Encoding and Decoding Hidden Meanings in Face-to-Face Communication: Understanding the Role of Verbal and Nonverbal Behaviours in Indirect Replies

Mingyuan Chu[1], Paul Tobin[1], Flora Ioannidou[2] and Jana Basnakova[3]

[1] University of Aberdeen, Aberdeen, UK

[2] University of Dundee, Dundee, UK

[3] Slovak Academy of Sciences, Bratislava, Slovakia

Author Note

Mingyuan Chu and Paul Tobin contributed equally to this work. Paul Tobin's current affiliation is Smart Innovation Norway, Halden, Norway. Jana Basnakova's current affiliation is Radboud University, Nijmegen, The Netherlands.

This research was supported in part by the British Academy/Leverhulme Small Research Grants SG151724 to Mingyuan Chu. We thank Michelle Sader, Rachael Greenhill, Pinja Suonpaa, Hannah Cannaux and Duan Wang for their help with data collection, Annabelle Olsson for her help with reliability check of nonverbal behaviour coding, and Yu Ho Wong for her help with proof reading of this article.

The raw data, SPSS analyses and the R scripts used for the generalized linear mixed models analyses can be found on the Open Science Framework: https://osf.io/ngsp6/ (Chu et al. 2022). Preliminary findings from this study were presented in an oral presentation session at the Experimental Psychology Society Meeting in St. John's, Canada (July 4-7, 2018).

Correspondence concerning this article should be addressed to Mingyuan Chu, School of Psychology, University of Aberdeen, UK, AB24 3FX, E-mail: mingyuan.chu@abdn.ac.uk#

Abstract

In everyday conversation, we often use indirect replies to save face of our interlocutor (e.g., "Your paper does have room for improvement"). Six experiments were conducted to examine the role of verbal and nonverbal behaviours in the production and comprehension of indirect replies. In Experiments 1a and 1b, participants engaged in question-answer exchanges designed to elicit four types of replies (i.e., direct, indirect, lie, and neutral). Results showed that uncertainty terms, discourse markers and head tilt were most uniquely associated with the production of indirect replies. In Experiments 2a, 2b, 3a and 3b, participants categorized the types of replies in video clips of real participants in Experiments 1a and 1b. Results showed that nonverbal behaviours enhanced the performance and boosted the confidence in the identification of indirect replies. Furthermore, uncertainty terms, discourse markers and head tilt were also the most reliable cues for identifying indirect replies. Finally, the extent to which people relied on verbal and nonverbal cues to identify an indirect reply was context dependent. The more informative the verbal/nonverbal information was, the fewer nonverbal/verbal cues that contributed to the identification of indirect replies. Our results demonstrated that people integrate verbal and nonverbal information to enhance their understanding of the intended meaning in indirect replies. Findings from the current research provide an initial step towards developing a comprehensive and unified model of the production and comprehension of indirect replies, which takes both verbal and nonverbal behaviours into account.

Keywords: indirect reply, verbal and nonverbal cues, head tilt, uncertainty term, discourse marker

Encoding and Decoding Hidden Meanings in Face-to-Face Communication: Understanding the Role of Verbal and Nonverbal Behaviours in Indirect Replies

Communication is essential in our social lives and serves as a building block in forming and maintaining relationships (Lakin et al., 2003). In everyday conversation, we often use indirect speech to convey our intention (Pinker et al., 2008). According to the speech act theory (Searle, 1975), indirect speech is when the sentence meaning is different from the intended meaning. For example, when saying “It’s really hot in this room.” to someone who is sitting next to the window, the true communicative intent is to request the hearer to open the window. Thus, effective communication requires encoding and decoding the hidden messages behind the surface meanings.

In daily communication a major motivation for using indirect speech is face management (Clark & Schunk, 1980; Holtgraves, 1998, 1999; Holtgraves & Yang, 1990; 1992). According to Brown and Levinson's influential politeness theory (Brown & Levinson, 1987), face is fragile and subject to constant threat during social interaction and humans have a need for maintaining their own faces as well as attending to others’ face. Failures in face management can lead to social interaction difficulties, misunderstandings, and even conflict (Limberg, 2009). Therefore, when delivering face-threatening information, speakers often use indirect speech in order to avoid criticism, disapproval or blame in a conversation. For example, after leaving a house party, if John asked Emma “Did they like my joke?”, Emma would avoid a blunt “No,” but say “Well, I think they probably have a different sense of humour to you,” instead. In this example, Emma used an indirect reply to suggest the jokes were not funny, but she did it in such a way that would save John’s face.

One frequently used face management strategy in indirect replies is uncertainty expressions. According to the politeness theory (Brown & Levinson, 1987), uncertainty expressions can soften a negative message by sacrificing accuracy for ambiguity. In a corpus

analysis, Youmans (2001) analysed the use of probability terms (e.g., maybe, could) in conversations from American English speakers. It was found that 39% of the probability terms were used for face-saving purposes rather than to communicate likelihood. Furthermore, there has been experimental evidence showing that uncertainty expressions are closely linked to the delivery and interpretation of face-threatening messages in conversations. For example, Juanchich and Sirota (2013) asked participants to communicate negative and positive outcomes to a partner either plainly or in a face-saving manner. Participants chose to use more uncertainty expressions in the face-saving condition than in the plain condition only when communicating negative outcomes.

Uncertainty expressions are not only used by the speakers when producing indirect speech, they are also used by the listener as a cue of indirectness when comprehending indirect speech. Bonnefon and Villejoubert (2006) asked participants to imagine that their doctor had informed them that they had 'possibly' developed one of two medical conditions, such as deafness or insomnia. Participants were then asked whether the doctor used the word "possibly" because she was not sure of the diagnosis or because she was being tactful. Results showed that the more severe the condition of deafness, participants were more likely to interpret that the probability term served for a face-saving purpose rather than for communicating likelihood. This suggests that people treat probability terms as a cue of indirectness.

In addition to probability terms, hedging terms (e.g., a little bit, sort of) and discourse markers (e.g., you know, I mean), can also be used for face management purposes. For example, Fraser (2010) found that hedging terms can serve a face-saving purpose by signalling a lack of commitment to what is said. Similarly, Tree and Schrock (2002) found that when using discourse markers, speakers can make their speech more casual and distance themselves from potentially face-threatening remarks. Thus, both hedging terms and

discourse markers can be used to express uncertainty and increase the politeness of a message. In addition to hedging terms and discourse markers, filled pauses (e.g., ‘um’, ‘uh’) and silent pauses are also used as markers of uncertainty in conversations (e.g., Brennan & Williams, 1995; Siegman & Pope, 1965; Smith & Clark, 1993; Swerts & Krahmer, 2005), although their roles in face management remains unclear.

Most existing research on indirect communication has focused predominantly on the verbal elements, whereas the role of nonverbal cues has been largely ignored (Bonnefon, Feeney & De Neys, 2011). Human communication is multi-modal in nature (Chovil, 1991; McNeill, 1998). When communicating with each other, we do not only speak but also use nonverbal behaviours, such as eye gaze, facial expression, hand gesture, and body posture (Kendon, 1997). Nonverbal behaviours account for a significant proportion of what we communicate during social interaction and play a vital role in revealing intentions (Argyle, 1975; Bates, 1976; Clark, 1996; McNeill, 1985, 1987, 1992). Considering that indirect communication requires encoding and decoding the hidden communicative intent behind the surface meaning of speech, nonverbal cues might play a crucial role in the production and comprehension of indirect messages. For example, it has been shown that the understanding of indirect requests (e.g., “It’s getting hot in here”, indicating a request to open the window) can be facilitated by a pointing gesture and eye gaze (Barr & Kelly, 1997; Kelly et al., 1999; Kelly, 2001).

So far, our understanding of the role of nonverbal cues in indirect replies remains extremely incomplete. We propose that nonverbal behaviours that are used to express uncertainty may play a vital role in the production and comprehension of indirect replies, as previous studies revealed that uncertainty is a key component of face-saving indirect communication (Bonnefon, et al., 2009; Holtgraves, 2008). Based on the existing research, the current study will focus on four types of nonverbal behaviours that have been identified

as markers of uncertainty, that is, palm-revealing gesture, head tilt, facial shrug and gaze aversion. These nonverbal cues may be used in indirect replies for the purpose of saving face.

In a palm-revealing gesture, the speaker shows an empty hand to the listener to indicate being uncertain of what to say (Bavelas et al., 1992; Kendon, 2004). This gesture is sometimes accompanied by a shoulder shrug. McNeill (1992) observed that the palm-revealing gesture is used when expressing uncertainty and often accompanies expressions, such as “I don’t know”, “it’s like...”, “well…”, “kind of…”. It is worth pointing out that gestures with a palm-revealing form can also serve other interactive functions. For example, the conduit gesture (Chu et al., 2014) is used when a speaker moves an open palm toward the listener as if to present a clearly formulated idea on the palm to the listener. The palm-up open hand gesture (Muller, 2004) can be used for pleading, requesting, and expressing openness to an imaginary object. However, in the current study, we only focused on the palm-revealing gesture that is used to express uncertainty.

The link between head tilt and the expression of uncertainty was demonstrated in Marono et al. (2018). In this study, the authors investigated the use of nonverbal cues in deception with real-life video recordings of politicians, criminals and people of high interest from various media sources. They found that head tilt was used when people were uncertain about their answers and needed further consideration and reflection.

A facial shrug usually consists of a retraction of the mouth corners (Chovil, 1991; Stone & Oh, 2008) and the raising of the cheek areas and the eyebrows (Debras, 2017). It has been shown that facial shrug is used when people communicate uncertainty such as “I don’t know”, “I’m not sure” and “I’m thinking about it” (Bitti et al., 2014; Chovil, 1991).

With regards to gaze aversion, Beattie (1978) found that speakers were more likely to avert their gaze in hesitant speech than in pre-planned speech. Similarly, Cegala et al. (1979)

found that gaze aversion occurs alongside discourse markers that are associated with uncertainty, such as “you know” and “like” as well as during pauses in speech.

In addition to the lack of understanding of the role of nonverbal cues in indirect replies, little is known about the relative contribution of verbal and nonverbal cues in indirect replies. Traditionally, it has often been assumed that information conveyed by nonverbal behaviour is only used as a last resort, for example, when speech is vague, inadequate, or absent (Vigliocco et al., 2014). In contrast, some researchers such as McNeill (1992), Clark (1996) and Kendon (1994) hold that speech and nonverbal behaviours together create a deeper and more meaningful message than just speech alone. Kelly et al. (2010) provided supporting evidence for the latter view by demonstrating that the integration of gesture and speech is obligatory during language comprehension. People cannot help but consider information conveyed by gesture when processing speech. Similarly, de Gelder and Vroomen (2000) demonstrated that the integration of affective information conveyed through the face and voice is mandatory during emotion identification. However, these two views may not be mutually exclusive. It is possible that listeners do integrate verbal and nonverbal information automatically to enhance their understanding of the intended meaning in indirect replies, but the relative contribution of the verbal and nonverbal cues may vary depending on the availability and informativeness of each channel.

Another important issue that remains poorly understood concerns the production-comprehension link of verbal and nonverbal behaviours in indirect replies. Most existing studies have investigated the production and comprehension of indirect speech separately. Thus, it remains unclear whether the verbal and nonverbal cues that are produced in indirect replies are also used as cues during the comprehension of indirect replies. In recent years, there has been an effort to build unified models of language production and comprehension (e.g., Dell & Chang, 2014; Chater et al., 2016; Pickering & Garrod, 2007, 2013). According

to these unified models, language production and comprehension are a single system and language comprehension relies on the production system to predict what the speaker is about to say. In contrast, some researchers disagree with the unified models, and hold that language production and comprehension are separable processes, and prediction can help but is not necessary for language comprehension (e.g., Huettig & Mani, 2016; Mani & Huettig, 2014; Mishra et al., 2012). Regarding nonverbal behaviours, there has been little research on their production-comprehension link. However, some theoretical accounts argued that comprehending meanings of nonverbal behaviours relies on the production system to simulate the perceived nonverbal behaviours (e.g., Alibali & Hostetter, 2010; Holle et al., 2008; Niedenthal et al., 2010).

Although the current study is not designed to test these competing hypotheses, our findings could provide useful insight into this debate. According to the integrated framework, the same verbal and nonverbal cues should be used both in the production and comprehension of indirect speech. In contrast, if the language production and comprehension processes were separable, different verbal and nonverbal cues might be used in these processes.

The Current Research

The overall aim of the current research is to provide a comprehensive understanding of the role of verbal and nonverbal behaviours in the production and comprehension of indirect replies. The verbal cues examined in the current study were the probability and hedging terms, discourse markers, and filled and silent pauses, and the nonverbal cues were palm-revealing gesture, head tilt, facial shrug and gaze aversion.

The first two experiments investigated the verbal and nonverbal cues used in the production of spontaneous (Experiment 1a) and scripted (Experiment 1b) indirect replies. Participants engaged in a face-to-face interaction which elicited either an indirect reply or other types of replies (i.e., direct, lie, and neutral). Verbal and nonverbal behaviours were

coded and compared between different types of replies to identify the verbal and nonverbal cues that are uniquely associated with the production of indirect replies.

Experiments 2a and 2b examined the independent contribution of nonverbal cues in the comprehension of indirect replies. Participants first viewed silent video clips of four types of replies (direct, indirect lie or neutral) collected in Experiments 1a and 1b. They were then asked to categorize the type of replies given in each video clip based on the nonverbal cues alone. Correlational and multiple regression analyses were conducted to determine the nonverbal cues that people used to identify indirect replies.

In Experiments 3a and 3b, participants first viewed clips of indirect and neutral replies collected in Experiments 1a and 1b either with both video and audio or with audio only. They were then asked to categorize the type of replies in each clip. To test whether nonverbal cues can facilitate the comprehension of indirect replies, participants' categorization performance was compared between the video-and-audio condition and the audio-only condition. Furthermore, correlational and multiple regression analyses were conducted to examine the relative contribution of verbal and nonverbal cues in the comprehension of indirect replies.

Open Science Statement

The raw data, SPSS analyses and the R scripts used for the generalized linear mixed models analyses can be found on the Open Science Framework: https://osf.io/ngsp6/. This study was not preregistered.

