It Doesn't Hurt to Ask:

Question-Asking Increases Liking

Karen Huang, Mike Yeomans, Alison Wood Brooks, Julia Minson, Francesca Gino

Harvard University

**** PRELIMINARY MANUSCRIPT ****

**** DO NOT CITE OR CIRCULATE WITHOUT PERMISSION ****

Abstract

Conversation is a fundamental human experience, one that is necessary to pursue intrapersonal and interpersonal goals across myriad contexts, relationships, and modes of communication. In the current research, we isolate the role of an understudied conversational behavior: question-asking. Across three studies of live dyadic conversations, we identify a robust and consistent relationship between question-asking and liking: people who ask more questions are better liked by their conversation partners. When people are instructed to ask more questions, they are perceived as higher in responsiveness, an interpersonal construct that captures listening, understanding, validation, and care. We measured responsiveness with an attitudinal measure from previous research as well as a behavioral measure: the number of follow-up questions one asks. In both cases, responsiveness explained the effect of question-asking on liking. In addition to analyzing live get-to-know-you conversations online, we also studied face-to-face speed-dating conversations. We find that speed daters who ask more questions during their dates are more likely to get second dates, a behavioral indicator of liking. We trained a natural language processing algorithm as a "follow-up question detector" that we applied to our speed-dating data (and can be applied to any text data to more deeply understand question-asking dynamics). The follow-up question rate established by the algorithm explained why question-asking led to speed-dating success. We also find that despite the persistent and beneficial effects of asking questions, people do not anticipate that question-asking increases interpersonal liking.

Keywords: Question-asking; Liking; Responsiveness; Conversation; Natural Language Processing

It Doesn't Hurt to Ask:

Question-Asking Increases Liking

Imagine this scenario: you meet a new colleague for the first time at a company party. You strike up a conversation, and the colleague tells you a funny story. You are interested and engaged, and you ask several questions during the story. After the story is over, you exchange pleasantries and part ways. Later you realize that your colleague didn't ask any questions about you, and you didn't reveal much information about yourself. Who made the better impression?

Conversation is a fundamental human experience, one that is necessary for pursuing intrapersonal and interpersonal goals across myriad contexts, relationships, and modes of communication (e.g., written, spoken). We converse with others to learn what they know—their information, stories, preferences, ideas, thoughts, and feelings—as well as to share what we know while managing others' perceptions of us. That is, two central goals of conversation are information exchange and impression management. In this paper, we examine an understudied conversational behavior that likely influences both of these goals: question-asking.

Although question-asking is ubiquitous, we know very little about the antecedents and consequences of asking questions during interpersonal interaction. In the current research, we investigate the psychology of question-asking as a social phenomenon. We measure people's natural rates of question-asking and explore how the propensity to ask questions influences interpersonal liking across experimental laboratory settings and an observational field setting. Compared to people who ask few questions, we expect that high question askers are better liked. Asking a question is an other-oriented conversational tool that both causes and conveys better listening, understanding, validation, and care (i.e., responsiveness, Reis & Patrick, 1996; Reis,

Maniaci, Caprariello, Eastwick, & Finkel, 2011). The question asker's responsiveness, in turn, is likely to cause him or her to be better liked by the question answerer.

Question-Asking in Conversation

A conversation is a cooperative interaction between people in which each person acts in coordination to contribute to a successful experience of shared understanding (Clark & Schaefer, 1989). It is an ongoing, sequential unfolding of actions and responses (Reis & Patrick, 1996), organized as speaker turns (Schegloff & Sacks, 1973). Most conversations are characterized by the transfer of information about beliefs, thoughts, or emotions from one person to another (Epley & Waytz, 2010).

There are many ways people can express themselves during a conversation, such as making a statement, telling a story, making a quip or joke, apologizing, or giving a compliment (Clark & Schaefer, 1989). In the current work, we explore a ubiquitous but understudied conversational tool: question-asking.

Question-asking directs conversations by encouraging another person to answer (Dillon, 1982, 1988). Though some people may ask questions to avoid disclosing information themselves, most questions function to solicit information from others (Chafe, 1970; Dillon, 1982; Kearsley, 1976). If one person asks a question, the other person should follow with a response that abides by basic conversational maxims (Graesser, 1985; Grice, 1975), such as responding at an appropriate length, and responding to the question at hand (Hilton, 1990), though some recent work suggests that people often violate these norms by dodging questions, responding with truth that is deliberately misleading (i.e., paltering), or refusing to answer altogether (John, Barasz, & Norton, 2016; Rogers & Norton, 2011; Rogers, Zeckhauser, Gino, Norton, & Schweitzer, 2016).

The type of question-asking we investigate—natural, conversational questions that elaborate on the question-responder's statements—differ categorically from the questions investigated in studies on experimentally-induced social closeness (e.g., Aron, Aron, Tudor, & Nelson, 1991; Aron, Melinat, Aron, Vallone, & Bator, 1997; Sedikides, Campbell, Reader, & Elliot, 1999; Sprecher, Treger, Wondra, Hilaire, & Wallpe, 2012). This work has defined social closeness as the inclusion of the other in the self (Aron, Aron, & Smollan, 1992; Aron et al., 1991). In this work, participants were instructed to ask a fixed list of questions that change topic but increase in intimacy over time, and partners take turns answering all questions (e.g., Aron et al., 1997). For example, each partner would take turns asking and answering the question "What do you value most in a friendship?" before moving on to asking and answering the question "What is your most treasured memory?" (Aron et al., 1997). In these studies, questions were provided by an experimenter, and participants were not instructed or encouraged to ask follow-up questions. In contrast, in our work, we investigate the effect of question-asking on liking in natural dyadic interactions.

In our research, we focus on information-seeking questions (e.g., Miles, 2013; Van der Meij, 1987) in which the question-asker is missing a piece of information and, by asking a question, seeks the answer from another person. People often ask information-seeking questions when meeting for the first time (Berger & Calbrese, 1975), and are more likely to seek information from others when they consider the information to have high utility (Swann, Stephenson, & Pittman, 1981). Because people often know very little about each other upon first meeting, individuals stand to learn a large amount about their conversation partners during first encounters.

Importantly, though, information exchange is not the only goal of conversation. Asking questions may serve and influence other motivations like impression management.

Question-Asking and Liking

Most people have an intrinsic desire to be liked by others (Baumeister, 1982; Jones & Pittman, 1982; Leary & Kowalski, 1990). Being liked by others influences interpersonal attraction, relationship development (Berscheid, 1985; Berscheid & Regan, 2005), and other important predictors of attraction such as in-group favoritism (Reis & Patrick, 1996).

It is important to examine conversation as a process that influences attraction (Davis & Perkowitz, 1979) and relationship development (Reis & Shaver, 1988; Miller, Berg, & Archer, 1983) because the content of a conversation can significantly influence the extent to which the participants like each other when the conversation is over. The effect of conversational content on interpersonal liking has been demonstrated across a wide array of conversational strategies, ranging from other-focused behaviors, such as giving a compliment or acknowledging another person's ideas, to self-focused behaviors, such as talking about oneself (Godfrey, Jones, & Lord, 1986; Laurenceau, Barrett, & Pietromonaco, 1998; Rosenfeld, 1966; Sprecher et al., 2013). However, to our best knowledge, no prior research has investigated whether and how asking questions may influence liking.

Though asking questions invites information disclosure, there are many reasons why people may not ask questions. First, people may not think to ask questions at all. Neglecting to ask questions altogether may happen because people are egocentric—focused on expressing their own thoughts, feelings, and beliefs (e.g., Gilovich, Medvec, & Savitsky, 2000) with little or no interest in hearing what another person has to say. On the other hand, some people may think to

ask questions, but may purposefully forgo asking because they are unsure about which question(s) to ask or worry about asking a question that is perceived as rude, inappropriate, intrusive, or incompetent. In these cases, it may be much easier to talk about oneself instead.

Indeed, in most conversations, people predominantly share information about themselves rather than discussing other possible topics (Landis & Burtt, 1924). A study of conversations in public settings like bars and trains suggests that people spend two-thirds of conversation time talking about their personal experiences (Dunbar, Marriott, & Duncan, 1997). Especially when meeting someone new, people tend to use self-focused presentation strategies like self-promotion (Godfrey et al., 1986). For example, Marr and Cable (2014) found that job candidates excessively attempt to "sell" themselves to make a favorable impression in job interviews.

The tendency to focus on the self when trying to impress others is misguided, as verbal behaviors that focus on the self, such as redirecting the topic of conversation to oneself, bragging, boasting, or dominating the conversation, tend to decrease liking (Berman, Levine, Barasch, & Small, 2015; Godfrey et al., 1986; Sezer, Gino, & Norton, 2015; Vangelisti, Knapp, & Daly, 1990). In contrast, verbal behaviors that focus on the other person, such as mirroring the other person's mannerisms (Ireland & Pennebaker, 2010), affirming the other's statements, or coaxing information from the other person, have been shown to increase liking (Godfrey et al., 1986; Rosenfeld, 1966).

