($M = 2.97$, $SD = 0.41$), eliciting lower accuracy than all conditions ($\text{LSD}_{\text{voice}}$ $p <$ .001; $d_{\text{Voice}} = 1.47$; $\text{LSD}_{\text{Combined}}$ $p < .001$; $d_{\text{Combined}} = 1.10$; $\text{LSD}_{\text{Visual}}$ $p < .001$; $d_{\text{Visual}} = .64$). These latter results provide additional evidence for our mechanistic account of the advantages of empathic accuracy in voice-only communication. Specifically, linguistic and paralinguistic vocal cues are critical for empathic accuracy—without these cues perceivers had the most difficulty perceiving emotions in interactions.

General Discussion

Social and biological sciences have demonstrated both the profound desire of individuals to connect with others (Bowlby, 1988), and the array of skills people possess to accurately discern other's emotions and intentions. And yet, in the presence of both will and skill to communicate, people often inaccurately perceive others' emotions. In the present research we examined the possibility that less is more—that voice-only communication, even though it involves only one of the modalities of emotion expression, will significantly improve empathic accuracy over communication across multiple senses. Results from five experiments support this view. This boost in empathic accuracy for voice-only communication occurred when compared to vision-only and across senses communication, and persisted when examining different kinds of emotional discussions, different ways of assessing self- and perceiver-reports of emotion, and even when the actual content of the interactions was held constant. Critically, follow-up correlational and experimental evidence supports the role of attention to linguistic and paralinguistic vocal cues in the enhancement of empathic accuracy in voice-only communication relative to accuracy across modalities. Overall, the findings align with a broader literature which finds that vocal cues are more critical to accurate emotion recognition than are facial cues of emotion (Gesn & Ickes, 1999; Hall & Schmid Mast, 2007; Zaki et al., 2009).

Broadly, these findings have direct applications to the ways in which both research and practice rely on facial cues of emotion to assess others' internal states—where much early work has focused (e.g., Ekman, 1989). The current research suggests that relying on a combination of vocal and facial cues, or solely facial cues, may not be the best strategy for accurate emotion recognition—particularly because facial expressions can sometimes be inconsistent with internal states or used to actively dissemble (e.g., Porter & ten Brinke, 2008).

Practically, these findings raise a number of questions about research on emotions, including if evidence for cross-cultural emotion recognition would reach similar conclusions if emotions were examined in other modalities aside from facial expressions (e.g., Gendron et al., 2014; Russell, 2003). It is also interesting to consider these findings in the context of policies aimed at training security and law enforcement in lie detection (e.g., Ekman, 2009). Although it is important to note that these data were collected where there were no incentives for lying, the enhanced accuracy in voice-only communication observed in this research suggests the voice domain as a potential area of future inquiry for understanding lie detection, as others have argued (Porter & ten Brinke, 2008).

We also report analyses only using mean level bias as our metric of empathic accuracy. In other work empathic accuracy can also be computed as a profile correlation between individual emotion self-reports and perceiver estimates. While we do not favor this approach because correlations are unstable without hundreds of observations (Schonbrodt & Perugini, 2013), we did compute profile correlations as an additional exploratory way to test our hypothesis that voice-only communication enhances empathic accuracy (see supplementary analyses). To summarize this analysis across ($k = 5$) studies, we calculated a combined Z score across studies using a fixed effects model with the effect sizes observed in each of the experiments (Goh, Hall, & Rosenthal, 2016). The result is a significant overall effect of voice-only communication in enhancing empathic accuracy relative to combined communication across the face and voice (Stouffer's combined $Z = 3.19$, $p = .001$). Thus, this exploratory supplementary analysis using profile correlations also supports our central hypothesis.