Experiments 1a & 1b

Experiments 1a and 1b aimed to identify the verbal and nonverbal cues that are uniquely produced with indirect replies. In both experiments, participants first heard either a face-threatening scenario or a non-face-threatening scenario. The questioner (confederate) then asked a question related to the scenario, and the responder (real participant) needed to answer the question in one of four reply conditions; direct, indirect, lie or neutral. In the

direct and indirect reply conditions, the responder needed to break the bad news either directly (e.g., “No, I don’t think they found your joke very funny.”) or indirectly without hurting the questioner’s feeling (e.g., “Well, I think they probably have a different sense of humour to you.”). Another way to save the questioner’s face is to tell a white lie (e.g., “Yes, I think they found your joke really funny”). However, the lie and indirect replies differ in their sincerity. In an indirect reply, the responder tries to tell the bad news without offending the person asking the question, whereas in a lie reply the responder tries to avoid telling the truth. Another difference between the indirect and lie reply is that a lie reply is simpler to produce than an indirect reply. In a lie, the responder only needs to provide a confirmative response to the question, whereas in an indirect reply the responder needs to find an appropriate excuse to save the questioner’s face. Finally, we also included a neutral reply condition as a control condition for non-face-saving fact-based replies (e.g., “People from different cultures often have different senses of humour.”). In Experiment 1a, participants replied to the question with their own spontaneous answers, whereas participants in Experiment 1b replied to the question with scripted answers provided to them. Using scripted answers would allow us to control for the language production complexity between different types of replies. Furthermore, due to the exploratory nature of the current research, Experiment 1b would also serve as a reliability measure for any findings observed in Experiment 1a.

Method

Participants

Thirty-four native English speakers (23 females) aged 17-29 years (M = 20.59, SD = 2.79) were recruited in Experiment 1a and thirty-two native English speakers (24 females) aged 17-30 years (M = 19.88, SD = 2.32) were recruited in Experiment 1b. All participants were recruited from the University of Aberdeen. They were compensated with course credits or £7 monetary reward. The experiments have received ethical approval from the Psychology

Ethics Committee at the University of Aberdeen (PEC/3540/2016/10). To our knowledge, no experimental studies have examined various verbal and nonverbal behaviours simultaneously in the production and comprehension of indirect replies, we cannot estimate our sample size based on previous research. Therefore, the sample sizes of all experiments reported in this paper were determined by the number of participants who volunteered to participate in each academic term.

Materials

In Experiment 1a, we created 40 context scenarios and 40 related questions. Some scenarios were adapted from Bašnáková et al. (2014). Each scenario and question had a face-threatening and a neutral version. In the face-threatening version, the scenario involved moderately bad news, whereas in the neutral version, the scenario described a similar situation without any bad news (see Table 1 for an example). All scenarios were presented auditorily and narrated by native English speakers.

In Experiment 1b, the same context scenarios and questions were used. In addition, four different types of scripted replies (i.e., direct, indirect, lie, natural) were provided to participants (see Table 1 for an example). These different types of scripted replies matched in sentence length (number of words), lexical frequencies of the content words based on the CELEX lexical database (Baayen et al., 1995), and the amount repeating content words from the context story and the lead-in question. In this way, the linguistic complexity and cognitive load were controlled across the four reply conditions.

Table 1

Example of the stimulus materials used in Experiments 1a and 1b

Face-threatening context scenario:
Your friend Simon is going to try out for the cross-country running team this year, however he is somewhat out of shape and you know he is unlikely to make the team. He asks for your thoughts on his chances of qualifying.
Question:
Are they likely to accept me onto the team?
Scripted replies (only in Experiment 1b):
Direct: No, I think your chances of making the team are very slim.
Indirect: They are a very fit and competitive team of runners.
Lie: In my opinion, you have every chance of getting on the team.
Neutral context scenario:
You and your roommate Simon are discussing the running team at your university. You are talking about how they seem to have improved dramatically from the previous year. Simon asks you why you think this is.
Question:
Why do you think the running team is successful this year?
Scripted reply (only in Experiment 1b):
Neutral: They are a very fit and competitive team of runners.

#

Procedure and Design

Upon arrival at the laboratory, the experimenter led the participant and a confederate to the testing room. After the participant and confederate were seated, they were told that the purpose of the study was to investigate how people talk to each other in sensitive social

situations, but there was no mention of verbal and nonverbal cues. They first completed a general consent form and then a separate consent form on how their video-audio recordings would be used. The study started with a warm-up phase, followed by the main task. During the warm-up phase, the experimenter asked the dyads to recall a situation in their own life in which they had to deliver moderately bad news, and then the experimenter would ask a series of short follow-up questions on how they dealt with the situation. The warm-up phase was designed to get the participant to talk about sensitive social situations and allow the participant and confederate to be more comfortable with each other in a video recording situation. All the warm-up questions were first asked to the real participant to remove any potential bias that might be imparted from the confederate's answers.

The main task involved the dyads first listening to context scenarios and then undertaking question-answer exchanges. The dyads were seated face-to-face at approximately 1.5 metres distance from one another. Two presentation screens were placed back to back between the dyads on a side table so they could not see each other's screen. Two video cameras (Canon Legria HF R606; frame rate of 25 frames per second) were used to record the participant and the confederate's verbal and nonverbal behaviours. Two single plug-in earphones were used to play context scenarios to each participant. The experimenter was seated approximately 2.5 metres behind the real participant. Figure 1 illustrates the setup of Experiments 1a and 1b.

Before the main task started, a fake group assignment procedure was used so that the confederate was always assigned the role of the questioner and the real participant was always assigned the role of the responder. A confederate was used because participants presumably responded more naturally when they believed that they were conversing with a fellow naive participant as opposed to the experimenter. In addition, the confederate remained neutral to the participant's reply to avoid biasing their verbal and nonverbal

behaviours. When asked at the end of the experiment, no participant reported that they were aware that their fellow participant was a confederate.

During the instruction phase, participants were told that they would hear some short scenarios describing a situation between two people (e.g., the cross-country running team scenarios displayed in Table 1). The questioner was required to ask the question presented on the questioner's screen. The responder was required to respond to the question according to the condition shown on the responder's screen. There were four reply conditions: direct, indirect, lie and neutral. In the direct condition, the responder needed to break the bad news directly. They were given an example answer (e.g., "No, I think your chances of making the team are very slim.") and an explanation of the condition (e.g., "In this way, you simply break the bad news directly."). In the indirect reply condition, the responder needed to break the bad news indirectly without hurting the questioner's feeling. They were given an example answer (e.g., "They are a very fit and competitive team of runners.") and an explanation of the condition (e.g., "In this way, you express the bad news indirectly without hurting Maria's feelings much."). In the lie condition, the responder needed to lie and avoid breaking any bad news. They were given an example answer (e.g., "You have every chance of getting on the team.") and an explanation of the condition (e.g., "In this way, you lied so as NOT to break the bad news and avoid hurting Maria's feelings."). In the neutral condition, the responder needed to provide a fact-based answer. They were given an example answer (e.g., "There's been a real focus on affective training sessions.") and an explanation of the condition (e.g., "Here, you do not evaluate or criticize anybody, just provide a factual answer to the question."). The questioner needed to guess the type of reply given by the responder and rate the naturalness of the reply. However, in fact the questioner's task was a 'fake' task with no data being collected. The purpose of this task was to increase the credibility of the confederate as a participant and encourage natural replies from the real participant.

Further instructions were given to the responder separately while the questioner was asked to wait outside the testing room. The responder was told that only they would hear a face-threatening scenario (e.g., “Your friend is unlikely to make the cross-country running team and asks your opinion”), whereas the questioner would hear a non-face-threatening version of the scenario (e.g., “You try to get into the cross-country running team and you ask your friends opinion on your chance to make it”). Since the scenarios were played to the questioner (confederate) and the responder (real participant) via separate earphones, all responders believed that only they heard a face-threatening version of the scenario, when in fact, the same scenario was played to the dyads. This extra instruction was used to elicit a more naturalistic indirect reply based on the logic that in everyday conversation people only break bad news indirectly if they know that their conversation partner is unaware of the bad news.

In each trial of the main experiment, the dyads were first presented with the context scenarios auditorily through the plug-in earphones, followed by the presentation of the question and the reply condition written on the questioner’s and the responder’s screens. The dyads were asked to signal to the experimenter once the questioner memorized the question and the responder was ready to answer the question according to the given reply condition. The responder was told that they should give a short reply with one or two sentences, although there was no time limit. Once the dyads signalled that they were ready, the experimenter turned both screens blank via a remote-controlled mouse and the participants could then start the question-answer exchange. This was done to ensure a natural conversation without distractions from the information on the screens. Once the responder had given the reply, the questioner subsequently guessed and wrote down the type of reply and rated the naturalness of the reply in a booklet.

Before the trials of the main experiment, the dyads practiced the four types of replies with one face-threatening and one neutral context scenario. In the main task, the four reply conditions were evenly distributed across 40 context scenarios presented in a random order. The main experiment took 45 to 60 minutes. After completing the main experiment, the participants were then debriefed and informed about the purpose of the study and the use of the confederate.

The same procedure and design were used in Experiments 1a and 1b. The only difference was that the responder in Experiment 1b was also provided with a scripted answer in the reply condition. The responder was instructed to follow the scripted answer as accurately as possible. However, they were also told that the purpose of the experiment was not a memory test and they did not have to follow every single word. Instead, it was important to deliver the scripted answer in a natural way as they would do in an everyday conversation.

Nonverbal Behaviour Coding

All nonverbal behaviour coding was done using the ELAN software (Wittenburg et al., 2006).

Palm-revealing Gesture. The categorization criteria for palm-revealing gesture coding was identical to those used in Chu et al. (2014). The hand orientation of palm-revealing gestures was typically palm up, or the hand turns to reveal more of the palm (see Figure 2a). They have very similar functions to the “open hand supine with lateral movement” described in Kendon (2004). These gestures are counted as a subtype of interactive gestures as defined in Bavelas et al. (1992). Palm-revealing gestures were sometimes produced with shoulder shrugs. We also counted a shoulder shrug alone as a palm-revealing gesture because it serves the same function of palm-revealing gestures (Ferré, 2012; Cooperrider et al., 2018).

Head Tilt. A head tilt consisted of a single lateral head movement to either the right or left (see Figure 2b).

Facial Shrug. Based on research, a facial shrug involved pulling back of the mouth corners and raising of the cheeks and eyebrow (Chovil, 1991; Takeuchi & Nagao, 1993; Stone & Oh, 2008; Debras, 2017; see Figure 2c).

Gaze Aversion. Gaze aversion was manually coded by segmenting the time when the participant did not make direct eye contact with the confederate.

Intercoder Reliability. To establish the intercoder reliability of the nonverbal behaviour coding, a second trained coder (who was blind to the reply conditions) independently identified all palm-revealing gestures, head tilts, facial shrugs and gaze aversion from eight randomly selected participants in Experiment 1a, and nine randomly selected participants in Experiment 1b. The proportions of the four types of nonverbal behaviours checked by the second coder out of all nonverbal behaviours identified by the first coder were 45%, 41%, 41% and 26% in Experiment 1a and were 51%, 55%, 49% and 25% in Experiment 1b. Agreement on the coding of palm-revealing gestures, head tilts, and facial shrugs was 94%, 96%, and 89% in Experiment 1a and was 93%, 91% and 89% in Experiment 1b, respectively. For gaze aversion, agreement was measured by calculating the correlations of the proportion of gaze aversion in each trial between the primary and secondary coders. The correlations were extremely high in both Experiments (Experiment 1a: r (355) = .99, $p < .001$; Experiment 1b: r (302) = .99, $p < .001$)

Verbal Behaviour Coding

Uncertainty Terms. Uncertainty terms are also named as probability and hedging terms in previous research (e.g., Bonnefon et al., 2009; Holtgraves & Perdew, 2016). The terms examined in the present study included ‘possibly’, ‘probably’, ‘could’, ‘might’,

'maybe', 'perhaps', 'may', 'I guess, 'kind of', 'not sure', 'don't know', 'suppose' and 'fairly'.

Discourse Markers. Discourse markers are words or phrases acting as transitions between different sections of conversation (Clark, 1996). The ones coded in the present study included 'you know', 'well', 'like', 'I think', and 'I mean'.

Filled Pause. The filled pauses included those ending with either an 'm' sound (e.g., mm, hmm, um) or an 'er' sound (e.g., uh, er, ah). They have been identified as the most commonly used filled pauses within the English language (Tottie, 2014).

Silent Pause. In accordance with previous research from Heldner and Edlund (2010), any silent period over 180 ms in speech was coded as a silent pause.

Results

To identify the unique nonverbal and verbal behaviours associated with the production of indirect replies, generalized linear mixed models (GLMMs) analyses were used for both Experiments 1a and 1b. For the majority of the analyses, apart from gaze aversion and silent pause, the assumptions of linearity, homoscedasticity and residuals normality were violated because both Experiments contained a high number of zeros values on the three nonverbal (palm-revealing gestures, head tilt, facial shrug) and the three verbal (uncertainty terms, discourse markers, filled pause) behaviours. Same problem persisted after data transformation (i.e., logarithmic). Therefore, we decided to convert the rates of these nonverbal and verbal behaviours into dichotomous variables, with zero indicating that a behaviour did not occur and one indicating that a behaviour occurred, and then perform binomial GLMMs to analyse the effect of the four reply conditions on the likelihood of a behaviour occurred in a trial. For this analysis we used the glmmTMB package (Magnusson et al., 2017) in R (4.1.2). For gaze aversion and silent pauses, as the assumptions of linearity, homoscedasticity and residuals normality were not violated, they were analysed with

GLMMs using lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages in R. For all the analyses, the four reply conditions (direct, indirect, lie and neutral) were treated as the fixed factor, and subjects and stimuli were treated as crossed random factors. The two random factors express the variability in the data set across the different participants and stimuli. The maximal fixed effect structure was determined by the likelihood structure where the null model, that contained only the random effects, was compared with the full model containing the fixed factor. Finally, for all the models, Cook's distance and Durbin-Watson test of independence were calculated with the use of the function that was included in car R-package (Fox & Weisberg, 2019). Both assumptions were met. For data wrangling and visualization packages tidyverse and ggplot2 (Wickham et al., 2019) were used.

Trials in which the participants did not give a reply or did not reply in the correct condition were categorized as error trials and were removed from the analysis. For Experiment 1a, 247 (18.16%) of all 1360 trials were error trials, of which 43 were in the direct reply condition, 144 were in the indirect reply condition, 24 were in the lie reply condition and 36 were in the neutral reply condition. For Experiment 1b, 15 (1.17%) of all 1280 trials were error trials, of which 2 were in the direct reply condition, 3 were in the indirect reply condition, 2 were in the lie reply condition and 8 were in the neutral reply condition. Table 2 presents the average word count and the duration for each reply condition in Experiments 1a and 1b.