We hypothesize that question-asking is an other-focused conversational strategy that is likely to increase liking for the question asker. This hypothesis is consistent with prior research. For example, at the trait level, people who tend to draw out more information from their conversation partners (termed "openers") are better liked by their partners in long-term relationships

(Miller et al., 1983). And studies of doctor-patient communication suggest that patients report

higher satisfaction with their visits when physicians ask more questions about the patients' expe-

riences (Bertakis, Roter, & Putnam, 1991; Robinson & Heritage, 2006). Furthermore, because

most people spend the majority of their conversations sharing their own views rather than focus-

ing on the other person, we hypothesize that people do not anticipate the effect of question-ask-

ing on liking.

Responsiveness Mediates the Effect of Question-Asking on Liking

We suggest that asking questions increases liking because doing so indicates *responsive-*

ness, a desirable interpersonal construct identified by prior research that encompasses the verbal

and nonverbal behaviors that fulfill the needs and wishes of one's conversation partner (Miller &

Berg, 1984; Davis, 1982). Responsive behavior in a conversation requires a set of skills for re-

sponding relevantly and appropriately. We argue that question-asking is one conversational be-

havior that is likely to convey high responsiveness.

Reis and Shaver (1988) developed a model of interpersonal intimacy that defines respon-

siveness as reflecting three components: understanding, validation, and care for the partner. First,

the understanding component of responsiveness refers to accurately comprehending the question-

responder's self-perceptions—their needs, goals, beliefs, emotions, and life situation (Reis &

Patrick, 1996; Reis & Shaver, 1988). By asking questions, one elicits information from the part-

ner, including facts, attitudes, preferences, and emotional expressions, which help to more accu-

rately and appropriately understand one's partner. Understanding cannot take place without being

well-informed about one's partner (Reis & Patrick, 1996), and question-asking is likely to in-

crease the disclosure and learning necessary for understanding.

Second, the validation component of responsiveness is defined as valuing and respecting the partner's self-perceptions and perspectives (Reis & Patrick, 1996; Reis & Shaver, 1988). Validation also involves affirming that the partner is accepted and valued (Reis & Shaver, 1988). We suggest that asking questions communicates respect and value for the partner's perspective. Ironically, even without responding to the partner with direct validation or affirmation, question-asking itself may be seen as a form of positive approval or validation (Cozby, 1973). By asking questions, you acknowledge that the partner's perspective is valuable enough that you want to know more. By soliciting more information from the partner, asking a question expresses interest in the partner's viewpoint (Chen, Minson, & Tormala, 2010). Indeed, previous research suggests that effective validation in marital communication can be successfully conveyed by asking open-ended questions (Notarius & Markman, 1981).

Finally, the caring component of responsiveness means showing affection and concern for the partner (Reis & Patrick, 1996; Reis & Shaver, 1988). Especially in initial interactions that are often devoid of prior relational information, asking questions is likely to signal care for the partner. Rather than talking about oneself, asking questions about the partner is likely to indicate warmth, positive affect, curiosity, and empathic concern—the question asker shows that he cares to know about the conversation partner's perspective. Expressing affection and care for the partner tends to increase liking by the partner, due to reciprocity effects (Montoya & Insko, 2008; Gouldner, 1960; Wilson & Henzlik, 1986; Sprecher, 1998).

According to Reis and Patrick's (1996) model of responsiveness, understanding is often a necessary requirement of validation and care. That is, one cannot validate and care for someone without first accurately recognizing and acknowledging his or her self-perceptions. In a study

that manipulated understanding and validation orthogonally, liking increased for validating partners when they were accurate rather than inaccurate (Patrick & Reis, 1995; Reis & Patrick, 1996). One needs to first accurately understand their partner's beliefs and attitudes in order to validate them.

The construct of responsiveness aligns closely with the concept of active listening discussed in fields such as communication and marital therapy (e.g., Bodie, 2011; Bodie, St. Cyr, Pence, Rold, & Honeycutt, 2012; Stanley, Bradbury, & Markman, 2000; Gordon, 1975; Weger, Bell, Minei, & Robinson, 2014; Lester, 2002; Rogers, 1951). Like the understanding and validation components of responsiveness, active listening requires paying full attention to the partner in the conversation (Bodie, 2011; Rogers & Farson, 2007; Hutchby, 2005; Rogers, 1955), and most definitions and studies of active listening emphasize the importance of asking questions that are relevant to the partner's statements (Paukert, Stagner, & Hope, 2004; Weger et al., 2014; Bodie, 2011; Minken et al., 1976). A listener's responses regulate the conversation (Duncan & Fiske, 1977; Patterson, 1994), such that responsive verbal behaviors can improve the fluency of the conversation, while unresponsive behaviors can end the conversation (Davis, 1982).

Taken together, we expect perceptions of responsiveness—understanding, validation, and care—to mediate the relationship between question-asking and liking. Asking more questions is likely to increase perceptions of responsiveness, and perceptions of responsiveness, in turn, are likely to increase interpersonal liking. Consistent with this theoretical model, the effect of question-asking on liking may only influence liking of the question-asker by the conversation partner himself (Davis & Perkowitz, 1979). Since we expect the benefits of question-asking on liking to

be explained by responsiveness to the other person in the conversation, we predict that increased

question-asking will not influence liking by third-party observers of the conversation.

Overview of the Present Research

In a series of four studies, we investigate the patterns and effects of question-asking in

dyadic conversation. In Study 1, we instruct one conversation partner in a dyad to ask a high or

low number of questions and measure the other partner's liking of the question asker. In Study

2A, we manipulate high or low question-asking for both conversation partners. In both of these

studies, we investigate responsiveness as a psychological mechanism. In Study 2B, we ask third-

party observers to rate conversation partners on liking. We also conduct a joint analysis of the

types of questions people asked in Studies 1-2. Finally, in Study 3, we investigate the effect of

question-asking in a field context (speed-dating) with a behavioral measure of liking (being

asked on a second date), and we develop a natural language processing algorithm that can classi-

fy question types automatically in any conversation data.

Analytical Strategy

The studies in this paper span a wide range of designs and methods. In general, our

analyses are always designed to test effects at the dyadic level—that is, how person A's level of

question-asking affects person B's evaluations of person A, or how person A's question-asking

affects how person A thinks they will be evaluated by person B. When appropriate, we conduct

these analyses using mixed effects regression models that can account for the nested structures in

some of our data, implemented though the *lme4* package in R (Bates, Maechler, Bolker, & Walk-

er, 2014). This framework allows us to model (and control for) dyad-level characteristics when

we estimate the effect of question-asking on multiple people in a single dyad, and evaluator-level

characteristics when we estimate the effect of question-asking on a single person across multiple dyads.

For each study, we report how we determined our sample size, all data exclusions, all manipulations, and all measures (Simmons, Nelson, & Simonsohn, 2011). In our Online Supplementary Materials on Open Science Framework we provide the data and analysis code from each study.

Study 1

In Study 1, we test the effect of question-asking by randomly assigning one participant in a two-person conversation to ask either a high or a low number of questions. The other conversation partner did not receive or know about the question-asking manipulation. After the conversation, both participants reported how they felt about the conversation and their partner, and how they thought their partner felt about them. To investigate the psychological mechanism, we coded the responsiveness of each conversation partner.

Method

Participants. We recruited 430 participants (215 dyads) to participate in a "Chat Study" in a behavioral lab. We applied several exclusion criteria that were determined *a priori* to ensure our analysis only considered dyads in which both participants completed the full survey. Accordingly, we excluded three dyads in which at least one partner did not finish the study, three dyads in which at least one partner indicated that he or she was not paired with another participant, and 10 dyads in which at least one partner reported that he or she was not able to complete a full conversation. These exclusions left a sample of 398 participants (194 male, 204 female), or 199 dyads, for our analyses.

We recruited participants in three waves due to lab recruiting constraints. In one recruitment wave, participants completed only our study and were paid $15. In two other recruiting waves, participants completed our study among a bundle of unrelated studies. In these latter cases, participants were paid $20 and $25 respectively. We found no differences in our results controlling for recruitment wave and report our results collapsed across all three waves.

Design and procedure. We asked participants to sit in separate cubicles in our behavioral lab. All study materials were presented on computers that were separated by dividers, and participants did not interact face to face before, during, or after the experimental session. Instead, participants interacted by sending each other instant messages using an interface called ChatPlat, an application that enables experimenters to pair people easily and allow them to chat with each other within an online survey. ChatPlat has been used and validated in previous research (e.g., Brooks & Schweitzer, 2011).