Although a growing body of work implicates the voice in emotion recognition, this research is the first, in our estimation to find evidence suggesting that voice-only communication enhances empathic accuracy above that observed in communication across senses. In two prior studies, no differences in empathic accuracy between voice-only and combined communication modalities were observed (Gesn & Ickes, 1999; Hall & Schmid Mast, 2007), whereas in one study voice-only communication elicited lower accuracy that communication across senses (Zaki et al., 2009). The discrepancy between these results and the current work may have occurred because the current work used larger samples than the two null studies (total $n = 72$ for Gesn & Ickes, 1999; highest cell $n = 24$ for Hall & Schmid Mast, 2007). Given the observed small to medium effect size of heightened empathic accuracy in voice-only communication relative to that across senses, the samples in these two prior studies may have been too small to detect an effect. It is also possible that the target stimuli developed in each of the prior studies may have artificially inflated links between target self-reports and nonverbal behaviors: In each of the prior studies, target self-reports were generated by having targets rate their emotions during a viewing of a video of themselves engaging in an interaction (e.g., Zaki et al., 2009), rather than retrospectively assessing their emotions as in the current research. It is possible that watching the interactions may have effectively forced targets to reconcile any discrepancies between their observed behavior and mental states (e.g., Festinger & Carlsmith, 1959), thereby inflating empathic accuracy in the combined senses modality. A future study that manipulates the channel of emotion targets use to rate their own emotions could directly test this prediction.

Limitations of this research and the future directions they suggest should also be noted. Participants were exclusively from the United States, constraining the generalizeability of the findings. It is possible, for instance, that cultures with stricter emotion display rules (e.g., Japanese cultural contexts) might show even higher empathic accuracy in voice-only communication in comparison to combined communication settings (Safdar et al., 2009). Intimacy is also an interesting potential moderator: It is possible that the advantage in empathic accuracy enjoyed by voice-only communication in the present work is confined only to stranger interactions, given the expertise in facial recognition of emotion that develops from increased contact. As well, prior work suggests that incentives enhance empathic accuracy particularly for attending to vocal cues (Hall, Andrzejewski, & Yopchick, 2009), raising the possibility that providing incentives might reduce the observed differences in empathic accuracy between voice-only and combined senses communication.

In this research, we have extended our understanding of the human capacity to read emotions into contexts involving the voice only. Reading the minds of others is simultaneously crucial for a well-functioning social life and a significant challenge for individuals. These findings suggest paradoxically that understanding others' mental states and emotions relies less on the amount of information provided, and more on the extent that people attend to the information being vocalized in interactions with others.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274. http://dx.doi.org/10.1037/0033-2909.111.2.256

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 42,* 241–251. http://dx.doi.org/10.1111/1469-7610.00715

Bowlby, J. (1988). *A secure base: Parent–child attachment and healthy human development*. New York, NY: Basic Books.

Côté, S., & Miners, C. T. (2006). Emotional intelligence, cognitive intelligence, and job performance. *Administrative Science Quarterly, 51,* 1–28.

DePaulo, B. M., & Rosenthal, R. (1979). Telling lies. *Journal of Personality and Social Psychology, 37,* 1713–1722. http://dx.doi.org/10.1037/0022-3514.37.10.1713

Ekman, P. (1989). The argument and evidence about universals in facial expressions. *Handbook of social psychophysiology,* 143–164.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (rev. ed.). New York, NY: W. W. Norton & Company.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology, 17,* 124–129.

Ekman, P., & Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS*). New York, NY: Oxford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203–210. http://dx.doi.org/10.1037/h0041593

Fletcher, G. J., & Kerr, P. S. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin, 136,* 627–658. http://dx.doi.org/10.1037/a0019792

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science, 25,* 911–920. http://dx.doi.org/10.1177/0956797613517239

Gesn, P. R., & Ickes, W. (1999). The development of meaning contexts for empathic accuracy: Channel and sequence effects. *Journal of Personality and Social Psychology, 77,* 746–761. http://dx.doi.org/10.1037/0022-3514.77.4.746

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10,* 535–549. http://dx.doi.org/10.1111/spc3.12267

Hall, J. A., Andrzejewski, S. A., & Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *Journal of Nonverbal Behavior, 33,* 149–180. http://dx.doi.org/10.1007/s10919-009-0070-5

Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin, 131,* 898–924. http://dx.doi.org/10.1037/0033-2909.131.6.898

Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion, 7,* 438–446. http://dx.doi.org/10.1037/1528-3542.7.2.438

Harlow, H. F. (1958). The nature of love. *American Psychologist, 13,* 673–685. http://dx.doi.org/10.1037/h0047884

Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of*

Behavioral Medicine, 40, 218–227. http://dx.doi.org/10.1007/s12160-010-9210-8

Hertenstein, M. J., Keltner, D., App, B., Bulleit, B. A., & Jaskolka, A. R. (2006). Touch communicates distinct emotions. *Emotion, 6,* 528–533. http://dx.doi.org/10.1037/1528-3542.6.3.528

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59,* 730–742. http://dx.doi.org/10.1037/0022-3514.59.4.730

Keltner, D., Capps, L., Kring, A. M., Young, R. C., & Heerey, E. A. (2001). Just teasing: A conceptual analysis and empirical review. *Psychological Bulletin, 127,* 229–248. http://dx.doi.org/10.1037/0033-2909.127.2.229

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.