Table 2

Average Word Count and Reply duration for Each Condition

	Experiment 1a		Experiment 1b	
Condition	Word count (*SD*)	Reply duration (*SD*)	Word count (*SD*)	Reply duration (*SD*)
Direct	11.48 (11.69)	4.03 (3.60)	11.53 (5.95)	3.39 (1.42)

Indirect	17.30 (10.43)	6.37 (3.38)	13.26 (7.36)	4.00 (1.81)
Lie	9.35 (5.87)	3.01 (1.76)	10.82 (4.65)	2.85 (1.05)
Neutral	14.96 (12.24)	5.98 (3.87)	11.50 (4.50)	3.56 (1.36)

Note. The reply duration was presented in seconds.

Nonverbal Behaviours Analysis

Palm-revealing Gesture. Table 3 presents the total number of palm-revealing gestures, the percentage of trials that contained at least one palm-revealing gesture and the mean rate of palm-revealing gestures (number of palm-revealing gestures per minute) in Experiments 1a and 1b.

Table 3

Descriptive Statistics for Palm-revealing Gestures

Condition	Experiment 1a			Experiment 1b		
	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	61	13.47%	2.85 (3.97)	40	9.12%	2.32 (6.51)
Indirect	72	22.96%	3.90 (4.88)	83	18.61%	4.14 (7.92)
Lie	27	7.28%	1.82 (2.74)	52	12.89%	3.34 (8.31)
Neutral	80	20.07%	2.96 (3.39)	72	16.99%	3.79 (7.25)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of palm revealing gestures occurrence in both Experiment 1a (*AIC* = 848.90, *BIC* = 879.00, *LogLik* = -418.50, x^2 (3) = 38.48, $p < .001$) and Experiment 1b (*AIC* = 680.50, *BIC* = 711.40, *LogLik* = -334.30, x^2 (3) = 24.57, $p < .001$). For both experiments, the likelihood of palm revealing gestures occurrence in the indirect reply condition was significantly higher than that in the direct and the lie reply conditions but was not different

from that in the neutral reply condition (see Table 4). Thus, there was no conclusive evidence that palm revealing gesture production was uniquely associated with the production of indirect replies.

Table 4

Results of Binomial GLMMs Analyses of Palm Revealing gesture

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 1a	(Intercept)	-1.56	0.30	-2.14	-0.99	-5.29	<.001
	Indirect vs Direct	-0.80	0.27	-1.33	-0.28	-2.98	0.003
	Indirect vs Lie	-1.56	0.30	-2.15	-0.97	-5.15	<.001
	Indirect vs Neutral	-0.18	0.25	-0.67	0.31	-0.72	0.473
Exp. 1b	(Intercept)	-2.66	0.55	-3.72	-1.59	-4.87	<.001
	Indirect vs Direct	-1.44	0.33	-2.09	-0.79	-4.37	<.001
	Indirect vs Lie	-0.78	0.30	-1.36	-0.20	-2.62	0.009
	Indirect vs Neutral	-0.24	0.28	-0.78	0.31	-0.85	0.398

Note. In all GLMMs analyses reported in this paper, Std. Error indicates the standard error of the estimate for each predictor. CIlow and CIhigh indicate the lower and upper limits of the 95% confidence interval for the estimate of each predictor.

Head Tilt. Table 5 presents the total number of head tilts, the percentage of trials that contained at least one head tilt and the mean rate of head tilts (number of head tilts per minute) in Experiments 1a and 1b.

Table 5

Descriptive Statistics for Head Tilts

Condition	Experiment 1a			Experiment 1b		
	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	53	12.46%	2.11 (2.79)	33	9.12%	1.79 (3.37)
Indirect	91	30.61%	3.97 (4.91)	82	22.08%	4.35 (5.57)
Lie	32	7.91%	1.71 (2.55)	33	9.43%	2.34 (4.36)
Neutral	60	13.16%	1.65 (1.94)	50	13.46%	2.65 (4.46)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of head tilt occurrence in both Experiment 1a (*AIC* = 773.50, *BIC* = 803.60, *LogLik* = -380.80, x^2 (3) = 57.93, $p < .001$) and Experiment 1b (*AIC* = 275.90, *BIC* = 300.00, *LogLik* = -131.90, x^2 (3) = 11.51, $p = 0.009$). For both experiments, the likelihood of head tilt occurrence in the indirect reply condition was significantly higher than that in the other three reply conditions (see Table 6). These results suggest that head tilt is a unique nonverbal behaviour that people produce during indirect communication.

Table 6

Results of Binomial GLMMs Analyses of Head Tilt

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)

Exp. 1a	(Intercept)	-1.11	0.31	-1.73	-0.50	-3.55	<.001
	Indirect vs Direct	-1.48	0.28	-2.02	-0.94	-5.36	<.001
	Indirect vs Lie	-2.07	0.30	-2.66	-1.48	-6.90	<.001
	Indirect vs Neutral	-1.37	0.27	-1.89	-0.85	-5.12	<.001
Exp. 1b	(Intercept)	-1.74	0.57	-2.85	-0.62	-3.06	0.002
	Indirect vs Direct	-1.04	0.45	-1.92	-0.17	-2.33	0.020
	Indirect vs Lie	-1.45	0.49	-2.41	-0.50	-2.98	0.003
	Indirect vs Neutral	-0.93	0.44	-1.79	-0.01	-2.1	0.035

Facial Shrug. Table 7 presents the total number of facial shrugs, the percentage of trials that contained at least one facial shrug and the mean rate of facial shrugs (number of facial shrugs per minute) in Experiments 1a and 1b.

Table 7

Descriptive Statistics for Facial Shrug

	Experiment 1a			Experiment 1b		
Condition	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	39	12.46%	2.14 (4.34)	20	5.97%	1.25 (3.01)

Indirect	32	14.80%	1.61 (2.57)	36	10.41%	1.73 (2.97)
Lie	9	2.53%	0.66 (1.81)	4	1.26%	0.36 (1.30)
Neutral	24	7.24%	0.78 (1.24)	20	6.09%	1.08 (2.10)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of facial shrug occurrence in both Experiment 1a (*AIC* = 547.80, *BIC* = 577.90, *LogLik* = -267.90, x^2 (3) = 37.32, p <.001) and Experiment 1b (*AIC* = 474.10, *BIC* = 504.90, *LogLik* = -231.00, x^2(3) = 32.29, p < .001). In Experiment 1a the likelihood of facial shrug occurrence in the indirect reply condition was significantly higher than that in the lie and the neutral reply conditions, whereas in Experiment 1b the likelihood of facial shrug occurrence in the indirect reply condition was significantly higher than that in the other three reply conditions (see Table 8). Thus, there was no conclusive evidence that facial shrug production was uniquely associated with indirect replies.

Table 8

Results of Binomial GLMMs Analyses of Facial Shrug

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 1a	(Intercept)	-2.46	0.41	-3.26	-1.67	-6.06	<.001
	Indirect vs Direct	-0.32	0.30	-0.92	0.28	-1.05	0.296
	Indirect vs Lie	-2.19	0.45	-3.07	-1.32	-4.92	<.001
	Indirect vs Neutral	-1.01	0.34	-1.67	-0.35	-3.00	0.003
Exp. 1b	(Intercept)	-3.15	0.46	-4.05	-2.25	-6.86	<.001

	Indirect vs Direct	-0.80	0.34	-1.46	-0.13	-2.35	0.019
	Indirect vs Lie	-2.58	0.57	-3.70	-1.46	-4.51	<.001
	Indirect vs Neutral	-0.79	0.34	-1.45	-0.12	-2.33	0.020

Gaze Aversion. Table 9 presents the total duration of gaze aversion (in minutes), the percentage of trials that contained at least one gaze aversion and the mean proportion of gaze aversion (total gaze aversion time divided by total reply duration) in Experiments 1a and 1b.

Table 9

Descriptive Statistics for Gaze Aversion

	Experiment 1a			Experiment 1b		
Condition	Total duration	Percentage of trials	Proportion (*SD*)	Total duration	Percentage of trials	Proportion (*SD*)
Direct	7.09	83.16 %	30.89% (16.51%)	5.45	78.62%	29.68% (18.04%)
Indirect	10.05	93.37 %	44.18% (22.65%)	6.17	83.91%	29.74% (19.06%)
Lie	5.40	83.86 %	31.57% (15.74%)	4.08	72.64%	26.57% (19.70%)

Neutral	14.64	95.39 %	47.19% (18.58%)	5.77	82.3 7%	31.87% (18.88%)

GLMMs analyses showed that the proportion of gaze aversion significantly differ between the four reply conditions in both Experiment 1a (*AIC* = -239.40, *BIC* = -204.30, *LogLik* = 126.70, x^2 (3) = 99.33, *p* <.001) and Experiment 1b (*AIC* = -966.50, *BIC* = -930.50, *LogLik* = 490.30, x^2 (3) = 14.38, *p* = 0.002). In Experiment 1a participants produced significantly more gaze aversion in the indirect reply condition than in the direct and lie reply conditions, whereas in Experiment 1b participants produced significantly more gaze aversion in the indirect reply condition than in the lie condition (see Table 10). Thus, there was no evidence that gaze aversion was uniquely associated with the production of indirect replies.

Table 10

Results of GLMMs Analyses of Gaze Aversion

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	t-value	Pr (>|t|)
Exp. 1a	(Intercept)	0.42	0.03	0.36	0.47	14.73	<.001
	Indirect vs Direct	-0.11	0.02	-0.15	-0.07	-5.91	<.001
	Indirect vs Lie	-0.10	0.02	-0.14	-0.06	-5.37	<.001
	Indirect vs Neutral	0.03	0.02	-0.01	0.07	1.58	0.116
Exp. 1b	(Intercept)	0.30	0.03	0.23	0.36	9.31	<.001
	Indirect vs Direct	-0.01	0.01	-0.03	0.02	-0.65	0.511

	Indirect vs Lie	-0.03	0.01	-0.06	-0.01	-2.55	0.011
	Indirect vs Neutral	0.02	0.01	-0.01	0.04	1.20	0.239

Verbal Behaviour Analysis

Uncertainty Terms. Table 11 presents the total number of uncertainty terms, the total number of trials that contained at least one uncertainty term and the mean rate of uncertainty terms (number of uncertainty terms per minute) in Experiments 1a and 1b. In Experiment 1b, there were three items that contained uncertainty terms within the scripted answers, and they were removed from the analysis.

Table 11

Descriptive Statistics for Uncertainty Terms

Condition	Experiment 1a			Experiment 1b		
	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	33	8.42%	1.45 (1.72)	24	6.29%	1.07 (1.98)
Indirect	126	45.92%	6.78 (6.06)	23	7.26%	0.88 (1.60)
Lie	18	5.06%	1.24 (2.30)	6	1.89%	0.39 (1.06)
Neutral	87	22.70%	3.33 (3.58)	8	2.56%	0.38 (0.87)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of uncertainty terms occurrence in both Experiment 1a (*AIC* = 851.00, *BIC* = 881.10, *LogLik* = -419.50, x^2 (3) = 162.4, p <.001) and Experiment 1b (*AIC* = 534.60, *BIC* = 565.40, *LogLik* = -261.30, x^2 (3) = 25.37, p < .001). For Experiment 1a, the likelihood of

uncertainty terms occurrence in the indirect reply condition was significantly higher than that in all the other three reply conditions. For Experiment 1b, the likelihood of uncertainty terms occurrence in the indirect reply condition was significantly higher than that in the direct and lie reply conditions, but was only marginally significantly higher than that in the neutral reply condition (p = .051) (see Table 12). It is worth noting that in Experiment 1b participants were asked to follow the scripted answers as accurately as possible, however, uncertainty terms were still used more frequently in indirect replies than other types of replies (except that the difference between the indirect and the neutral reply conditions was only marginal). Taken together, the results strongly support the idea that uncertainty terms (e.g., maybe, perhaps, not sure) are a crucial cue in expressing politeness in a face-threatening situation during indirect communication (e.g., Bonnefon et al., 2009; Brown & Levinson, 1987; Holtgraves, 2016; Youmans, 2001).

Table 12

Results of Binomial GLMMs Analyses of Uncertainty Terms

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 1a	(Intercept)	-0.26	0.26	-0.76	0.25	-1.01	0.315
	Indirect vs Direct	-2.61	0.29	-3.18	-2.04	-8.98	<.001
	Indirect vs Lie	-3.12	0.32	-3.76	-2.49	-9.61	<.001
	Indirect vs Neutral	-1.27	0.23	-1.72	-0.81	-5.46	<.001
Exp. 1b	(Intercept)	-2.76	0.34	-3.43	-2.10	-8.18	<.001

Indirect vs Direct	-0.69	0.33	-1.33	-0.05	-2.11	0.035
Indirect vs Lie	-2.06	0.78	-3.00	-1.12	-4.31	<.001
Indirect vs Neutral	-0.63	0.32	-1.27	0.003	-1.95	0.051

Discourse Markers. Table 13 presents the total number of discourse markers, total number of trials that contained at least one discourse marker and the mean rate of discourse markers (number of discourse markers per minute) for all participants in Experiments 1a and 1b.

Table 13

Descriptive Statistics for Discourse Markers

Condition	Experiment 1a			Experiment 1b		
	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	46	8.42%	1.18 (2.45)	19	5.97%	1.04 (1.55)
Indirect	111	36.74%	4.11 (4.30)	85	26.81%	4.74 (4.30)
Lie	33	7.59%	1.32 (2.31)	13	4.09%	0.72 (1.71)
Neutral	72	16.78%	2.08 (2.92)	40	12.82%	2.43 (3.32)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of discourse markers occurrence in both Experiment 1a (*AIC* = 733.90, *BIC* = 764.00, *LogLik* = -360.90, x^2 *(3)* = 96.44, $p < .001$) and Experiment 1b (*AIC:* 792.30, *BIC* = 823.10, *LogLik* = -390.10, x^2 *(3)* = 99.98, $p < .001$). For both experiments the likelihood of

discourse markers occurrence in the indirect reply condition was significantly higher than that in the other three reply conditions (see Table 14). Thus, results from both experiments clearly support the idea that discourse markers are used for a politeness function in face-threatening conversations (Tree &Schrock, 2002).