Participants were paired with another participant in the room based on their arrival time at the ChatPlat chat window (i.e., the first-arriving participant was matched with second arriver, and so on). Participants were anonymous and unknown to one another. To get the conversation started, they both read these instructions: "You will chat for 15 minutes. During the conversation, your objectives are for you and your partner to get to know each other and learn about each other's interests." Participants were also told to pay attention during the chat because they would be asked to complete several questionnaires about their partner after they finished chatting. After chatting for 15 minutes, the chat window closed automatically. Participants received a notification one minute before the end of the chat.

After participants were paired, each dyad was randomly assigned to one of two condi-

tions: *many-questions* or *few-questions*. In the many-questions condition, one participant in each

dyad was told that he or she needed to ask "at least 9 questions." In the few-questions condition,

the participant was told that he or she needed to ask "at most 4 questions." These question-asking

values were determined based on the natural base rate of question-asking from a separate pilot

study conducted in the same behavioral laboratory ($N = 193$). We used the 25th (4 questions) and

75th percentiles (9 questions) of question-asking from conversations in the pilot study to ensure

the number of questions would be noticeably different from an average conversation ($M = 6.72$,

$SD = 4.16$), but still natural. The participants who received question-asking instructions were

also told not to let their partners know they had been given additional instructions. None of the

participants were told the purpose of these instructions, and they were blind to our hypotheses.

Dependent variables. At the end of the chat, participants in all conditions reported their

liking for their partner, and predicted their partner's liking of them, using the same four-item sur-

vey measure of interpersonal liking (see Appendix A for a full list of measures). In addition to

liking, we also measured learning. We measured participants' knowledge of their partner using

the Activity Preferences Questionnaire (APQ; Surra & Longstreth, 1990; Swann & Gill, 1997), a

nine-item block of Likert responses that ask participants to indicate enjoyment of common activi-

ities (i.e. cooking, sports, reading, etc.). Each participant gave their own answers to the APQ and

predicted how their partner would answer the APQ items (order counterbalanced). At the end of

the survey, we included a manipulation check, asking participants if they were instructed to ask

questions and, if so, how many.

Coding of conversations. We coded the text written by each participant for responsiveness. Coding of all 199 conversations was split among six research assistants, who were blind to condition and hypotheses, such that every conversation was coded by three independent raters. Research assistants read conversations in randomized order, and rated the degree to which they thought that each person in the conversation perceived their partner as responsive, on a 1 (not at all) to 7 (very much so) scale, using the perceived responsiveness scale described by Reis et al. (2011), which captures the three components of responsiveness: understanding, validation, and care.

Results

We compared all of the analyses below across recruitment wave and found no meaningful differences. We present our results collapsed across recruitment wave.

Question-asking. Throughout this paper, we measure question-asking using a simple algorithm that counted conversational turns that included question marks. This method produced virtually identical results when compared to human coders (Cronbach's α: Study 1 = 0.95, Study 2 = 0.97). Using this scheme, if someone asked multiple questions in a single turn (i.e. before their partner responded), this was counted as a single question. However, the results below are identical if we account for multiple-question turns, as well. We used this algorithm to compute the total number of turns in which a question was asked (number of questions asked) as well as proportion of all conversational turns that included a question (question rate).

Consistent with our intended manipulation, participants who were instructed to ask many questions did in fact ask more questions ($M = 10.23$, $SD = 4.94$) than participants who were instructed to ask few questions ($M = 4.34$, $SD = 2.16$, two-sample t-test: $t(296) = 6.31$, $p < .001$).

Participants who received no instructions fell in between ($M = 7.03$, $SD = 3.95$). The same pattern held when questions were measured as a percentage of all conversational turns: those assigned to ask many questions had a higher question rate ($M = 39.1\%$, $SD = 18.9\%$) than those assigned to ask few questions ($M = 21.8\%$, $SD = 14.7\%$; two-sample t-test: $t(197) = 7.16$, $p < .001$, Cohen's $d = .78$).

Liking. The primary dependent measure for this study was a block of four items about how much participants liked their partner after the conversation had ended (see Appendix A). These items were aggregated into a single standardized index of liking (Cronbach's $\alpha = 0.87$), and we plot the averages by condition in Figure 1.

We test our primary hypothesis by estimating the simple effects of the high (vs. low) question-asking instructions on the partner who did not receive instructions. Confirming our prediction, participants paired with high question-askers liked their partners more ($M = 5.79$, $SD = 1.21$) than did participants paired with low question-askers ($M = 5.31$, $SD = 1.48$; $t(197) = 2.47$, $p = .014$; Cohen's $d = 0.35$). There was no difference in liking among those who received the instructions. Those who were instructed to ask many questions liked their partners just as much ($M = 5.76$, $SD = 0.94$) as did participants who were instructed to ask few questions ($M = 5.67$, $SD = 1.27$, two-sample t-test: $t(197) = 0.51$, $p = .612$).

Predicted liking. Predicted liking was reported on the same four items used to measure liking, but participants were asked to anticipate their partner's liking of them (Cronbach's $a = 0.85$, see Appendix A). There was no difference in predicted liking between participants who were instructed to ask many questions ($M = 5.27$, $SD = .93$) or few questions ($M = 5.19$, $SD = 1.05$; $t(197) = .61$, $p = .544$). This was confirmed in a hierarchical linear model that controlled

for person fixed effects. Across conditions, askers' predictions of their partner's liking were not well-calibrated to the actual difference in liking among those partners (interaction term: β = -.386, SE = .223, $t(197)$ = 1.73, p = .084). Across all conditions, there was no correlation between participants' predicted liking and their question-asking rate (r = -0.018, $t(396)$ = 0.35, p = .724). These results suggest that participants did not think that question-asking had an effect on liking.

Learning. The nine APQ items were condensed into a single metric of interpersonal perception accuracy (learning), based on previous research (Shrout & Fleiss, 1979; Swann & Gill, 1997). Though participants assigned to ask many questions were not significantly more accurate (M = 0.327, SD = 0.315) than participants assigned to ask few questions (M = 0.270, SD = 0.317, two-sample t-test: $t(197)$ = 1.26, p = .211), there was a significant correlation between question-asking rate and learning (r = 0.254, $t(196)$ = 3.68, p < .001) among those who did not receive question-asking instructions. It may be the case that the instructions to generate additional questions interfered with participants' ability to retain the information that their partners shared.

Responsiveness. There was high agreement among the coders on ratings of responsiveness (ICC = .75). In line with our hypotheses, participants who were instructed to ask many questions were rated as more responsive to their partner (M = 4.68, SD = 1.08) than participants who were instructed to ask few questions (M = 4.37, SD = .99, t = 2.14, p = .034). There was no difference in rated responsiveness between the un-manipulated participants who were partnered with a high-question-asker (M = 4.55, SD = 0.99) and those who were partnered with a low-question-asker (M = 4.61, SD = .98, $t(197)$ = .38, p = .707).

As a test of our proposed model, we conducted a mediation analysis using a nonparametric bootstrap sampling procedure (Tingley, Yamamoto, Hirose, Kelle, & Imai, 2014). We estimat-

ed the causal pathway linking question-asking with liking for the question asker, as mediated by the responsiveness of the question asker. Across 5,000 simulations, this procedure estimated an effect that was significantly different from zero (standardized effect = 0.121, *95% CI* = [0.007, 0.248], *p* = .034). This suggests that responsiveness plays an important role in explaining why asking someone more questions leads to interpersonal liking.

Discussion

In this study, we tested the effect of asking many or few questions in a conversation. When one conversation partner was instructed to ask many (9) or few (4) questions, their conversation partners liked the high question-askers more than they liked the low question-askers. The participants did not anticipate the effect of question-asking on liking. Furthermore, we found evidence for the mechanism underlying the effect. Participants assigned to ask a high number of questions were perceived as more responsive to their partners, which predicted higher liking by their conversation partners.

Study 2

In Studies 2A-B, we address two limitations in Study 1. First, in Study 1, we manipulated the question-asking of one of the partners in every dyad. In this design, we could not test how the number of questions asked by oneself could moderate the effect of partner question-asking on liking for the partner. Second, we only measured liking by the people who were actively involved in the conversation. This meant that we could not determine whether liking was driven by indirect trait inferences (e.g. question-asking serves as an indicator of likability broadly) or by the direct experience of the conversationalists during the conversation itself.

In Study 2, we address these issues directly. Participants in Study 2A again chatted with each another in dyads, but all participants were assigned to either high or low question-asking conditions in a 2 (self: high v. low question asking) x 2 (partner: high v. low question asking) design. In Study 2B, we recruited a separate sample of observers to read the transcripts of the conversations in Study 2A and rate both participants on the same dependent measures. These observers were able to take an outside view of the conversation, without having to focus on maintaining the dialogue.