Kraus, M. W., Côté, S., & Keltner, D. (2010). Social class, contextualism, and empathic accuracy. *Psychological Science, 21,* 1716–1723. http://dx.doi.org/10.1177/0956797610387613

Kraus, M. W., Horberg, E. J., Goetz, J. L., & Keltner, D. (2011). Social class rank, threat vigilance, and hostile reactivity. *Personality and Social Psychology Bulletin, 37,* 1376–1388. http://dx.doi.org/10.1177/0146167211410987

Kraus, M. W., Park, J. W., & Tan, J. J. X. (in press). Signs of social class: The experience of economic inequality in everyday life. *Perspectives on Psychological Science*.

Lewis, K. L., Hodges, S. D., Laurent, S. M., Srivastava, S., & Biancarosa, G. (2012). Reading between the minds: The use of stereotypes in empathic accuracy. *Psychological Science, 23,* 1040–1046. http://dx.doi.org/10.1177/0956797612439719

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science, 19,* 898–903. http://dx.doi.org/10.1111/j.1467-9280.2008.02174.x

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist, 63,* 503–517. http://dx.doi.org/10.1037/0003-066X.63.6.503

Meltzoff, A. N. (1985). Immediate and deferred imitation in fourteen-and twenty-four-month-old infants. *Child Development, 56,* 62–72.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Pt. 1. Basic mechanisms. *Psychological Review, 104,* 3–65. http://dx.doi.org/10.1037/0033-295X.104.1.3

Montoya, A. K., & Hayes, A. F. (2016). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*. Advance online publication. http://dx.doi.org/10.1037/met0000086

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18,* 566–570. http://dx.doi.org/10.1016/j.tics.2014.09.007

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77,* 1296–1312. http://dx.doi.org/10.1037/0022-3514.77.6.1296

Porter, S., & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19,* 508–514. http://dx.doi.org/10.1111/j.1467-9280.2008.02116.x

Richard, F. D., Bond, Jr., C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7,* 331–363.

Richards, J. M., Butler, E. A., & Gross, J. J. (2003). Emotion regulation in romantic relationships: The cognitive consequences of concealing feelings. *Journal of Social and Personal Relationships, 20,* 599–620. http://dx.doi.org/10.1177/02654075030205002

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763–797. http://dx.doi.org/10.1037/0096-1523.27.4.763

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115,* 102–141.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. http://dx.doi.org/10.1037/0033-295X.110.1.145

Safdar, S., Friedlmeier, W., Matsumoto, D., Yoo, S. H., Kwantes, C. T., Kakai, H., & Shigemasu, E. (2009). Variations of emotional display rules within and across cultures: A comparison between Canada, USA, and Japan. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 41,* 1.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47,* 609–612. http://dx.doi.org/10.1016/j.jrp.2013.05.009

Simon-Thomas, E. R., Keltner, D. J., Sauter, D., Sinicropi-Yao, L., & Abramson, A. (2009). The voice conveys specific emotions: Evidence from vocal burst displays. *Emotion, 9,* 838–846. http://dx.doi.org/10.1037/a0017810

Smith, E. R. (2012). The beginning of a new editorial term at the Attitudes and Social Cognition section of the Journal of Personality and Social Psychology (JPSP) is a natural occasion for reflection on this journal's mission: Introduction. *Journal of Personality and Social Psychology, 102,* 1–3. http://dx.doi.org/10.1037/a0026676

Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology, 62,* 787–797. http://dx.doi.org/10.1037/0022-3514.62.5.787

Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2016). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000226

Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion, 9,* 478–487. http://dx.doi.org/10.1037/a0016551

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2,* 1497–1517. http://dx.doi.org/10.1111/j.1751-9004.2008.00109.x

Received August 3, 2016
Revision received February 1, 2017
Accepted February 12, 2017 ■