Table 14

Results of Binomial GLMMs Analyses of Discourse Markers

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 1a	(Intercept)	-0.88	0.32	-1.50	-0.25	-2.76	0.006
	Indirect vs Direct	-2.37	0.31	-2.98	-1.75	-7.59	<.001
	Indirect vs Lie	-2.45	0.31	-3.06	-1.84	-7.86	<.001
	Indirect vs Neutral	-1.31	0.26	-1.81	-0.81	-5.12	<.001
Exp. 1b	(Intercept)	-1.40	0.30	-1.99	-0.81	-4.62	<.001
	Indirect vs Direct	-2.00	0.29	-2.57	-1.44	-6.98	<.001
	Indirect vs Lie	-2.44	0.33	-3.08	-1.80	-7.45	<.001
	Indirect vs Neutral	-1.10	0.23	-0.65	0.003	-4.76	<.001

Filled Pause. Table 15 presents the total number of filled pauses, the total number of trials that contained at least one filled pause and the mean rate of filled pauses (number of filled pauses per minute) in Experiments 1a and 1b.

Table 15

Descriptive Statistics for Filled Pauses

Condition	Experiment 1a			Experiment 1b		
	Total number	Percentage of trials	Rate (*SD*)	Total number	Percentage of trials	Rate (*SD*)
Direct	82	18.52%	2.80 (3.46)	74	15.41%	2.94 (4.76)
Indirect	111	35.71%	3.68 (3.72)	79	19.24%	2.97 (4.28)
Lie	39	9.81%	1.83 (2.68)	40	7.55%	1.51 (4.04)
Neutral	136	32.24%	3.96 (3.42)	101	22.76%	4.35 (6.52)

Binomial analyses showed that the four reply conditions had a significant effect on the likelihood of filled pauses occurrence in both Experiment 1a (*AIC* = 930.60, *BIC* = 960.70, *LogLik* = -459.30, x^2 *(3)* = 86.17, $p < .001$) and Experiment 1b (*AIC* = 702.20, *BIC* = 733.10, *LogLik* = -345.10, x^2 *(3)* = 58.28, $p < .001$). In Experiment 1a participants produced significantly more filled pauses in the indirect reply condition than in the direct and lie reply conditions, whereas in Experiment 1b participants produced significantly more filled pauses in the indirect reply condition than in the lie reply condition (see Table 16). Thus, there was no evidence that filled pauses were uniquely associated with the production of indirect replies.

Table 16

Results of Binomial GLMMs Analyses of Filled Pauses

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 1a	(Intercept)	-0.96	0.36	-1.66	-0.26	-2.68	0.007
	Indirect vs Direct	-1.14	0.26	-1.65	-0.63	-4.41	<.001
	Indirect vs Lie	-2.16	0.30	-2.74	-1.57	-7.20	<.001
	Indirect vs Neutral	-0.17	0.24	-0.64	0.30	-0.71	0.480
Exp. 1b	(Intercept)	-2.31	0.49	-3.27	-1.36	-4.75	<.001
	Indirect vs Direct	-0.49	0.29	-1.05	0.08	-1.68	0.093
	Indirect vs Lie	-2.11	0.39	-2.88	-1.34	-5.38	<.001
	Indirect vs Neutral	0.38	0.27	-0.15	0.90	1.42	0.156

Silent Pause. Table 17 presents the total duration of silent pause (in minutes), the total number of trials that contained at least one silent pause and the mean proportion of silent pause (total silent pause duration divided by total reply duration) in Experiments 1a and 1b.

Table 17

Descriptive Statistics for Silent Pause

Condition	Experiment 1a			Experiment 1b		
	Total duration	Percentage of trials	Proportion (SD)	Total duration	Percentage of trials	Proportion (SD)
Direct	5.22	95.96%	28.74% (8.51%)	3.90	96.23%	21.73% (5.71%)
Indirect	6.36	100%	29.02% (10.44%)	4.45	98.42%	21.13% (5.99%)
Lie	3.90	84.18%	25.46% (11.35%)	2.75	93.71%	18.21% (5.70%)
Neutral	9.52	99.01%	32.40% (12.90%)	4.00	98.40%	20.96% (5.02%)

GLMMs analyses showed that the four reply conditions had a significant effect on the proportion of silent pauses in both Experiment 1a (*AIC* = -1262.40, *BIC* =-1227.30, *LogLik* = 638.20, x^2 (3) = 53.64, p <.001) and in Experiment 1b (*AIC* = -2751.80, *BIC* = -2715.80, *LogLik* = 1382.90, x^2 (3) = 37.18, p <.001). In Experiment 1a participants produced significantly more silent pause in the indirect reply condition than in the lie and neutral reply conditions, whereas in Experiment 1b participants produced significantly more silent pause in the indirect reply condition than in the lie reply condition (Table 18). Thus, there was no evidence that silent pause was uniquely associated with the production of indirect replies.

Table 18

Results of GLMMs Analyses of Silent Pause

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	t-value	Pr (>|t|)
Exp. 1a	(Intercept)	0.28	0.02	0.25	0.32	16.27	<.001
	Indirect vs Direct	0.01	0.01	-0.01	0.04	1.08	0.279
	Indirect vs Lie	-0.04	0.01	-0.06	-0.01	-3.16	0.002
	Indirect vs Neutral	0.04	0.01	0.01	0.06	3.20	0.001
Exp. 1b	(Intercept)	0.21	0.01	0.19	0.23	22.26	<.001
	Indirect vs Direct	0.01	0.01	-0.01	0.02	0.98	0.328
	Indirect vs Lie	-0.03	0.01	-0.04	-0.02	-4.66	<.001
	Indirect vs Neutral	-0.001	0.01	-0.01	0.01	-0.23	0.818

Discussion

Experiments 1a and 1b aimed to identify the verbal and nonverbal behaviours that are uniquely associated with the production of indirect replies. Previous research has shown that speakers often use uncertainty expressions when delivering negative messages indirectly to preserve ‘face’ of those involved in the conversational exchanges (Bonnefon & Villejoubert, 2006; Holtgraves & Perdew, 2016). Therefore, we focused on the verbal and nonverbal

behaviours that have been shown to express uncertainty in conversations. For the nonverbal behaviours, only the head tilt was uniquely produced in indirect replies. With regards to the verbal behaviours, we found that both uncertainty terms and discourse markers were uniquely associated with the production of indirect replies.

Experiments 2a & 2b

The results of Experiments 1a and 1b showed that the head tilt was a unique nonverbal behaviour associated with the production of indirect replies. However, it is unclear which nonverbal cues people used in the comprehension of indirect replies. Experiments 2a and 2b aimed to address this issue by examining the independent contribution of different nonverbal behaviours in identifying indirect replies. Participants viewed *silent* video clips of people giving four types of replies (direct, indirect, lie or neutral) and then categorized the type of reply presented in each clip. This would allow us to determine the pure contribution of nonverbal cues in identifying indirect replies without any impact from the verbal information. Correlational and multiple regression analyses were conducted with the proportion of indirect categorization as the dependent variable, and the frequencies of palm-revealing gesture, head tilt, facial shrug, the proportion of gaze aversion, and the reply duration as the predictor variables.

Method

Participants

Forty-one participants (36 females) aged 18-24 years (M = 19.51, SD = 1.33) were recruited in Experiment 2a. Fifty-eight participants (49 females) aged 18-25 years (M = 19.76, SD = 1.51) were recruited in Experiment 2b. All participants were recruited from the University of Aberdeen and were compensated with course credits or £7 monetary reward. The experiments have received ethical approval from the Psychology Ethics Committee at the University of Aberdeen (PEC/3837/2018/1).

Materials

Video clips of the spontaneous replies recorded in Experiment 1a (a total of 1360 video clips with 340 clips in each reply condition) and the scripted replies recorded in Experiment 1b (a total of 1280 video clips with 320 clips in each reply condition) were used for material selection in Experiments 2a and 2b, respectively. The video clips were edited using Adobe Premiere Pro 5. Audio was removed from all clips. The face of the experimenter in the background was blocked out (see illustration from phase 1 in Figure 3) so as not to distract the participant from focusing on the responder's nonverbal behaviours in the video clip.

For Experiment 2a, we first removed the video clips with error responses from Experiment 1a (43, 144, 24, 36 video clips from the direct, indirect, lie, and neutral reply condition, respectively). We then randomly removed some clips so that there were equal number of video clips in each reply condition from each speaker. Thus, the final stimulus set consisted of 716 video clips with 179 in each reply condition. All clips were distributed into 4 testing lists of equal running time to keep the length of the experiment reasonable for participants. Each list contained an equal number of clips in each reply condition.

For Experiment 2b, we used video clips of all replies (320 clips in each reply condition) from Experiment 1b. These clips were distributed into 8 lists, with each list containing an equal number of clips in each reply condition. Among the 1280 trials used in Experiment 2b, 15 of them were error trials. These error trials were excluded from the final analyses.

Procedure

For both Experiments 2a and 2b, the participants were told that they would view silent video clips containing replies collected from real participants in a prior study, and their task was to categorize the type of reply (i.e., direct, indirect, lie and neutral) and indicate their

confidence level in their categorization. The instructions contained the same example scenarios and replies used in Experiments 1a and 1b to ensure that participants fully understood the four types of replies.

For both Experiments, the trials consisted of three phases (see Figure 3). In the first phase, the participant viewed the clips in silence. In the second phase, the participant categorized the type of reply in a clip by pressing the corresponding number key (1=direct, 2=indirect, 3=lie, 4=neutral). Participants could re-watch the clip as many times as they wish by pressing the Space key. In the third phase, participants rated how confident they were about their categorization on a 7-point scale (1= not confident at all; 7 = extremely confident). The inter-trial interval was 500 ms. The video clips of each reply condition were randomly presented. Prior to the main experiment, the participant undertook 3 practice trials to familiarize themselves with the task.

Results

Table 19 presents the descriptive statistics of the mean categorization accuracy and total number of categorizations for all clips in Experiments 2a and 2b.

Table 19

Descriptive Statistics of the Mean Categorization Accuracy and the Proportion of Each Type of Categorizations

Condition	Experiment 2a		Experiment 2b	
	Categorization accuracy (*SD*)	Proportion of categorizations	Categorization accuracy (*SD*)	Proportion of categorizations
Direct	.41 (.10)	26.43%	.38 (.10)	27.66%
Indirect	.37 (.10)	26.39%	.36 (.10)	28.03%
Lie	.49 (.16)	28.68%	.39 (.17)	24.64%
Neutral	.23 (.09)	18.50%	.25 (.09)	19.67%

Predicting Indirect Categorization from Nonverbal Cues

To determine the contributions of each type of nonverbal behaviours in identifying indirect replies, binomial GLMMs were used for both Experiments. Participants' response was coded either as 1 (indirect categorization) or 0 (direct, lie, or neutral categorizations). For all the models, the rate of the four nonverbal cues (i.e., palm-revealing gesture, head tilt, facial shrug, gaze aversion) as well as the reply duration were used as fixed factors, and participants and stimuli were treated as random factors. We additionally included reply duration as a fixed factor because the average reply durations of the indirect replies were descriptively the longest in among all four types of replies (Experiment 2a: 4.85, 6.43, 3.62, 6.01 seconds in the direct, indirect, lie and neutral reply condition, respectively; Experiment 2b: 3.38, 4.01, 2.85, 3.57 seconds in the direct, indirect, lie and neutral reply condition, respectively). All assumptions, including VIF test for multicollinearity were met.

Tables 20 presents the descriptive statistics of the fixed factor variables in Experiments 2a and 2b.

Table 20

Descriptive statistics for the Predictor Variables in Experiments 2a and 2b

	Experiment 2a				Experiment 2b			
Predictor variable	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max
Palm-revealing gesture	2.95	8.13	0	83.16	3.33	9.91	0	100
Head tilt	2.4	6.38	0	46.69	2.68	7.55	0	55.4
Facial shrug	1.5	5.16	0	35.78	1.12	4.85	0	41.52

Gaze aversion	0.37	0.25	0	1	0.29	0.23	0	1
Reply duration (in seconds)	5.23	4.17	0.72	30.79	3.45	1.87	0.95	20.38

Tables 21 and 22 present the correlation coefficients among the fixed factors in Experiments 2a and 2b, respectively.

Table 21

Pearson Correlation Coefficients among the fixed factors in Experiment 2a

Variable	1	2	3	4
1. Palm-revealing gesture	-			
2. Head tilt	.12***	-		
3. Facial shrug	.09**	.09**	-	
4. Gaze aversion	-.06	.02	.01	-
5. Reply duration	.00	.10**	-.05	.24***

Note. $N = 716$; *$p < .05$, **$p < .01$, ***$p < .001$

Table 22

Pearson Correlation Coefficients among the fixed factors in Experiment 2b

Variable	1	2	3	4
1. Palm-revealing gesture	-			
2. Head tilt	.41***	-		
3. Facial shrug	.02	.11***	-	
4. Gaze aversion	-.03	-.02	.08**	-
5. Reply duration	.01	.01	-.01	.06*

Note. $N = 1265$; *$p < .05$, **$p < .01$, ***$p < .001$

Binomial analyses were significant in both Experiment 2a (*AIC* = 7624.90, *BIC* = 7679.80, *LogLik* = -3804.40, x^2 (5) = 237. 06, $p < .001$) and Experiment 2b (*AIC* = 10533.80, *BIC* = 10590.00, *LogLik* = -5258.90, $x^2(5) = 177.55$, $p < .001$). The models revealed that all four types of nonverbal behaviours (palm revealing, head tilt, facial shrug, and gaze aversion), as well as the reply duration, were significant predictors of the indirect categorization for both Experiment 2a and 2b (See Table 23).