Our analyses throughout Study 2 use hierarchical linear modeling to control for dyad fixed effects, as we randomized condition at the individual level (as opposed to the dyadic level as in Study 1). This allowed us to estimate the effect of high (vs. low) question-asking instructions in person A on four hypothesized outcome measures: how much person B likes A, how much A thinks they are liked by B, how much a neutral observer C likes A, and how much C thinks A is liked by B. Finally, we again tested our full theoretical model by analyzing responsiveness as a mediator of the relationship between question-asking and liking.

Study 2A Method

Participants. We recruited participants from Amazon's Mechanical Turk (MTurk) for a "Chat Study." We recruited a total of 446 participants to target a sample that was the same size as the sample size in Study 1 ($N = 430$). From that group, we applied the same *a priori* exclusion criteria as in Study 1 (because the sample was collected online, there were naturally more technical challenges that lead to more exclusions compared to Study 1, which was conducted in a behavioral lab). We excluded 15 dyads where at least one participant did not finish the study. We excluded 28 dyads that contained a duplicate IP address and three dyads that contained a dupli-

cate MTurk ID. Finally, we excluded eight dyads in which participants reported that they were

not able to complete a full conversation. After exclusions, we analyzed data from 338 partici-

pants (177 male, 161 female), or 169 dyads.

Design and procedure. Before starting their conversations, all participants were given

the same instructions as in Study 1 and told their objective was "to get to know each other." The

text of the question-asking instructions was also the same (i.e. "ask [at most 4/at least 9] ques-

tions"). As in Study 1, we measured liking and predicted liking as our main dependent variables

(see Appendix A for all measures collected.)

The most important difference from Study 1 is that *both* participants in every conversa-

tion were given question-asking instructions (compared to just one participant). Each participant

was assigned randomly to ask either many or few questions, and they were assigned to condition

independently from their conversation partner. This ensured that in one-quarter of the pairs both

participants were assigned to ask many questions, in one quarter of the pairs both participants

were assigned to ask few questions, and the remaining pairs included one partner assigned to ask

many questions and one partner assigned to ask few.

After data collection ended, we recruited four independent raters to code responsiveness

for each participant in all 169 conversations in randomized order, using the same measure of re-

sponsiveness used in Study 1 (Reis et al., 2011).

Study 2A Results

Manipulation check: Questions asked. Across all conditions, individual participants

asked 5.98 questions ($SD = 3.63$), on average. Following the instructions, participants who were

told to ask many questions did in fact ask more questions ($M = 8.77$, $SD = 3.15$) than participants

who were told to ask few questions ($M = 3.52$, $SD = 1.78$; HLM: $\beta = 2.68$, $SE = 0.33$, $t(184) =$ 8.05, $p < .001$; Cohen's $d = 1.45$). This difference also held when we computed question-asking as a fraction of total conversational turns: high-question participants had a higher question rate ($M = 46.1\%$, $SD = 19.8\%$) than did low-question participants ($M = 22.5\%$, $SD = 15.6\%$; HLM: $\beta = 12.2\%$, $SE = 2.0\%$, $t(223) = 6.13$, $p < .001$; Cohen's $d = 1.11$). There was no effect of the partner's condition on one's own question rate (all $ps > 0.3$). This confirms that our manipulation had its intended effect on how participants conducted their conversations.

Liking. In these data, both participants in each dyad were subject to a manipulation, so we tested our hypotheses using nested hierarchical linear models. Participants reported their liking for their partner using the same four items as in Study 1 (see Appendix A). These items were aggregated into a single standardized index of liking (Cronbach's $a = 0.92$), and the results by condition are depicted in Figure 2. The results replicate the effect found in Study 1: participants liked high question-askers ($M = 6.02$, $SD = 0.74$) more than low question-askers ($M = 5.79$, $SD = 0.97$; HLM: $\beta = 0.279$, $SE = 0.089$, $t(306) = 3.13$, $p = .001$; Cohen's $d = 0.27$). We also conducted a multiple regression model to test for an interaction between experimental conditions. This analysis revealed that the effect of question-asking was no different between the two conditions (interaction term: $\beta = .108$, $SE = .190$; $t(166) = .49$, $p = .628$), which implies that the effect of question-asking is robust to variations in question-asking from the person being asked (no matter how many questions you asked your partner, the number of questions he/she asked you influenced liking).

Predicted liking. We tested whether participants anticipated the effects of question-asking, using the same standardized index of predicted liking as in Study 1 (Cronbach's $a = 0.92$).

Participants assigned to ask many questions again did not think they would be liked any more (*M* = 5.16, *SD* = 0.90) than participants assigned to ask few questions (*M* = 5.22, *SD* = 1.00; HLM interaction term: *β* = -.289, *SE* = 0.113, *t*(504) = 2.55, *p* = .011). And there was no correlation between question-asking and predicted liking across all conditions (*r* = -0.064, *t*(336) = 1.18, *p* = .238). These results provide further evidence that the positive effect of question-asking on liking is not anticipated by the askers.

Responsiveness. There was once again high agreement among the four coders' ratings of responsiveness (ICC = .75). Replicating the results from Study 1, participants who were instructed to ask many questions were rated as more responsive to their partner (*M* = 4.69, *SD* = 0.67), compared to participants who were instructed to ask few questions (*M* = 4.62, *SD* = 0.69; HLM: *β* = 0.107, *SE* = 0.051, *t*(221) = 2.1, *p* = .035). We conducted another test of our proposed mediation model, and once again found support for our hypotheses. That is, the effect of question-asking instructions on partner liking was significantly mediated by the responsiveness of the question-asker to their partner (standardized effect = 0.072, *95% CI* = [0.004, 0.143], *p* = .041).

Study 2B Method

Participants. We recruited 644 participants from Amazon's Mechanical Turk who participated in exchange for $0.50. As in Studies 1 and 2A, exclusion criteria were determined *a priori*. We excluded 30 participants with duplicate IP addresses and two participants who reported that they could not read the chat conversation. We included 612 participants (373 male, 239 female) in the analysis.

Design and procedure. Participants were randomly assigned to read the transcript of one of the 169 conversations from Study 2A, and were told they would answer some questions about

the conversation partners. Afterwards, participants reported their own liking of both partners and

their prediction of how much each person liked their partner, using the same sets of measures as

in Study 2A. These questions were grouped into two blocks – reported liking and predicted lik-

ing – and the order of these blocks was counterbalanced. Next, participants reported their esti-

mates of how many questions each partner asked. Importantly, all participants in this study were

unaware of the question-asking manipulations of these conversation partners as well as the pur-

pose of the original study and our hypotheses.

Study 2B Results

Each of the 169 conversations was viewed by at least three different independent ob-

servers. We combined the observers' ratings exactly as in Study 2A: that is, every observer's own

liking and predicted liking for both people in the conversation were calculated at the individual

level, as a standardized index across the set of four liking items.

In general, observers who rated the same conversations tended to agree with one another,

with high intra-class correlations for reported liking (ICC = .810) and predicted liking by each

partner toward the other partner (ICC = .876). To test the effect of question-asking on third-party

ratings, the observer ratings were entered into a hierarchical linear model, with their rating as the

dependent variable, and rater and dyad fixed effects included as control variables. This allowed

us to make a precise estimate of how third-party perceptions were influenced by the question-

asking instructions.

Third-party liking. Across the conversations, third-party observers reported a mean lik-

ing of 5.41 ($SD = 1.03$) toward participants. When each person's question-asking condition was

entered as a predictor in the hierarchical model (controlling for dyad and rater fixed effects), the

results showed that people who were assigned to ask more questions were not liked any more than people who were assigned to ask fewer questions (HLM: β = -0.043, SE = 0.037, $t(744)$ = 1.18, p = .240). We also tested this effect in a nested model that compared the third-party raters' liking to the liking reported by the partners in the conversation (and adding target fixed effects) – this, too, showed that third-party liking was not influenced by question-asking conditions (inter-action term: β = -0.180, SE = 0.055, $t(1669)$ = 3.25, p = .001). These results show that question-asking did not induce liking in outside observers in the same way that it affected the people engaged in the conversation.

Though it is not central to our hypotheses, we also tested how *answering* questions affected liking by third-party observers. In our data, outside observers liked people whose partner was instructed to ask many questions more than people whose partner was instructed to ask few questions (HLM: β = 0.075, SE = 0.037, $t(745)$ = 2.05, p = .041). That is, they liked question answerers better than question askers. Again, we tested a nested model that compared the pattern in third-party raters' liking to the liking reported by the partners in the conversation. This analysis showed that third-party liking was more responsive to someone's question-asking than the asker herself (interaction term: β = 0.222, SE = 0.055, $t(1670)$ = 4.02, $p < .001$). This again suggests that question-answering, rather than question-asking, makes a person more appealing to third parties.