Table 23

Summary of Binomial GLMMs Analyses in Experiments 2a and 2b

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 2a	(Intercept)	-2.22	0.11	-2.43	-2.01	-20.90	<.001
	Palm revealing	0.01	0.01	0.001	0.02	2.22	0.026
	Head Tilt	0.02	0.01	0.01	0.03	3.40	0.001
	Facial Shrug	0.02	0.01	0.01	0.04	2.91	0.004
	Gaze Aversion	0.49	0.17	0.16	0.82	2.91	0.004
	Reply Duration	0.13	0.01	0.11	0.15	13.49	<.001
Exp. 2b	(Intercept)	-1.79	0.08	-1.95	-1.63	-21.35	<.001
	Palm revealing	0.01	0.003	0.004	0.02	3.19	0.001
	Head Tilt	0.02	0.004	0.01	0.03	5.43	<.001

Facial Shrug	0.02	0.01	0.01	0.03	2.79	0.005
Gaze Aversion	0.26	0.13	0.01	0.50	2.05	0.041
Reply Duration	0.16	0.02	0.13	0.20	9.90	<.001

Discussion

Experiments 2a and 2b aimed to identify the nonverbal cues that participants used in categorizing indirect replies when no verbal information was available. Results showed that all selected nonverbal behaviours, as well as the reply duration, were significant predictors of the indirect categorization. Participants were more likely to categorize a reply as indirect if it contained more palm-revealing gesture, head tilt, facial shrug, gaze aversion, and longer reply duration.

Among all the predictor variables, reply duration was the most significant predictor in both experiments. This was presumably because not all replies contained all nonverbal cues, whereas the duration of most indirect replies was longer than other types of replies, as shown in Experiments 1a and 1b.

Among the four nonverbal behaviours, head tilt was the most salient cue that participants relied on when categorizing a reply as indirect in both experiments. This corresponded to our findings from Experiments 1a and 1b, which showed that head tilt was the only nonverbal behaviour that was produced more frequently in indirect replies compared to all other types of replies. Thus, head tilt was strongly associated with both the production and comprehension of indirect replies.

Results from Experiments 2a and 2b also revealed that people took palm-revealing gesture, facial shrug and gaze aversion as cues for indirect replies, which provides some supporting evidence for the idea that these nonverbal behaviours signify uncertainty and serve as cues for indirectness during conversations. However, unlike head tilt, these three types of nonverbal behaviours were not uniquely associated with the production of indirect replies, as shown in Experiments 1a and 1b.

Experiments 2a and 2b showed the importance of nonverbal behaviours for comprehending an indirect reply. However, in everyday communication people rarely communicate in silence. Thus, it is important to investigate the role of nonverbal behaviours in the comprehension of indirect replies when both verbal and nonverbal cues are available. Experiments 3a and 3b were designed to address this issue.

Experiments 3a & 3b

The first goal of Experiments 3a and 3b was to investigate whether nonverbal behaviours could facilitate detecting indirect replies when both verbal and nonverbal cues were available. The second goal was to examine the relative contributions of different verbal and nonverbal cues in identifying indirect replies. We presented video clips of indirect and neutral replies in one of the two modality conditions: a video-and-audio condition and an audio-only condition. Participants were asked to categorize the type of reply (indirect vs. neutral) and indicate the confidence in their categorization.

Unlike Experiments 2a and 2b, where participants were asked to categorize a clip from four types of replies (direct, indirect, lie and neutral), participants in Experiments 3a and 3b were only asked to categorize indirect and neutral replies. This was because performance on categorizing direct and lie replies was likely to be at ceiling when verbal information was available, as all direct replies were negative remarks and were likely to contain the word “no” and all lie replies were positive remarks and were likely to contain the word “yes”. In

contrast, the verbal contents of indirect and neutral replies were not tied to any specific words and could not be easily distinguished. If nonverbal cues could facilitate the detection of indirect replies, participants should be better able to differentiate between the indirect and neutral replies in the video-and-audio condition than in the audio-only condition. Furthermore, we performed binomial GLMMs analyses to explore the relative contribution of verbal and nonverbal cues in identifying indirect replies.

Method

Participants

Sixty participants (47 females) aged 18-57 years (M = 24.02, SD = 7.34) participated in Experiment 3a. Sixty-two participants (33 females) aged 18-58 years (M = 27.42, SD = 10.55) participated in Experiment 3b. All participants were recruited from the University of Aberdeen and were compensated with course credits or £7 monetary reward. The study received ethical approval from the Psychology Ethics Committee at the University of Aberdeen (PEC/3837/2018/1). In Experiment 3b, two participants were removed from the analyses. One failed to understand the instructions and always pressed the same response key for all trials in each modality condition. The other decided to withdraw from the experiment before it was finished.

Materials

Stimuli used in Experiments 3a and 3b were selected from Experiments 1a and 1b, respectively. The selection procedure in Experiments 3a and 3b was similar to that used in Experiments 2a and 2b. We first removed the video clips with incorrect replies from Experiments 1a and 1b. We then further removed video clips of one participant from Experiment 1a and one from Experiment 1b because their average reply durations were more than 3 standard deviations above the average reply duration across all participants. The final stimulus set consisted of 344 video clips in Experiment 3a and 608 video clips in Experiment

3b. Each set of video clips contained an equal number of indirect and neutral replies. For each selected video clip, we also created an audio-only version of the clip that only contained audio with a screen shot of the person sitting still in the clip without showing any nonverbal cue.

Selected video and audio clips were distributed into two lists in Experiment 3a and four in Experiment 3b. In each list, there were an equal number of indirect and neutral replies and equal number of video-and-audio and audio-only clips. Each selected clip was only presented in one of the two modality conditions (video-and-audio or audio-only) in each list.

Procedure and Design

The task and instruction procedures were identical to those used in Experiments 2a and 2b, except that participants only needed to categorize two types of replies (indirect and neutral), and they viewed the clips in the video-and-audio and audio-only conditions. No participant viewed the same reply clip in both modality conditions. Both reply type and modality condition were within-participants factors. The two reply types and two modality conditions were mixed and randomly presented.

Results

Table 24 presents the descriptive statistics of the mean categorization accuracy and total number of categorizations to all clips in each modality and reply type condition in Experiments 3a and 3b.

Table 24

Descriptive Statistics of the Mean Categorization Accuracy and the Proportion of Each Type of Categorizations in Each Modality and Reply Condition

		Experiment 3a		Experiment 3b	
Modality	Reply	Categorization accuracy (*SD*)	Proportion of categorizations	Categorization accuracy (*SD*)	Proportion of categorizations

video-and-audio	Indirect	.63 (.14)	50.52%	.56 (.11)	51.17%
	Neutral	.62 (.12)	49.48%	.54 (.11)	48.83%
Audio-only	Indirect	.61 (.13)	50.70%	.49 (.13)	47.00%
	Neutral	.60 (.12)	49.30%	.55 (.14)	53.00%

Categorization Performance (Accuracy) analysis

To examine whether nonverbal behaviours can help participants identify the indirect replies more accurately, we conducted binomial GLMMs, with one being a correct indirect categorization and zero being an incorrect indirect categorization. The modality (video-and-audio and audio-only) was used as the fixed factor and participants and stimuli were treated as random factors. Results showed no significant differences between the video-and-audio and audio-only conditions in Experiment 3a (*AIC* = 26171.30, *BIC* = 26203.00, *LogLik* = -13081.60, x^2 (1) = 2.42, *p* = 0.120). However, participants could identify indirect replies more accurately in the video-and-audio condition than in the audio-only condition in Experiment 3b (*AIC* = 24698.10, *BIC* = 24729.30, *LogLik* = -12345.00, x^2 (1) = 9.94, *p* = 0.002). This suggests that nonverbal behaviours can improve the accuracy in the identification of indirect replies, but only when the speech was less informative.

Categorization Performance (d prime) Analysis

To further take the response bias into account, we conducted paired-samples t-tests comparing *d'* scores between the video-and-audio and audio-only conditions to examine whether nonverbal behaviours can help participants better differentiate the indirect and neutral replies. Results revealed that, in both Experiments, *d's* were significantly higher in the video-and-audio condition (Experiment 3a: *M* = 0.70, *SD* = 0.30; Experiment 3b: *M* = 0.28, *SD* = 0.21) than in the audio-only condition (Experiment 3a: *M* = 0.58, *SD* = 0.28;

Experiment 3b: $M = 0.12$, $SD = 0.23$; Experiment 3a: $t\ (59) = 3.03$, $p = .004$, Cohen's $d = 0.39$; Experiment 3b: $t\ (59) = 3.70$, $p < .001$, Cohen's $d = 0.49$). Thus, participants could better differentiate indirect replies from neutral replies when they could see nonverbal cues in the clips than when they could not. This suggests that, regardless of the informativeness of speech, nonverbal behaviours can facilitate the detection of indirect replies when taking response bias into account.

Confidence Rating Analysis

To examine whether nonverbal behaviours can enhance participants' confidence in identifying indirect replies (regardless of the categorization accuracy), we carried out GLMMs analyses on confidence ratings of the indirect categorizations, with the presentation modality as the fixed factor and participants and stimuli as random factors. Confidence ratings of the indirect categorizations, in both Experiments, were significantly higher in the video-and-audio condition (Experiment 3a: $M = 4.75$, $SD = 0.88$; Experiment 3b: $M = 5.23$, $SD = 0.88$) than the audio-only condition (Experiment 3a: $M = 4.59$, $SD = 0.93$; Experiment 3b: $M = 5.06$, $SD = 1.01$; Experiment 3a: $AIC = 69927.40$, $BIC = 69974.90$, $LogLik = -34957.70$, $x^2\ (1) = 20.44$, $p < .001$; Experiment 3b: $AIC = 59132.80$, $BIC = 59179.70$, $LogLik = -29560.40$, $x^2\ (1) = 19.26$, $p < .001$). This suggests that nonverbal behaviours can enhance participants' confidence in identifying indirect replies, regardless of the informativeness of speech.

Predicting Indirect Categorization from Verbal and Nonverbal Cues in the Video-and-audio Condition

Table 25 presents the descriptive statistics of the predictor variables for Experiments 3a and 3b.

Table 25

Descriptive Statistics for the Predictor Variables in Experiments 3a and 3b

Predictor variable	Experiment 3a				Experiment 3b			
	M	*SD*	Min	Max	*M*	*SD*	Min	Max
Palm-revealing gesture	3.59	8.37	0	51.72	3.97	10.84	0	100
Head tilt	2.66	6.62	0	46.69	3.52	8.44	0	55.40
Facial shrug	1.46	4.48	0	35.19	1.49	5.50	0	34.68
Gaze aversion	.56	.25	0	1	.69	.24	0	1
Uncertainty terms	4.75	7.60	0	38.59	1.27	4.64	0	39.72
Discourse markers	2.76	5.70	0	34.40	3.16	7.27	0	42.86
Filled pauses	3.12	5.52	0	25.50	3.11	7.16	0	43
Silent pause	.30	.16	0	.86	.22	.09	0	.69
Reply duration (in seconds)	5.64	3.16	0.86	20.78	3.60	1.64	1.25	20.38

Tables 26 and 27 present the correlation coefficients among the predictor variables in the video-and-audio condition in Experiments 3a and 3b.

Table 26

Pearson Correlation Coefficients among the Predicator Variables in the Video-and-Audio Condition in Experiment 3a

	Variable	1	2	3	4	5	6	7	8
1.	Palm-revealing gesture	-							
2.	Head Tilt	.19***	-						
3.	Facial shrug	.19***	.24***	-					
4.	Gaze aversion	-.11*	-.02	.04	-				
5.	Uncertainty terms	.07	.17***	.03	.00	-			
6.	Discourse markers	.03	.19***	.08	.06	.06	-		
7.	Filled pauses	.06	.05	-.03	.18***	.02	.06	-	
8.	Silent pause	.03	-.04	-.06	.34***	-.16**	-.14***	.11**	-
9.	Reply duration	-.07	-.02	-.02	.33***	-.03	.18***	.19***	.16**

Note. $N = 344$; *$p < .05$, **$p < .01$, ***$p < .001$

Table 27

Pearson Correlation Coefficients among the Predicator Variables in the Video-and-Audio Condition in Experiment 3b

	Variable	1	2	3	4	5	6	7	8

1.	Palm-revealing gesture	-							
2.	Head Tilt	.35***	-						
3.	Facial shrug	.06	.12**	-					
4.	Gaze aversion	-.04	-.05	.07*	-				
5.	Uncertainty terms	.09*	.03	.03	-.08*	-			
6.	Discourse markers	.00	.04	-.01	-.04	.00	-		
7.	Filled pauses	.00	-.05	-.01	.02	-.03	-.02	-	
8.	Silent pause	-.10**	-.09*	.04	.07*	.00	-.01	-.04	-
9.	Reply duration	-.03	-.03	-.04	.16***	.11**	.16***	.04	.04

Note. $N = 608$; *$p < .05$, **$p < .01$, ***$p < .001$

To examine the relative contributions of different verbal and nonverbal cues in identifying indirect replies in the video-and-audio condition, binomial GLMMs were used for both Experiments. Participants' response was coded either as 1 (indirect categorization) or 0 (neutral categorizations). For all the models, the rate of the four nonverbal cues (i.e., palm-revealing gesture, head tilt, facial shrug, gaze aversion), the rate of the four verbal cues (i.e., uncertainty terms, discourse markers, filled pauses and silent pause), as well as the reply duration were used as fixed factors, and participants and stimuli were treated as random factors. All assumptions were met.

Binomial analyses were significant in both Experiment 3a (*AIC* = 12585.40, *BIC* = 12672.20, *LogLik* = -6280.70, x^2 (9) = 95.48, $p < .001$) and Experiment 3b (*AIC* = 11972.90,

BIC = 12058.30, *LogLik* = -5974.40, x^2 (9) = 152.66, $p < .001$). Results revealed that the rate of head tilt, discourse markers and the reply durations were significant predictors in both experiments. Furthermore, the rate of uncertainty terms was a significant predictor in Experiment 3a, and the rate of facial shrug and the proportion of silent pause were significant predictors in Experiment 3b. The relative contributions of predictor variables are shown in Table 28.