Third-party predicted liking. Across the conversations, third-party observers reported a mean predicted liking of 5.59 (SD = 0.95) by one partner for the other partner. When each conversation partner's condition was entered as a predictor in the hierarchical model, the results showed that neither own question-asking (β = 0.009, SE = 0.030, $t(699)$ = 0.29, p = .773), nor

partner question-asking (β = -0.006, SE = 0.029, t(699) = 0.19, p = .847) influenced predicted liking. We once again tested a nested model that compared the pattern in third-party raters' predicted liking to partners' true liking, and confirmed that third-party predictions did not reflect the effect of question-asking on actual partner liking (interaction term: β = -.209, SE = 0.053, t(1670) = 3.95, p < .001). These results show that outside observers, like conversation participants, did not anticipate that people who receive many questions would like their partner more.

Discussion

In Study 2A, we again found that question-asking increased liking in a 15-minute conversation between two people, replicating the results from Study 1. People who asked a high number of questions were liked by their partner more than people who asked a low number of questions, and this was not moderated by the number of questions the recipient had asked during the conversation. This rules out matching or mismatching hypotheses that might suggest that we like people who ask a similar or different number of questions than we ask. We also replicate our Study 1 results with regard to the underlying psychological mechanism: high question-asking leads to increased responsiveness, which explains the effect of question asking on liking.

Furthermore, in Study 2B, third-party observers did not like high question-askers any more than low question-askers. Interestingly, third-party observers liked high question-responders (those who received many questions) more than low question-responders (those who received few questions). In other words, when you are participating in a conversation, you like people who *ask* more questions. When you are observing a conversation, you like people who *answer* more questions. These results suggest that people like question-askers when the ques-

tions are directed toward them personally. This further supports the mechanism of responsive-ness—we like people who seem responsive to us personally (not outside observers).

In Studies 2A-B, we also replicated our finding that question-askers did not anticipate the effect their questions would have on liking. We extended this result to show that it is also true of third-party observers. That is, even neutral observers did not predict that high question-askers would be liked more by their partners than low question-askers.

Classification of Question Types Across Studies 1-2

Studies 1-2 provide converging empirical support for our proposed model: people like their partners more when those partners ask more questions, because they are more responsive. In this section, we analyze not just the *amount* of questions people ask, but also the *types* of questions they ask one another across all data from Studies 1-2. We investigate several question types, but we focus our predictions on one type in particular: follow-up questions.

We define follow-up questions as questions that encourage the partner to elaborate on the content of their prior conversational turn (Davis, 1982). This definition underscores previous conceptualizations of follow-up questions identified in the active listening literature (Paukert et al., 2004; Weger et al., 2014) Follow-up questions are only possible if an individual asks an original question, listens to the answer, and probes for more information (i.e., understands the answer, validates the partner, and cares to know more—the definition of responsiveness). We predict that follow-up questions are an indicator of responsiveness and will correlate positively with liking.

Method

We first developed a classification scheme for the most common question types in our dataset. Based on several small-scale qualitative coding trials, we developed a classification scheme with six question types: follow-up, full-switch, partial-switch, mirror, rhetorical, and introductory (see Table 1 for examples). The first three question types reflected the semantic content of the dialogues, while the latter three types captured common structural elements. These question types align with some prior work on question types that categorize questions based on function and type of information sought. For example, follow-up questions reflect general-inquiry questions that begin with *why* or *how*, which request that the other person provide information, while rhetorical questions function to make a point rather than request information (Graesser, 1985; Schegloff, 2007; Miles, 2013).

Follow-up questions were those that followed up on the topic the partner had mentioned earlier in the conversation (almost always in the previous turn). One needs to listen and understand what the partner said in order to ask a follow-up question. *Full-switch questions* were those that asked about a new topic, one that was unrelated to what the partner had already discussed. We also included a separate category, *partial-switch questions*, for questions that changed topics somewhat, but not entirely.

Mirror questions describe a situation wherein an individual asks the same question that was just asked by their partner. This type of question is always preceded by a question. Mirror questions are distinct from follow-up questions, which are preceded by a statement. *Rhetorical questions* were defined as questions where one does not expect a response from the partner; these speech acts take the grammatical form of a question, but are used to make a point rather than

elicit information. Finally, *introductory questions* were the most superficial, routine questions at the beginning of the conversation.

We trained six research assistants to read through the 368 conversation transcripts from Studies 1 and 2A, such that each transcript would be read by three independent coders. Each person classified every question as one of the six types, and their scores were compiled at the level of every turn. Each turn with a question was assigned a single label based on majority rule among the three coders' labels (ties occurred in <1% of all cases, and were broken in favor of the least prevalent label).

Results and Discussion

Across Studies 1 and 2A, raters had high agreement across most of the question types they were assigned to code. Follow-up questions were the most common, comprising 39.9% of all questions (Cronbach's $\alpha = 0.87$), followed by full-switch questions (28.1%, $\alpha = 0.86$), mirror questions (18.9%, $\alpha = 0.94$), introductory questions (5.8%, $\alpha = 0.93$), partial-switch questions (5.5%, $\alpha = 0.47$), and then rhetorical questions (1.9%, $\alpha = 0.74$). Due to low prevalence and lack of coder consensus, we do not report results for partial switch and rhetorical questions. However, the results for the other question types are unchanged if these question types are included or excluded from the analyses.

We first wondered how the counts of question types mapped onto the constructs in our theoretical model, across all participants in all conditions. We estimated multiple regressions that regressed each of the dependent measures from our model – coder-rated responsiveness and partner-reported liking – onto each participant's set of question counts as independent variables. These results are summarized in Table 2. We find that responsiveness was higher among people

who asked more follow-up questions, and lower among people who asked more full-switch questions. We also find that follow-up questions alone predicted increased partner liking.

We then compared question types across our two question-asking conditions, and found that not only did people in the high question-asking conditions ask more questions, they asked different kinds of questions. Specifically, the proportion of questions that were follow-up questions was higher among people who were instructed to ask many questions ($M = 41.9\%$, $SD = 19.8\%$) than among people who were instructed to ask few questions ($M = 28.8\%$, $SD = 25.7\%$, HLM: $\beta = 0.132$, $SE = 0.020$, $t(526) = 6.67$, $p < .001$). Furthermore, this proportion significantly mediated the effect of condition on partner liking (estimated effect $= .059$, $95\%\ CI = [.011, .113]$, $p = .006$).

These analyses of question types suggest a specific prescription for conversationalists. Asking more questions broadly leads to more interpersonal liking, but follow-up questions are particularly likely to increase liking because they require responsiveness from the question-asker, and signal responsiveness to the question-asker's partner.

Study 3

In Studies 1 and 2, we found that question-asking increases responsiveness and liking, and these effects were driven by increases in follow-up questions, rather than other types of questions. In Study 3, we build on these results by examining the effect of question-asking on a behavioral measure of liking—getting asked on a second date—in a real-world setting of speed-dating.

In this study, we use data from a previous investigation of speed-dating (Jurafsky, Ranganath, & McFarland, 2009; Ranganath, Jurafsky, & McFarland, 2009, 2013). Speed-dating is an

ideal environment to test our hypotheses because: 1) speed daters are highly motivated to make a positive first impression on their conversational partners, and 2) each speed dater interacts with many other partners, which allows us to estimate individual differences in question-asking. Furthermore, the behavioral outcome of getting asked out on a second date allows us to estimate the real effects of question asking and quantify those effects relative to the effects of other variables observable in this setting.

Method

We examine data from a previously published investigation of speed-dating (Jurafsky et al., 2009; Ranganath et al., 2009, 2013). One hundred and ten men and women were gathered in three different sessions. Each dater went on 15-19 four-minute-long speed dates during a session. Every person wore a microphone to capture the dialogue during the dates. After each speed date, participants filled out a brief survey about their most recent partner, and indicated whether they would want a follow-up date. If both people in a pair wanted a second date, the experimenters provided them with one another's contact information (see Jurafsky et al. [2009] for full description of methods).

Our primary dependent variable was participants' willingness to go on a second date, recorded as a binary response (yes/no). Our primary independent variable was the number of questions each person asked on each date. As in previous studies, we measured question-asking by counting conversational turns that included a question mark. However, in this study, we excluded "repair questions" (asking someone to repeat themselves because they were not heard correctly). These did not exist in the text-based conversations in Studies 1-3, but they did consti-

tute 1.2% of all questions in Study 3, according to the detection algorithm reported by Ranganath

and colleagues (2013).

Based on the results of our question type coding in Studies 1-2, we wanted to again test

the role of follow-up questions in this dataset. However, the size of the text corpus ($n = 987$) was

much larger than in Studies 1 and 2 ($n = 368$ combined). So instead of using human coders, we

leveraged the coded data from the earlier studies to build an algorithm that could classify ques-

tion types directly from the text in Study 3.

The coded data from the first two studies contained 4,209 conversational turns that were

classified by human coders to contain one of the six question types described earlier. We pro-

cessed the text of those questions to create a dataset that contained counts of common words and

phrases, as well as some holistic features like turn length and questions in the previous turn (full

details in Appendix B). This was used as training data for a LASSO classification algorithm

(Tibshirani, 1996; Friedman, Hastie & Tibshirani, 2010), which could classify new questions

based on how their text matched examples from the training data. When we conducted out-of-

sample tests using hand-labeled data; the algorithm achieved an 86.9% accuracy rate detecting

follow-up questions.

In this study, we used the algorithm to label the questions extracted from each speed

dater's transcript. The new question turns were processed in the same way as the turns in the

training data, to extract the same set of features. Then, each question was assigned four probabil-

ities, based on the predicted likelihood that it was either a follow-up, switch, introductory, or

mirror question. Our analyses use these probabilities directly, to account for model uncertainty.

However, our analyses are similar if we instead assign a single label to each new question turn.

Results

Our analyses focus on the 1,961 unique observations for which we had a complete tran-

script matched to a post-date survey (after the speed-date). Across these observations, 46.8% of

participants were willing to go on a second date with their partner, although there was also a

large gender gap: men wanted second dates with 56.8% of their partners, while women only

wanted second dates with 36.9% of their partners, on average. All of the dates were filled with

active conversation. Though these speed dates were shorter in duration than the conversations in

Studies 1 and 2, they produced a similar amount of dialogue, presumably because participants

were interacting verbally face-to-face.

Questions were common. On an average date, each person asked 9.8 questions ($SD =$

5.3), which means that roughly 24% of their conversational turns included a question. Our algo-

rithm estimated that the average date included 4.51 follow-up questions from each person ($SD =$

2.91), 3.21 switch questions ($SD = 1.81$), 0.32 introductory questions ($SD = 0.37$), and 1.75 mir-

ror questions ($SD = 1.09$).

Question-asking was a relatively stable trait across individuals. That is, the best predictor

of someone's question rate on any given speed date was their question rate on other dates ($r = .$

535, $t(1959) = 28$, $p < .001$). Likewise, the best predictor of someone's follow-up question rate

was their follow-up question rate on other dates ($r = .556$, $t(1959) = 28$, $p < .001$).

We compared the rate of question-asking to the date "success rate" (as an indicator of lik-

ing). The average question rate per person ($N = 110$) is plotted in Figure 3 against the adjusted

measure of success rate (percentage of second dates per first date, accounting for partner-level

fixed effects). The adjusted rate accounts for the fact that some speed daters are pickier than oth-

ers; in particular, female partners are far pickier than male partners. In the plot, the trend is clear: people who ask more questions get more dates.

We estimated a series of hierarchical logistic regression models to formally test whether the extent of a person's question-asking affects the chances that their partner will want a second date with them. We treat every yes/no second-date decision as an observed outcome nested within dyads ($N = 1961$), and we report several model specifications as robustness checks (see Table 3). We operationalize our independent variable (question-asking) in three different ways: total questions per date (in column 1), total questions per turn (in columns 2, 4, 6), and question types per turn (in columns 3, 5, 7). We also report three specifications of partner effects. There are simple regressions that only control for dyad effects (columns 1-3). There are also regressions that control for gender differences in willingness to date (columns 4-5). And columns that account for all partner-level differences – including gender – using person fixed effects (columns 6-7).

The results are consistent throughout: asking more questions increased date success. This finding was driven by follow-up questions: people who ask a higher proportion of follow-up questions have increased date success, even though other types of question-asking are unrelated to date success. In additional tests, we find no interaction between gender and question-asking. That is, question-asking was similarly effective for both men and women.

Discussion

The results of Study 3 demonstrate a similar main effect as in Studies 1 and 2 in a real-world setting: speed dating. People who asked more follow-up questions were asked on more second dates, a straightforward indicator of interpersonal liking, and this result was not true for

other types of questions. The regression estimates imply that an average person in this sample who asks three more follow-up questions per date over an evening of 20 speed dates should expect to achieve "success" on one additional date. Despite the correlational nature of these data, we believe that they provide further evidence that question-asking can improve impression formation, even in high-stakes, real-world settings.

General Discussion

Conversations are complex social interactions, fraught with decisions about what to say and how to behave. Although most adults have decades of experience conversing with others, our data suggest that people often fail to engage in behaviors that will help them make the most positive impression. While prior data demonstrate that people tend to talk about themselves (Dunbar et al, 1997; Marr & Cable, 2014), our results suggest this may not be an optimal strategy. Instead, across several studies, we find that question-asking reliably increases liking. Furthermore, we identify an important psychological mechanism driving this effect: the effect of question-asking is driven by an increase in responsiveness, which leads question-receivers to like the partner more. Across our studies, we find support for this mechanism using both an attitudinal and a behavioral measure (follow-up questions) of responsiveness.

By asking third-party observers to rate conversations, we tested whether question askers are liked because they are more responsive, or because question-asking is a signal of traits that are inherently likable – for example by signaling interest, empathy, or superior social skills. Instead, we found that high question-askers were not liked any more by third-party observers than were low question-askers. This result dovetails with our findings on follow-up questions, perhaps suggesting a boundary condition. Because a third-party observer is not present in the conversa-

tion by definition, none of the questions being asked can follow up on anything they have said. These results provide converging evidence that people like question-askers because they perceive question-askers as more responsive.

In this research, we find evidence that people do not anticipate the effect of question-asking on liking. This is true when they are in the role of question asker, question receiver, and third-party observer. Thus, it seems that people are largely unaware that asking questions has social benefits. Indeed, third party raters seemed to like question respondents more, perhaps because they were interested in the information that responses provide. This result may explain why low question-asking persists in social interaction: individuals may recall that while observing conversation they liked or found interesting the person who responded to questions.

Lastly, we investigated conversations in the ecologically valid context of speed-dating, where people are motivated to make a good impression during a first encounter. We found that for both men and women, question-asking robustly predicts getting asked on a second date. Using a machine learning analysis of the question contents, we find that follow-up questions, in particular, have a strong relationship with liking in this context, while other types of questions have no relationship at all. Furthermore, we measured question-asking across many interactions for each person, and found evidence that asking follow-up questions is a relatively stable trait: Question-asking is a desirable trait that people seek in potential partners.

Theoretical Implications

Our work makes several fundamental theoretical contributions to an array of existing literatures. First, our work contributes to the understanding of responsiveness within the context of conversations. When person A asks person B more questions, particularly follow-up questions,

person B will like person A more as a result. People want to be heard and validated by others (Davis & Perkowitz, 1979; Laurenceau et al., 1998; Reis & Patrick, 1996; Reis & Shaver, 1988). The question-asker, by expressing interest and engagement, serves a validating role as a valuable conversation partner—indeed, one with whom people want to interact in the future. Prior research has conceptualized responsiveness as understanding, validation, and care (Reis & Patrick, 1996; Reis & Shaver, 1988), and we show that an important behavioral indicator of responsiveness is asking more follow-up questions in a conversation. Follow-up questions appropriately elaborate on the content of the partner's message, and signal that the content is worth continuing to discuss. Responsiveness has previously been conceptualized in close relationships, and we extend this construct to casual encounters and first meetings, by identifying a behavior that can signal responsiveness during conversations in general.

Second, our work contributes to extant research about active listening, which has been previously investigated across fields such as communications, crisis communication, and marital therapy (e.g., Bodie, 2011; Bodie et al., 2012; Stanley, Bradbury, & Markman, 2000; Gordon, 1975; Weger et al., 2014; Lester, 2002; Rogers, 1951), but has been largely overlooked in social psychology. We identify and show evidence that question-asking is a critical component of active listening. Bridging the literatures of communications and social psychology, we suggest that question-asking is an important indicator of both active listening and responsiveness, and we open the pathway for future research to investigate active listening as a social psychological construct.

Third, our finding that people fail to predict the effect of question-asking on liking contributes to previous work on prediction and forecasting errors (e.g., Gilbert & Wilson, 2007). We

suspect that people may show a truncation error specific to conversational experiences: when people simulate conversations, they tend to only imagine what they would say at one time point, rather than the timing and responsiveness of dialogue as the conversation unfolds. This prediction error could additionally help to explain why people tend to talk about themselves during a conversation rather than ask many questions. At any one time point, it is easier to offer statements about the self, since that information is more easily accessible compared to responses that are contingent on the partner's response. Thus, people fail to predict the effect of question-asking on liking when reflecting on conversation and when engaged in a conversation.