Table 28

Summary of Binomial GLMMs Analyses in the Video-and-Audio condition in Experiments 3a and 3b

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 3a	(Intercept)	-0.33	0.16	-0.65	-0.02	-2.08	0.037
	Palm Revealing	0.003	0.01	-0.01	0.02	0.55	0.583
	Head Tilt	0.02	0.01	0.01	0.04	2.55	0.011
	Facial Shrug	0.01	0.01	-0.01	0.04	1.17	0.242
	Gaze Aversion	-0.26	0.23	-0.71	0.19	-1.13	0.259
	Uncertainty terms	0.04	0.01	0.03	0.06	6.64	<.001
	Discourse Markers	0.03	0.01	0.02	0.05	3.64	<.001

	Filled Pauses	-0.002	0.01	-0.02	0.02	-0.17	0.867
	Silent Pause	0.28	0.35	-0.41	0.96	0.79	0.428
	Reply Duration	0.23	0.05	0.12	0.33	4.24	<.001
Exp. 3b	(Intercept)	-0.40	0.11	-0.62	-0.18	-3.60	<.001
	Palm Revealing	0.01	0.003	-0.001	0.01	1.62	0.105
	Head Tilt	0.01	0.004	0.001	0.02	2.21	0.027
	Facial Shrug	0.01	0.01	0.002	0.03	2.32	0.020
	Gaze aversion	0.18	0.14	-0.09	0.45	1.28	0.201
	Uncertainty terms	0.01	0.01	-0.01	0.02	1.07	0.286
	Discourse Markers	0.02	0.004	0.01	0.03	4.65	<.001
	Filled Pauses	-0.002	0.01	-0.01	0.01	-0.48	0.631
	Silent Pauses	1.23	0.36	0.53	1.94	3.41	0.001
	Reply Duration	0.32	0.03	0.25	0.38	9.03	<.001

Predicting Indirect Categorization from Verbal Cues in the Audio-only Condition

Although the main focus of the present study was on nonverbal behaviours, they are not always available in everyday communication, such as when people speak on the phone. Thus, it is also interesting to examine the independent contributions of different verbal cues in identifying indirect replies in the audio-only condition.

Tables 29 and 30 present the correlation coefficients among the predictor variables in the audio-only condition in Experiments 3a and 3b.

Table 29

Pearson Correlation Coefficients among the Predicator Variables in the Audio-Only Condition in Experiment 3a

Variable	1	2	3	4
1. Uncertainty terms	-			
2. Discourse markers	.06	-		
3. Filled pauses	.02	.06	-	
4. Silent pauses	-.16**	-.14**	.11*	-
5. Reply duration	-.03	.18***	.19***	.16**

Note. $N = 344$; *$p < .05$, **$p < .01$, ***$p < .001$

Table 30

Pearson Correlation Coefficients among the Predicator Variables in the Audio-Only Condition in Experiment 3b

Variable	1	2	3	4
1. Uncertainty terms	-			
2. Discourse markers	.00	-		
3. Filled pauses	-.03	-.02	-	

4. Silent pauses	.00	-.01	-.04	-
5. Reply duration	.11**	.16***	.04	.04

Note. $N = 608$; *$p < .05$, **$p < .01$, ***$p < .001$

To examine the relative contributions of different verbal cues in identifying indirect replies in the audio-only condition, binomial GLMMs were used for both Experiments. Participants' response was coded either as 1 (indirect categorization) or 0 (neutral categorizations). For all the models, the rate of the four verbal cues (i.e., uncertainty terms, discourse markers, filled pauses and silent pause) as well as the reply duration were used as fixed factors with participants and stimuli as random factors. All assumptions were met.

Binomial analyses were significant in both Experiment 3a (*AIC* = 12875.50, *BIC* = 12933.30, *LogLik* = -6429.70, x^2 (5) = 96.86, $p < .001$) and Experiment 3b (*AIC* = 11783.40, *BIC* = 11840.30, *LogLik* = -5883.70, x^2 (5) = 199.34, $p < .001$). Results revealed that the rate of uncertainty terms, discourse markers, and the reply durations were significant predictors in both Experiments. The relative contributions of each predictor are shown in Table 31.

Table 31

Summary of Binomial GLMMs Analyses in the Audio-Only Condition in Experiments 3a and 3b

Experiment	Fixed factor	Estimate	Std. Error	CIlow	CIhigh	z-value	Pr (>|z|)
Exp. 3a	(Intercept)	-0.29	0.13	-0.55	-0.03	-2.22	0.027
	Uncertainty terms	0.05	0.01	0.04	0.06	7.23	<.001
	Discourse Markers	0.03	0.01	0.01	0.05	3.63	<.001

	Filled Pauses	0.01	0.01	-0.004	0.03	1.45	0.147
	Silent Pause	-0.06	0.30	-0.64	0.52	-0.21	0.833
	Reply Duration	0.19	0.05	0.10	0.28	4.07	<.001
Exp. 3b	(Intercept)	-0.35	0.11	-0.56	-0.14	-3.28	0.001
	Uncertainty terms	0.02	0.01	0.01	0.03	3.48	<.001
	Discourse Markers	0.03	0.004	0.02	0.03	6.70	<.001
	Filled Pauses	0.002	0.004	-0.01	0.01	0.52	0.600
	Silent Pauses	0.48	0.32	-0.77	1.01	1.50	0.133
	Reply Duration	0.35	0.03	0.28	0.41	10.90	<.001

Discussion

The first goal of Experiments 3a and 3b was to examine whether nonverbal cues can enhance the performance and confidence in identifying indirect replies. The results of the d' performance and confidence rating analyses revealed that regardless of the informativeness of speech, participants could better discriminate indirect replies from neutral replies and were more confident in identifying indirect replies in the video-and-audio condition than in the audio-only condition. However, the results of the categorization accuracy analysis showed

that participants were only more accurate in categorizing indirect replies in the video-and-audio condition than in the audio-only condition when speech was less informative (Experiment 3b). These results indicate that nonverbal cues could help people detect indirectness and boost the confidence of their detection, especially when speech is less informative.

The second goal of Experiments 3a and 3b was to examine the relative contributions of verbal and nonverbal cues in identifying indirect replies. With regards to nonverbal cues, the GLMMs analyses in both Experiments 3a and 3b showed that head tilt was the most salient nonverbal cue for identifying indirect replies. These findings indicate that participants took head tilt as a cue of indirect replies regardless of the informativeness of speech. For facial shrug, this nonverbal behaviour was a significant predictor of the indirect categorization in Experiment 3b but not in Experiment 3a. These results suggest that people might only rely on facial shrugs to identify indirect replies when speech was less informative. For palm-revealing gestures and gaze aversion, neither of these nonverbal behaviours were significant predictors in Experiments 3a or 3b. These results indicate that participants did not rely on palm-revealing gestures or gaze aversion to identify indirect replies when verbal information was available. This might be because these two types of nonverbal cues can be used for other functions than indicating uncertainty in daily conversations. For example, a gesture with a palm revealing form can be associated with presenting, pleading, requesting and expressing openness to an abstract idea (Kendon, 2004; Muller, 2004), whereas gaze aversion can be used for reducing to the cognitive load during speaking (Doherty-Sneddon & Phelps, 2005; Glenberg et al., 1998; Phelps et al., 2006).

With regards to verbal cues, the GLMMs analyses from both video-and-audio and audio-only modalities in both Experiments 3a and 3b showed that discourse markers were a salient verbal cue in identifying indirect replies. These findings were consistent with previous

research highlighting discourse markers as a sign of uncertainty and politeness (e.g., Tree & Schrock, 2002). For uncertainty terms, this verbal behaviour was the most salient verbal cue in predicting indirect categorization proportion in Experiment 3a. This was consistent with previous literature which has shown the importance of uncertainty terms in indirect communication (Bonnefon & Villejoubert, 2009; Holtgraves & Perdew, 2016; Youmans, 2001). Uncertainty terms were not a reliable cue for identifying indirect replies in Experiment 3b, as it was a significant predictor only in the audio-only condition, but not in the video-and-audio condition. This was unsurprising because clips in Experiment 3b consisted of scripted replies and consequently, the rate of uncertainty terms was low. For silent pause, this was a significant predictor in the video-and-audio condition in Experiment 3b. This suggests that silent pause is a less salient verbal cue of indirect replies compared to uncertainty terms and discourse markers. For filled pause, this verbal behaviour was never a significant predictor of indirect categorization proportion in Experiment 3a or 3b. Thus, participants did not use it as a cue to identify indirect replies.

General Discussion

During social interactions, we often need to reply to a potential face-threatening question indirectly to save face for our conversational partner and preserve social harmony. The current research aimed to identify the verbal and nonverbal cues that are uniquely associated with the production of indirect replies and to examine the contribution of verbal and nonverbal behaviours for comprehending indirect replies. Experiments 1a and 1b showed that uncertainty terms, discourse markers, and head tilt were the verbal and nonverbal cues uniquely associated with the production of indirect replies. Experiments 2a and 2b revealed that when verbal information was not available, people took head tilt, palm-revealing gesture, facial shrug and gaze aversion as cues for indirect replies. Experiments 3a and 3b found that nonverbal cues can enhance the performance and confidence in identifying indirect replies,

especially when speech is less informative. When both verbal and nonverbal cues were available, uncertainty terms, discourse markers and head tilt were the most salient verbal and nonverbal cues for identifying indirect replies.

Nonverbal Cues and Indirect Reply

Among the four types of nonverbal cues under investigation, head tilt was the only cue that showed strong association with both the production and comprehension of indirect replies. It was produced more frequently in indirect replies than other types of replies and was a significant cue for identifying indirect replies regardless of the availability and informativeness of verbal information. Previous literature highlighted that head tilt manifests when we experience uncertainty or indecision during communication (e.g., Marono et al., 2018). In the current study, we found a significant positive relationship between the rate of head tilt and the rate of uncertainty expressions (i.e., uncertainty terms and discourse markers; see Table 26). Thus, the current results suggest head tilt was not only a marker of uncertainty, but it also serves as a communicative tool for face-saving and signifies politeness in indirect replies.

For palm-revealing gesture, facial shrug and gaze aversion, our results showed that none of these nonverbal cues were uniquely associated with the production of indirect replies. These findings suggest that these nonverbal cues might not serve for face-saving purposes in the production of indirect replies, although they have been shown as markers of uncertainty in communication (e.g., McNeill, 1992; Chovil, 1991; Cegala et al., 1979). Regarding the comprehension of indirect replies, the results suggest that the role of these nonverbal cues in identifying indirect replies was context dependent. When identifying indirect replies without any verbal information, all of them were taken as significant nonverbal cues for indirect replies. However, when both verbal and nonverbal information were available, only facial shrug was a significant cue for identifying indirect replies when the verbal information was

less informative (e.g., when the speech content was almost identical across conditions). Taken together, the current study failed to provide conclusive evidence that supports a strong link between these three types of nonverbal cues and the production and comprehension of indirect replies.

Our findings could also shed some light on the compositional view of shrug (Streeck, 2009; Morris, 1994; Givens, 1977), which proposes that shrugs consist of multiple components, including palm-revealing gesture, facial shrug, head tilt and their combinations. According to Streeck (2009), the core meaning of shrugs is to display distancing and disengagement, such as incapacity (I cannot do anything about it), inaction (I am not doing it), indifference (I don't care about it), and uncertainty (I am not sure). However, there has been evidence that different forms of shrugs are used to express different types of disengagment (Debras, 2017). For example, palm-revealing gestures have been suggested to be associated with the expression of incapacity and inaction (Müller, 2004), facial shrug is used as a disclaimer, and head tilt is associated with the expression of uncertainty (Beaupoil-Hourdel & Debras, 2017). Findings from the current study provide further insights into the compositional view of shrug and the specificity of the meanings of different shrug components. Our results showed that head tilt was the only shrug component that specifically used for expressing and comprehending uncertainty across all Experiments regardless the informativeness of speech. However, when speech was less informative (Experiment 3b), both head tilt and facial shrug were considered as cues for expression of uncertainty. Finally, when speech was absent (Experiments 2a and 2b), all three shrug components (i.e., palm-revealing gesture, facial shrug and head tilt) were used to identify the expression of uncertainty.

Verbal Cues and Indirect Reply

There has been much evidence that uncertainty terms and discourse markers play a prominent role in the production and comprehension of indirect communication (Bonnefon & Villejoubert, 2009; Holtgraves, 2016; Tree & Schrock, 2002; Youmans, 2001). The current study has further confirmed the significance of these verbal cues in both how people produce and comprehend indirect replies. Our results showed these verbal cues were uniquely associated with the production of indirect replies and were the most reliable cues for the comprehension of indirect replies. Based on previous research and the current findings, there seems to be little doubt that uncertainty terms and discourse markers signify uncertainty and serve for a politeness function in indirect replies.

Our study failed to find supporting evidence that people use filled or silent pause as a marker that signals uncertainty for face-saving purposes during the production and comprehension of indirect replies. Our results showed that the filled and silent pauses were not uniquely associated with the production of indirect replies. With regards to the comprehension of indirect replies, filled pauses were never a significant cue for indirect replies and silent pause was only taken as a cue for indirect replies in the video-and-audio condition in Experiment 3b where speech was less informative. Previous research has demonstrated that filled and silent pauses were not merely a reflection of uncertainty in speech (e.g., Brennan & Williams, 1995; Siegman & Pope, 1965; Smith & Clark, 1993; Swerts & Krahmer, 2005), but they also serve for pragmatic functions, such as signalling the speaker's intention to continue a turn in a conversation (Maclay & Osgood, 1959), to make a delay in speech (Clark & Tree, 2002), to indicate a topic change (Rendle-Short, 2004) and to end a conversation (Schegloff, 2010). However, our results suggest that there is very limited evidence that filled and silent pauses serve for politeness functions in indirect replies.

The Independent and Relative Contribution of Nonverbal Cues in Comprehending Indirect Replies

Until recently, most research on indirect replies has primarily focused on the linguistic elements while neglecting the role of nonverbal behaviours in revealing the true communicative intent in indirect replies. Results from Experiments 2a and 2b showed that participants could still correctly identify indirect replies at a rate of approximately 10% above the chance level (see Table 19) from nonverbal cues alone, and all four types of nonverbal cues significantly contributed to the identification of indirect replies. To our knowledge, our study was the first to investigate the independent contribution of nonverbal cues for comprehending indirect replies. Our findings clearly showed the value of nonverbal behaviours as cues for revealing the true communicative intent in indirect replies.

In everyday interaction, however, people usually do not communicate in silence. To fully understand the role of nonverbal cues in comprehending indirect replies, it is necessary to understand their relative contribution alongside the verbal cues. Experiments 3a and 3b first compared participants' performance and confidence in categorizing indirect replies between the video-and-audio condition and the audio-only condition. It has been found that nonverbal cues significantly boosted the performance and confidence not only when speech was uninformative but also when it was informative. These results are clearly consistent with the view that verbal and nonverbal information together can create a deeper and more meaningful message than just the verbal information alone (e.g., de Gelder & Vroomen, 2000; Kelly et al., 2010). The findings showed the importance for considering communication as a multi-modal system including both the verbal and nonverbal channels (McNeill, 1992; Clark, 1996; Kendon, 1994).