Fourth, these findings contribute to a literature about interpersonal interaction and intimacy in longer-term relationships. Although our work only looked at first-encounters, it is likely that the effect of question-asking on responsiveness and liking extends to repeated interactions and longer-term relationships. Previous work has found that social closeness increases as people reciprocally answer questions that grow increasingly more intimate (Aron et al., 1997; Sedikides et al., 1999). But this work has focused on full-switch questions, provided in advance by an experimenter. Our results focus on follow-up questions, and suggests that responsiveness may also be an important factor in a relationship over repeated interactions. The process of interacting with another person affects outcomes such as attraction and intimacy in longer-term relationships (Davis & Perkowitz, 1979; McAllister, 1980; Miller et al., 1983; Reis & Shaver, 1988; Reis, Clark, & Holmes, 2004).

Finally, even though question-asking seems to be a relatively stable trait across individuals, people can learn to ask more questions. In Studies 1-2, we manipulated question-asking, showing that it is remarkably easy to induce people to ask more questions. This suggests that

question-asking is a skill that can be learned. For example, if a person consistently receives positive feedback (such as liking from the other person, or more second dates) for asking more questions during a conversation, she could learn to associate question-asking with positive outcomes, and change her conversational behavior. Over time, given a choice to talk about oneself or ask more questions in a conversation, a person could tend to choose the latter. This is consistent with work showing that active listening skills can be improved with training (Paukert et al., 2004).

Limitations and Future Directions

Our findings are qualified by several limitations that suggest fruitful avenues for future research. For example, in our experiments, we randomly induced the decision to ask questions in conversation. However, in the real world, this decision is not exogenous and may be influenced by many contextual or person-level factors. In our speed-dating study, we find evidence that question-asking is a relatively stable trait across individuals. This result is consistent with prior work showing that being an "opener" in conversation is a stable trait (Miller et al., 1983). Although context was controlled in our experiments, we also found evidence for one contextual factor: time. Specifically, in every one of our studies, people asked more questions early in their conversation, and the question rate drifted downward throughout (though the rate of speech remained constant). However, the number of follow-up questions increased over time, even as other question types became much less common. Of course, time is not the only contextual driver of question-asking, and more research is needed.

Next, the relationship between question-asking and interpersonal liking may not be monotonic—it may be curvilinear. That is, it may be the case that one can ask too many questions, annoying the conversation partner while not revealing sufficient information about oneself.

When one asks too many questions without reciprocating self-disclosure, liking may decrease.

People do not want to talk about themselves indefinitely without reciprocity, since self-disclosure

should induce the other person to self-disclose as well (Cozby, 1972; Dindia, 1988, 2002;

Jourard, 1971; Sprecher et al., 2013). Indeed, prior research suggests that higher turn-taking in

reciprocal self-disclosure increases liking (Sprecher et al., 2013). A person who asks too many

questions would likely be contributing less to the content of the conversation, by not presenting

enough challenge or novelty (e.g., Silva, 2008). In our data, third-party observers preferred ques-

tion recipients, presumably because those seemed like more complete, interesting people. There-

fore, there is likely an optimal balance between question-asking and question-answering. To test

this idea, future work should test question-asking dynamics at the extremes. What happens when

someone asks zero questions? What happens when someone asks fifty?

Broadly, our work opens a new research topic and methodology on the psychology of

conversation, a pervasive human experience. We contribute a new methodology for studying

conversations using software (ChatPlat) that facilitates and captures the text of live conversa-

tions. We conducted a set of large-sample experimental studies involving live dyadic interaction,

which is unprecedented. The use of these methods is important for studying live dyadic interac-

tions between people, since previous research has been limited to small sample sizes of dyads,

hypothetical scenarios, and confederates. Furthermore, we show how basic features of open-end-

ed conversations can be classified automatically, allowing precise measurement of experimental

data without imposing artificial constraints on the participants themselves. We used natural lan-

guage processing and machine learning algorithms to build a "follow-up question detector" (see

Appendix B) that can be applied to any text data to more deeply understand question-asking dynamics.

Our analyses demonstrate a clear effect of the quantity of questions, and especially of asking follow-up questions. Our evidence suggests the importance of question type in conversation, and contributes to a more comprehensive question typology (Graesser, 1985; Miles, 2013; Schegloff, 2007), which may include disclosure triggers that function like questions. Although we have identified that follow-up questions increase liking, there are certainly some types of questions that would not engender liking. For example, rude questions could degrade the quality of a conversation. But rude statements can be equally jarring, so it is not clear whether this is a moderator of question-asking *per se* or a simple effect of semantic content (i.e., rude vs. polite). Similarly, one might expect questions that would embarrass the question-responder to decrease liking. The effects of question-asking on liking are likely to be moderated by the social sensitivity of the questions being asked. We leave ideas about question type as a fruitful avenue for future research.

Furthermore, question-asking is likely to function differently in competitive contexts. People are more likely to be skeptical or defensive when asked questions that solicit information. For example, people may reply honestly or dishonestly depending on the type of question asked (Minson, Ruedy, & Schweitzer, 2011). Indeed, they may require the question-asker to mutually disclose information to feel comfortable disclosing. The reciprocity of disclosure becomes more salient and important in competitive compared to cooperative contexts.

Although there are benefits of question-asking, our findings suggest that people fail to ask enough questions. Why might people forego asking questions in dyadic conversation? First,

given that people tend to be egocentric, they may not think to ask questions of their conversation partners at all because they are too focused on their own thoughts, feelings, and beliefs (e.g., Gilovich et al., 2000). Second, people may realize they can ask questions but have too little interest, care, or curiosity to hear the answers. They may lack social curiosity, defined as "interest in how other people behave, think, and feel" (Renner, 2006, p. 305). Third, people may want to ask questions but perceive social costs to asking questions. For instance, asking for help feels awkward and embarrassing (Bohns & Flynn, 2010; DePaulo & Fisher, 1980). People fear appearing incompetent when asking others for advice (Brooks, Gino, & Schweitzer, 2015) and may hesitate to ask questions because they fear rejection from others (Downey & Feldman, 1996). In the classroom, students could worry that question-asking shows ignorance (Graesser, McMahen, & Johnson, 1994). Indeed, when placed in private tutoring settings compared to the public arena of the classroom, student question-asking increases to ten questions per hour (Graesser, Person, & Huber, 1993). Furthermore, people may worry about making a negative impression by asking the "wrong" questions—those that could be perceived as rude, inappropriate, or intrusive. In sum, people may underweight the benefits of question-asking, overweight the costs, or both.

Conclusion

People spend most of their time during conversations talking about their own viewpoints and tend to self-promote when meeting people for the first time. In contrast, high question-askers —those that probe for information from others—are perceived as more responsive and are better liked. Although most people don't anticipate the effects of question-asking and do not ask enough questions, people would do well to learn that it doesn't hurt to ask.

References

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596-612.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology*, *60*, 241-253.

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, *23*(4), 363-377.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(7).

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*(1), 3-26.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1*(2)*, 99–112.

Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research*, *52*(1), 90-104.

Berscheid, E. (1985). Interpersonal attraction. In G. Lindzey & K Aronson (Eds.), *Handbook of Social Psychology* (3rd ed., pp. 413-484). New York: Random House.

Berscheid, E., & Regan, P. (2005). *The psychology of interpersonal relationships.* Upper Saddle

River, NJ: Prentice Hall.

Bertakis, K. D., Roter, D., & Putnam, S. M. (1991). The relationship of physician medical inter-

view style to patient satisfaction. *The Journal of Family Practice*, *32*(2), 175-181.

Bodie, G. D. (2011). The Active-Empathic Listening Scale (AELS): Conceptualization and evi-

dence of validity within the interpersonal domain. *Communication Quarterly*, *59*, 277–

295.

Bodie, G. D., St. Cyr, K., Pence, M., Rold, M., & Honeycutt, J. (2012). Listening competence in

initial interactions I: Distinguishing between what listening is and what listeners do. *In-

ternational Journal of Listening*, *26*, 1–28.

Bohns, V., & Flynn, F. J. (2010). "Why didn't you just ask?" Underestimating the discomfort of

help-seeking. *Journal of Experimental Social Psychology, 46*(2), 402–409.

Brooks, A. W., & Schweitzer, M. E. (2011). Can Nervous Nelly negotiate? How anxiety causes

negotiators to make low first offers, exit early, and earn less profit. *Organizational Be-

havior and Human Decision Processes*, *115*(1), 43-54.