Our results, however, also revealed that people do not always automatically integrate all verbal and nonverbal cues when comprehending indirect messages. We found that the

extent to which people relied on nonverbal cues to identify indirect replies depends on the informativeness of the speech. The more informative the speech was, the fewer types of nonverbal cues that contributed to the identification of indirect replies. Furthermore, we also found evidence that people relied on fewer types of verbal cues to identify indirect replies when nonverbal cues were available than when they were not. Results from Experiment 3b showed that uncertainty terms were a significant verbal cue for comprehending indirect replies in the audio-only condition but not in the video-and-audio condition.

The Production-Comprehension Link of Verbal and Nonverbal Behaviours

Although the current study was not designed to directly test whether language production and comprehension are separate or a single system, some of our results are clearly consistent with the unified views of the production and comprehension of verbal and nonverbal behaviours, which proposes that comprehending speech and nonverbal behaviours requires simulating the production of the perceived speech and nonverbal behaviours (e.g., Alibali & Hostetter, 2010; Pickering & Garrod, 2013). We found that uncertainty terms, discourse markers and head tilt were uniquely associated with the production of indirect replies, and they were also prominent cues of indirectness during the comprehension of indirect replies. These findings suggest that people not only use these cues to express uncertainty and politeness when producing indirect replies, but they also rely on these cues to identify the indirect replies during comprehension.

Much experimental evidence of the unified views comes from studies using prediction-encouraging and highly restricted lab-based tasks (e.g., Dell & Chang, 2014; Chater et al., 2016; Pickering & Garrod, 2007, 2013). The present study extended these findings to a more “real world” situation by showing a strong link between the production and comprehension of both verbal and nonverbal behaviours in spontaneous conversations. To provide more insights into this issue, future studies could use the same set of participants

to examine whether people who produce certain verbal and nonverbal behaviours when producing indirect replies also rely on the same behaviours when comprehending indirect messages.

Despite the evidence supporting a unified system, the current investigation also found evidence of mismatches between the production and comprehension of verbal and nonverbal behaviours. For example, when attempting to identify indirect replies, silent pause was used when nonverbal information was unavailable and palm-revealing gestures, facial shrug and gaze aversion were used when verbal information was unavailable. None of these cues, however, were uniquely associated with the production of the indirect replies. Thus, people do not always rely on the verbal and nonverbal cues that they used in the production to identify indirect replies. Similar mismatches have also been found in deception research. Vrij and Semin (1996) found that professionals such as police detectives, prison guards and customs officers frequently used averted eye gaze to detect deception, whereas meta-analyses have revealed a non-significant relationship between gaze aversion and lie production (DePaulo et al., 2003; Sporer & Schwandt 2007; Vrij, 2008). These mismatches suggest that people also rely on their belief or knowledge to interpret the meaning of verbal or nonverbal behaviours.

Limitations and Future Directions

There are several limitations to the present study which are important to consider. Firstly, the stimulus used in the comprehension experiments (Experiments 2a to 3b) were based on the video recording of different types of replies produced in the production experiments (Experiments 1a and 1b). This allowed us to manipulate the presentation modality of the same set of replies to test the independent and relative contributions of verbal and nonverbal cues in comprehending indirect replies. However, it can be argued that these recorded clips may be less ecologically valid than live face-to-face interactions. Risko et al.

(2012) pointed out that video clips, although 'socially relevant,' might be lacking in some aspects of real-life social interactions. Järveläinen et al. (2001) found that although observing videotaped hand movements clearly activated the primary motor cortex, the activation was stronger during the observation of live rather than video recorded hand movements. This is presumably because live hand movement was more representative of actions in daily life and thus more likely to improve participant's attention, motivation, or interest. To increase the ecological validity, future studies could use a live face-to-face interaction paradigm to study the role of verbal and nonverbal cues in comprehending indirect messages.

Secondly, all replies were produced by native English speakers in the present study. Thus, the extent to which the current findings can be generalized to speakers of other languages or cultures remains unclear. Tannen (1981) pointed out that indirect communication varied across different cultures and can often cause confusion and misunderstanding. It has been shown that speakers from high-context cultures (e.g., Japan, Korea) are more likely to speak indirectly and to look for indirect meanings than speakers from low-context cultures (e.g., United States, Germany; Hall, 1976; Holtgraves, 1997). Based on these findings, speakers from high-context cultures, as compared to those from low-context cultures, may use more verbal and nonverbal cues when producing indirect replies and may be better at decoding indirectness from these cues. In addition, previous research has found that the way we observe nonverbal behaviours varies between cultures (Matsumoto, 2007). For example, Jack et al. (2009) found that when decoding different emotions from facial expressions, Western Caucasian observers distributed their eye fixations evenly across the face, whereas Eastern Asian observers persistently fixated on the eye region. Therefore, future research should further investigate how cultural factors can influence the use of verbal and nonverbal behaviours in the production and comprehension of indirect messages.

Thirdly, although the results of the present study clearly showed that verbal and nonverbal cues expressing uncertainty play a crucial role in conveying and comprehending indirect replies, there are certainly more components that contribute to politeness and indirectness. For example, there has been ample evidence that prosody conveys a vast variety of meaning during communication (Prieto, 2015). Characteristic changes in pitch, loudness, duration of an utterance or voice quality of the speaker are instrumental in modulating the attitudinal stance of the speaker towards the listener (e.g. Brown and Prieto, 2017). For example, confident voices are associated with increased intensity, and increased variation in both intensity and pitch, whereas doubtful voices tend to display a higher pitch (Jiang & Pell, 2017). Furthermore, Caballero et al. (2018) found that a rude request was marked by a lower pitch and tended to display a falling pitch contour, whereas a polite request had a higher pitch and a rising pitch contour. In addition to prosody, violation of the Maxim of Relevance rule is also an important cue for indirect replies (Grice, 1975; Searle, 1975). That is, indirect replies are often unrelated to their respective questions. In a recent study, Boux et al. (in press) showed that people perceive indirect replies as less coherent, predictable and less semantically similar to their respective questions that direct replies with the same surface meaning. Thus, future studies should further investigate the contribution of crucial components and their interaction with uncertainty verbal and nonverbal cues in the delivery and decoding of indirect replies.

Finally, we cannot estimate the sample size of each experiment based on previous research as to our knowledge no experimental studies have examined various verbal and nonverbal behaviours simultaneously in the production and comprehension of indirect replies. Therefore, the sample size of each experiment was determined by the number of participants we could recruit in each academic term. However, each pair of experiments had an identical design and procedure except that one involved spontaneous replies and the other

involved scripted replies, and therefore they could serve as a replication of each other. Since the results were mostly replicated in each pair of experiments, they were unlikely to be underpowered.

Conclusion

Our results showed that among different verbal and nonverbal cues, uncertainty terms, discourse markers and head tilt were those that were most closely related to indirect replies. They were not only the unique cues used by the speakers during the production of indirect replies, but they were also the most reliable cues used by the listeners during the comprehension of indirect replies. Moreover, it was found that nonverbal cues facilitated the identification of indirect replies and the extent to which people relied on verbal and nonverbal cues to identify indirect replies was context dependent. The more informative the verbal/nonverbal information was, the fewer types of nonverbal/verbal cues that contributed to the identification of indirect replies. Taken together, findings from the current research could provide an initial step towards developing a comprehensive and unified model of the production and comprehension of indirect replies, which would take both verbal and nonverbal behaviours into account. In addition, the outcomes of the current research could be used to inform the general public as well as professionals on how to express opinions and concerns more tactfully and to “read” the verbal and nonverbal cues in indirect communication, which would allow greater levels of trust and avoid conflict and tension due to misunderstandings of the verbal and nonverbal cues.

References

Alibali, M. W., & Hostetter, A. B. (2010). Mimicry and simulation in gesture comprehension. *Behavioral and Brain Sciences, 33(6),* 433. http://doi.org/10.1017/S0140525X10001445

Argyle, M. (1975). The syntaxes of bodily communication. In J. Benthal & T. Polhemus (Eds.), *The body as a medium of expression* (pp. 143–161). E.P. Dutton & Co., Inc. http://doi.org/10.1515/ling.1973.11.112.71

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Linguistic Data Consortium.

Barr, D. J. & Kelly, S. D. (1997) Nonverbal factors in understanding and remembering indirect requests. In M. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Bašnáková, J., Weber, K., Petersson, K. M., van Berkum, J., & Hagoort, P. (2014). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex, 24(10),* 2572-2578. http://doi.org/10.1093/cercor/bht112

Bates, E. (1976). *Language and context: The acquisition of pragmatics.* Academic Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint.* arXiv:1406.5823. https://doi.org/10.48550/arXiv.1406.5823

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse processes, 15(4),* 469-489. http://doi.org/10.1080/01638539209544823

Beattie, G. W. (1978). Sequential temporal patterns of speech and gaze in dialogue. *Semiotica, 23(1-2),* 29-52. http://doi.org/10.1515/semi.1978.23.1-2.29

Beaupoil-Hourdel, P., & Debras, C. (2017). Developing communicative postures: The emergence of shrugging in child communication. *Language, Interaction and Acquisition, 8(1),* 89-116. https://doi.org/10.1075/lia.8.1.05bea

Bitti, P. E. R., Bonfiglioli, L., Melani, P., Caterina, R., & Garotti, P. (2014). Expression and communication of doubt/uncertainty through facial expression. *Journal of Theories and Research in Education*, *9*(1), 159-177. https://doi.org/10.6092/issn.1970-2221/4296

Bonnefon, J. F., Feeney, A., & De Neys, W. (2011). The risk of polite misunderstandings. *Current Directions in Psychological Science, 20*(5), 321-324. http://doi.org/10.1177/0963721411418472

Bonnefon, J. F., & Villejoubert, G. (2006). Tactful or doubtful? Expectations of politeness explain the severity bias in the interpretation of probability phrases. *Psychological Science, 17(9),* 747-751. https://doi.org/10.1111/j.1467-9280.2006.01776.x

Bonnefon, J. F., Feeney, A., & Villejoubert, G. (2009). When some is actually all: Scalar inferences in face-threatening contexts. *Cognition, 112(2),* 249-258. http://doi.org/10.1016/j.cognition.2009.05.005

Boux, I. P., Margiotoudi, K., Dreyer, F. R., Tomasello, R., & Pulvermüller, F. (2022). Cognitive features of indirect speech acts. *Language, Cognition and Neuroscience,* 1-25.# http://doi.org/10.1080/23273798.2022.2077396

Brennan, S. E., & Williams, M. (1995). The feeling of another′ s knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of memory and language, 34(3),* 383-398. https://doi.org/10.1006/jmla.1995.1017

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge university press. http://doi.org/10.1017/CBO9780511813085

Brown, L., & Prieto, P. (2017). (Im) politeness: Prosody and gesture. In *The Palgrave handbook of linguistic (im) politeness (pp. 357-379).* Palgrave Macmillan, London. http://doi.org/10.1057/978-1-137-37508-7_14

Caballero, J. A., Vergis, N., Jiang, X., & Pell, M. D. (2018). The sound of im/politeness. *Speech Communication*, *102*, 39-53. http://doi.org/10.1016/j.specom.2018.06.004

Cegala, D. J., Sokuvitz, S., & Alexander, A. F. (1979). An investigation of eye gaze and its relation to selected verbal behavior. *Human Communication Research, 5(2),* 99-108. http://doi.org/10.1111/j.1468-2958.1979.tb00625.x

Chater, N., McCauley, S. M., & Christiansen, M. H. (2016). Language as skill: Intertwining comprehension and production. *Journal of Memory and Language, 89,* 244-254. http://doi.org/10.1016/j.jml.2015.11.004

Chovil, N. (1991). Discourse-oriented facial displays in conversation. *Research on Language & Social Interaction*, *25*(1-4), 163-194. http://doi.org/10.1080/08351819109389361

Chu, M., Tobin, P., Ioannidou, F. & Basnakova, J. (2022, April 22). *Encoding and Decoding Hidden Meanings in Face-to-Face Communication: Understanding the Role of Verbal and Nonverbal Behaviours in Indirect Replies.* Retrieved from osf.io/ngsp6

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694. http://doi.org/10.1037/a0033861

Clark, H. H. (1996). *Using language*. Cambridge university press. http://doi.org/10.1017/CBO9780511620539

Clark, H. H., & Schunk, D. H. (1980). Polite responses to polite requests. *Cognition, 8(2),* 111-143. http://doi.org/10.1016/0010-0277(80)90009-8

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition, 84(1),* 73-111. http://doi.org/10.1016/S0010-0277(02)00017-3

Cooperrider, K., Abner, N., & Goldin-Meadow, S. (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Frontiers in Communication, 3,* 23. http://doi.org/10.3389/fcomm.2018.00023

Debras, C. (2017). The shrug: Forms and meanings of a compound enactment. *Gesture, 16(1),* 1-34. http://doi.org/10.1075/gest.16.1.01deb

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1634),* 20120394. http://doi.org/10.1098/rstb.2012.0394

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion, 14(3),* 289-311. https://doi.org/10.1080/026999300378824

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin, 129(1),* 74-118. https://doi.org/10.1037/0033-2909.129.1.74

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & cognition, 33(4),* 727-733. https://doi.org/10.3758/BF03195338

Ferré, G. (2011). Functions of three open-palm hand gestures. *Multimodal Communication, 1(1),* 5-20. https://doi.org/10.1515/mc-2012-0002

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third edition). Sage, Thousand Oaks CA.

Fraser, B. (2010). Pragmatic Competence: The Case of Hedging. In Kaltenbock, G., Mihatsch, W., & Schneider, S. (Eds.), *New Approaches to Hedging* (pp. 15-34). Emerald Group Publishing Ltd. https://doi.org/10.1163/9789004253247_003

Givens, D. (1977). Shoulder shrugging: a densely communicative behavior. *Semiotica, 19,* 13–29. https://doi.org/10.1515/semi.1977.19.1-2.13

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & cognition, 26(4),* 651-658. https://doi.org/10.3758/BF03211385

Grice, H. P. (1975). Logic and conversation. In *Speech acts (pp. 41-58).* Brill. https://doi.org/10.1163/9789004368811_003

Hall, E. (1976). *Beyond culture.* New York: Anchor Press/Double day.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics, 38(4),* 555-568. https://doi.org/10.1016/j.wocn.2010.08.002

Holle, H., Gunter, T. C., Rüschemeyer, S. A., Hennenlotter, A., & Iacoboni, M. (2008). Neural correlates of the processing of co-speech gestures. *Neuroimage, 39(4),* 2010-2024. https://doi.org/10.1016/j.neuroimage.2007.10.055

Holtgraves, T. (1997). Styles of language use: Individual and cultural variability in conversational indirectness. *Journal of personality and social psychology, 73(3),* 624-637. https://doi.org/10.1037/0022-3514.73.3.624

Holtgraves, T. (1998). Interpreting indirect replies.#*Cognitive Psychology*,#*37*(1), 1-27. https://doi.org/10.1006/cogp.1998.0689

Holtgraves, T. (1999). Comprehending indirect replies: When and how are their conveyed meanings activated? *Journal of Memory and Language, 41(4),* 519-540. https://doi.org/10.1006/jmla.1999.2657

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, *58*(3), 627-645. https://doi.org/10.1016/j.jml.2007.06.001

Holtgraves, T., & Joong-Nam, Y. (1990). Politeness as universal: Cross-cultural perceptions of request strategies and inferences based on their use. *Journal of personality and social psychology, 59(4),* 719. https://doi.org/10.1037/0022-3514.59.4.719

Holtgraves, T., & Perdew, A. (2016). Politeness and the communication of uncertainty. *Cognition*, *154*, 1-10. https://doi.org/10.1016/j.cognition.2016.05.005

Holtgraves, T., & Yang, J. N. (1992). Interpersonal underpinnings of request strategies: General principles and differences due to culture and gender. *Journal of personality and social psychology, 62(2),* 246. https://doi.org/10.1037/0022-3514.62.2.246

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience, 31(1),* 19-31. https://doi.org/10.1080/23273798.2015.1072223

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current biology, 19(18),* 1543-1548. https://doi.org/10.1016/j.cub.2009.07.051

Järveläinen, J., Schürmann, M., Avikainen, S., & Hari, R. (2001). Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts. *Neuroreport,* 12(16), 3493-3495. https://doi.org/10.1097/00001756-200111160-00024

Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, *88*, 106-126. https://doi.org/10.1016/j.specom.2017.01.011

Juanchich, M., & Sirota, M. (2013). Do people really say it is "likely" when they believe it is only "possible"? Effect of politeness on risk communication. *Quarterly Journal of Experimental Psychology, 66(7),* 1268–1275. https://doi.org/10.1080/17470218.2013.804582

Kelly, S. D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension.#*Journal of Child Language*,#*28*(02), 325-349. https://doi.org/10.1017/S0305000901004664

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of memory and Language, 40(4),* 577-592. https://doi.org/10.1006/jmla.1999.2634

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*(2), 260-267. https://doi.org/10.1177/0956797609357327

Kendon, A. (1994). Do gestures communicate? A review. *Research on language and social interaction, 27(3),* 175-200. https://doi.org/10.1207/s15327973rlsi2703_2

Kendon, A. (1997). Gesture. *Annual review of anthropology, 26(1),* 109-128. https://doi.org/10.1146/annurev.anthro.26.1.109

Kendon, A. (2004). *Gesture: Visible action as utterance.* Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511807572

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software,* 82, 1-26. https://doi.org/10.18637/jss.v082.i13

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, *27*(3), 145-162. https://doi.org/10.1023/A:1025389814290

Limberg, H. (2009). Impoliteness and threat responses. *Journal of pragmatics, 41(7),* 1376-1394. https://doi.org/10.1016/j.pragma.2009.02.003

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15(1),* 19-44. https://doi.org/10.1080/00437956.1959.11659682

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., van Bentham, K., Bolker, B.M., and Brooks, M. E. (2017). *glmmTMB: Generalized Linear Mixed Models using Template Model Builder*. R package version 0.1.3. https://github.com/glmmTMB

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of experimental child psychology, 126,* 264-279. https://doi.org/10.1016/j.jecp.2014.05.004

Marono, A., Clarke, D. D., Navarro, J., & Keatley, D. A. (2018). A sequence analysis of nonverbal behaviour and deception. *Journal of Police and Criminal Psychology, 33(2),* 109-117. https://doi.org/10.1007/s11896-017-9238-9

Matsumoto, D. (2007). Culture, context, and behavior. *Journal of personality, 75(6),* 1285-1320. https://doi.org/10.1111/j.1467-6494.2007.00476.x

McNeill, D. (1985). So you think gestures are nonverbal? Psychological Review, 92(3), 350–371. https://doi.org/10.1037/0033-295X.92.3.350

McNeill, D. (1987). *Psycholinguistics: A new approach.* Harper & Row Publishers.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press.

McNeill, D. (1998). Speech and gesture integration. In J. M.Iverson, S. Goldin-Meadow, et al. (Eds.), *The nature and functions of gesture in children's communication* (pp. 11-27). Jossey-Bass. https://doi.org/10.1002/cd.23219987902

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian

low and high literates. *Journal of Eye Movement Research, 5(1),* 1-10. https://doi.org/10.16910/jemr.5.1.3

Morris, D. (2015). *Bodytalk: A world guide to gestures.* Random House.

Müller, C. (2004). Forms and uses of the Palm Up Open Hand: A case of a gesture family? In Cornelia Müller & Roland Posner (Eds.), *The semantics and pragmatics of everyday gestures, Proceedings of the Berlin conference, April 1998 (pp. 233–256).* Berlin: Weidler Buchverlag.

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and brain sciences, 33(6),* 417. https://doi.org/10.1017/S0140525X10000865

Phelps, F., Doherty-Sneddon, G., & Warnock, H. (2006). Helping children think: Gaze aversion and teaching. *British Journal of Developmental Psychology, 24,* 577-588. https://doi.org/10.1348/026151005X49872

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in cognitive sciences, 11(3),* 105-110. https://doi.org/10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences, 36(4),* 329-347. https://doi.org/10.1017/S0140525X12001495

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech.#*Proceedings of the National Academy of sciences,*#*105*(3), 833-838. https://doi.org/10.1073/pnas.0707192105

Prieto, P. (2015). Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science, 6(4),* 371-381. https://doi.org/10.1002/wcs.1352

Rendle-Short, J. (2004). Showing structure: Using um in the academic seminar. *Pragmatics, 14(4),* 479-498. https://doi.org/10.1075/prag.14.4.04ren

Risko, E. F., Laidlaw, K. E., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in human neuroscience,* 6, 143. https://doi.org/10.3389/fnhum.2012.00143

Schegloff, E. A. (2010). Some other "uh (m)" s. *Discourse Processes, 47(2),* 130-174. https://doi.org/10.1080/01638530903223380

Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 59–82). New York: Academic Press. http://dx.doi.org/10.1017/CBO9780511609213.004

Siegman, A. W., & Pope, B. (1965). Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. *Journal of Personality and Social Psychology, 2(4),* 522. https://doi.org/10.1037/h0022491

Stone, M., & Oh, I. (2008). Modeling facial expression of uncertainty in conversational animation. In Wachsmuth, I., & Knoblich, G. (Eds.), *Modeling communication with robots and virtual humans* (pp. 57-76). Berlin: Heidelberg Springer. http://dx.doi.org/10.1007/978-3-540-79037-2_4

Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning (Vol. 2).* John Benjamins Publishing. https://doi.org/10.1075/gs.2

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of memory and language, 32(1),* 25-38. https://doi.org/10.1006/jmla.1993.1002

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13(1),* 1-34. https://doi.org/10.1037/1076-8971.13.1.1

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language, 53(1),* 81-94. https://doi.org/10.1016/j.jml.2005.02.003

Takeuchi, A., & Nagao, K. (1993). Communicative facial displays as a new conversational modality. In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems* (pp. 187-193). https://doi.org/10.1145/169059.16915

Tannen, D. (1981). Indirectness in discourse: Ethnicity as conversational style. *Discourse processes, 4(3),* 221-238. https://doi.org/10.1080/01638538109544517

Tottie, G. (2014). On the use of uh and um in American English. *Functions of Language, 21(1),* 6-29. https://doi.org/10.1075/fol.21.1.02tot

Tree, J. E. F., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics, 34(6),* 727-747. https://doi.org/10.1016/S0378-2166(02)00027-9

Vigliocco G, Perniss P, & Vinson D. (2014) Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B, 369,* 20130292. https://doi.org/10.1098/rstb.2013.0292

Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of nonverbal behavior, 20(1),* 65-80. https://doi.org/10.1007/BF02248715

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

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of open source software,* 4(43), 1686. https://doi.org/10.21105/joss.01686

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. In: *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation.*

Youmans, M. (2001). Cross-cultural differences in polite epistemic modal use in American English. *Journal of Multilingual and Multicultural Development, 22*, 57–73. https://doi.org/10.1080/01434630108666425

Figure 1

Setup for Experiments 1a and 1b

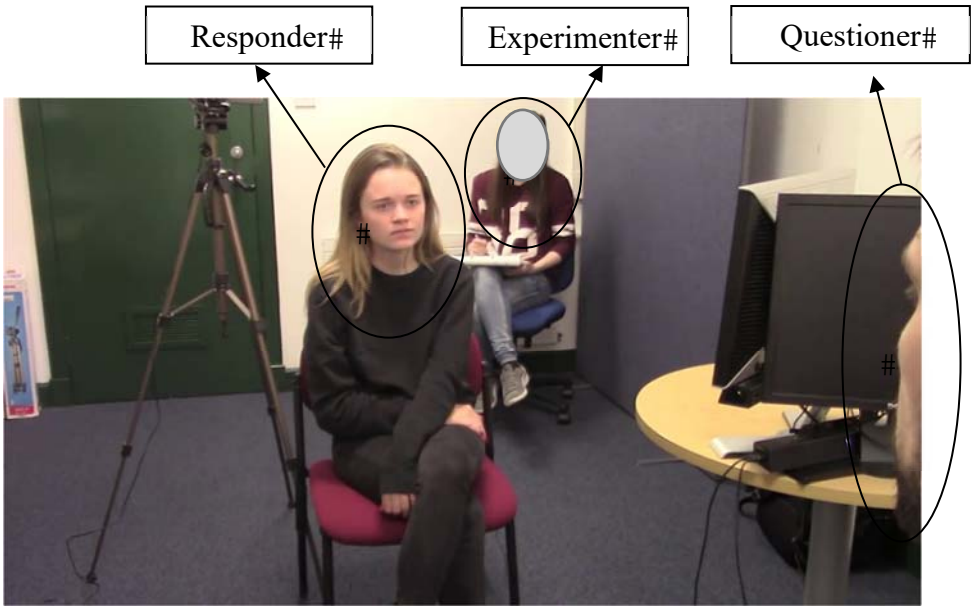


Figure 2

Examples of (a) Palm-revealing Gesture, (b) Head Tilt, and (c) Facial Shrug

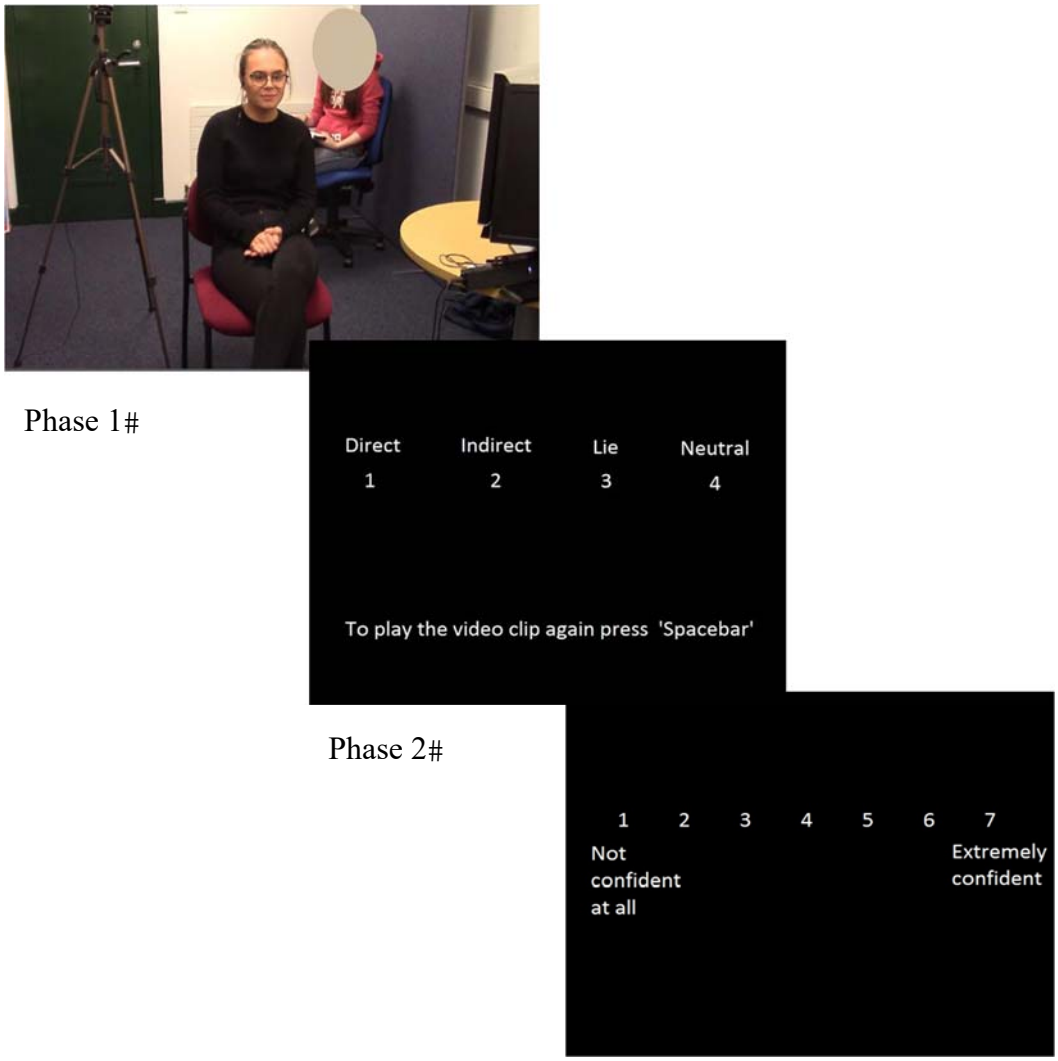
a) palm-revealing gesture

b) head tilt

c) facial shrug

Figure 3

An Illustration of the Three Phases in a Trial



Phase 1#

Phase 2#

Phase 3#