Brooks, A. W., Gino, F., & Schweitzer, M. E. (2015). Smart people ask for (my) advice: Seeking

advice boosts perceptions of competence. *Management Science*, *61*(6), 1421-1435.

Chafe, W. L. (1970). *Meaning and the structure of language.* Chicago, IL; University of Chicago

Press.

Chen, F. S., Minson, J. A., & Tormala, Z. L. (2010). Tell me more: The effects of expressed in-

terest on receptiveness during dialog. *Journal of Experimental Social Psychology*, *46*(5),

850-853.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*(2)*,* 259-294.

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin*, *79*(2), 73-91.

Davis, D. (1982). Determinants of responsiveness in dyadic interaction. In W. Ickes & E. S. Knowles (Eds.), *Personality, roles, and social behavior* (pp. 85-139). New York: Springer-Veriag.

Davis, D., & Perkowitz, W. T. (1979). Consequences of responsiveness in dyadic interaction: Effects of probability of response and proportion of content-related responses on interpersonal attraction. *Journal of Personality and Social Psychology*, *37*(4), 534-550.

DePaulo, B., & Fisher, J. (1980). The costs of asking for help. *Basic and Applied Social Psychology, 1,* 23–35.

Dillon, J. T. (1982). The multidisciplinary study of questioning. *Journal of Educational Psychology*, *74*(2), 147-161.

Dillon, J. T. (1988). *Questioning and teaching: A manual of practice.* New York: Teachers College Press.

Dindia, K. (1988). A comparison of several statistical tests of reciprocity of self-disclosure. *Communication Research, 15*, 726–752.

Dindia, K. (2002). Self-disclosure research: Knowledge through meta-analyses. In M. Allen, R. W. Preiss, B. M. Gagle, & N. Burrell (Eds.), *Interpersonal communication: Advances through meta-analyses* (pp. 169–186). Mahwah, NJ: Erlbaum.

Downey, G., & Feldman, S. I. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology*, *70*(6), 1327-1343.

Dunbar, R. I., Marriott, A., & Duncan, N. D. (1997). Human conversational behavior. *Human Nature*, *8*(3), 231-246.

Duncan, S. D., Jr., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory.* Hillsdale, NJ: Erlbaum.

Epley, N., & Waytz, A. (2010). Mind perception. In Fiske, S. T., Gilbert, D. T., & Lindzey, G. (Eds.), *Handbook of social psychology* (5th ed.) (pp. 498-541). New York: John Wiley & Sons, Inc.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33(1),* 1-22.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351-1354.

Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of Personality and Social Psychology, 78*(2)*,* 211–222.

Godfrey, D. K., Jones, E. E., & Lord, C. G. (1986). Self-promotion is not ingratiating. *Journal of Personality and Social Psychology*, *50*(1), 106-115.

Gordon, T. (1975). *P.E.T.: Parent effectiveness training*. New York, NY: New American Library.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*, 161–178.

Graesser, A. C. (1985). An introduction to the study of questioning. *The psychology of questions*, 1-14.

Graesser, A. C., McMahen, C. L., & Johnson, B. (1994). Question asking and answering. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics.* San Diego, CA: Academic Press.

Graesser, A. C., Person, N. K., & Huber, J. (1993). Question asking during tutoring and the design of educational software. In M. Rabinowitz, *Cognitive science foundations for instruction* (pp. 149–172). Hillsdale, NJ: Lawrence Erlbaum.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). New York: Academic Press.

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis, 21(3),* 298-313.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, *107*(1), 65-81.

Hutchby, I. (2005). "Active listening": Formulations and the elicitation of feelings-talk in child counseling. *Research on Language and Social Interaction*, *38*, 303–329.

Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: Synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology*, *99*(3), 549-571.

John, L. K., Barasz, K., & Norton, M. I. (2016). Hiding personal information reveals the worst. *Proceedings of the National Academy of Sciences*, *113*(4), 954-959.

Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 231–262). Hillsdale, NJ: Erlbaum.

Jourard, S. M. (1971). *The transparent self* (rev. ed.). New York, NY: Van Nostrand Reinhold.

Jurafsky, D., Ranganath, R., & McFarland, D. (2009, May). Extracting social meaning: Identify-
ing interactional style in spoken conversation. In *Proceedings of Human Language Tech-
nologies: The 2009 Annual Conference of the North American Chapter of the Association
for Computational Linguistics* (pp. 638-646). Association for Computational Linguistics.

Jurafsky, D. & Martin, J. (2009). *Speech and natural language processing: An introduction to
natural language processing, computational linguistics, and speech recognition.* Upper
Saddle River, NJ: Prentice Hall.

Kearsley, G. P. (1976). Questions and question asking in verbal discourse: A cross-disciplinary
review. *Journal of Psycholinguistic Research*, *5*(4), 355-375.

Landis, M. H., & Burtt, H. E. (1924). A Study of Conversations. *Journal of Comparative Psy-
chology*, *4*(1), 81-89.

Laurenceau, J. P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal
process: The importance of self-disclosure, partner disclosure, and perceived partner re-
sponsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*,
74(5), 1238-1251.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-
component model. *Psychological Bulletin*, *107*(1), 34-47.

Lester, D. (2002). Active listening. In D. Lester (Ed.), *Crisis intervention and counseling by tele-
phone* (2nd ed., pp. 92–98). Springfield, IL: Charles C. Thomas.

Marr, J. C., & Cable, D. M. (2014). Do interviewers sell themselves short? The effects of selling
orientation on interviewers' judgments. *Academy of Management Journal*, *57*(3),
624-651.

McAllister, H. A. (1980). Self-disclosure and liking: Effects for senders and receivers. *Journal of Personality*, *48*(4), 409-418.

Miles, E. W. (2013). Developing strategies for asking questions in negotiation. *Negotiation Journal*, *29*(4), 383-412.

Miller, L. C, & Berg, J. (1984). Selectivity and urgency in interpersonal exchange. In V. J. Derlega (Ed.), *Communication, intimacy, and close relationships* (pp. 161-206). Orlando, FL: Academic Press.

Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology*, *44*(6), 1234-1244.

Minkin, N., Braukmann, C. J., Minkin, B. L., Timbers, G. D., Timbers, B. J., Fixsen, D. L., Phillips, E. L., & Wolf, M. M. (1976). The social validation and training of conversational skills. *Journal of Applied Behavior Analysis*, *9*(2), 127-139.

Minson, J. A., Ruedy, N. E., & Schweitzer, M. E. (2011, August). *Ask (the right way) and you shall receive: The effect of question type on information disclosure and deception.* Unpublished paper presented at the Academy of Management meeting, San Antonio, TX.

Montoya, R. M., & Insko, C. A. (2008). Toward a more complete understanding of the reciprocity of liking effect. *European Journal of Social Psychology*, *38*(3), 477-498.

Notarius, C., & Markman, H. (l981). Couples interaction scoring system. In E. Filsinger & R. Lewis (Eds.), *Observing marriage: New behavioral approaches.* Beverly Hills, CA: Sage.

Patrick, B. C., & Reis, H. T. (1995). *Responsiveness to self-disclosure as a mechanism for inti-
macy development: The relative effects of interpersonal understanding and validation.*
Unpublished manuscript, University of Rochester, NY.

Patterson, M. L. (1994). Strategic functions of nonverbal exchange. In J. A. Daly & J. M. Wie-
mann (Eds.), *Strategic interpersonal communication* (pp. 273-29.3). Hillsdale, NJ: Erl-
baum.

Paukert, A., Stagner, B., & Hope, K. (2004). The assessment of active listening skills in helpline
volunteers. *Stress, Trauma, and Crisis*, *7*, 61–76.

Ranganath, R., Jurafsky, D., & McFarland, D. (2009). It's not you, it's me: Detecting flirting and
its misperception in speed-dates. In *Proceedings of the 2009 Conference on Empirical
Methods in Natural Language Processing: Volume 1-Volume 1* (pp. 334-342). Association
for Computational Linguistics.

Ranganath, R., Jurafsky, D., & McFarland, D. A. (2013). Detecting friendly, flirtatious, awk-
ward, and assertive speech in speed-dates. *Computer Speech & Language*, *27*(1), 89-115.

Reis, H. T., Clark, M. S., & Holmes, J. G. (2004). Perceived partner responsiveness as an orga-
nizing construct in the study of intimacy and closeness. In D. Mashek & A. P. Aron
(Eds.), *Handbook of closeness and intimacy* (pp. 201–225). Mahwah, NJ: Erlbaum.

Reis, H. T., Maniaci, M. R., Caprariello, P. A., Eastwick, P. W., & Finkel, E. J. (2011). Familiari-
ty does indeed promote attraction in live interaction. *Journal of Personality and Social
Psychology*, *101*(3), 557-569.

Reis, H. T., & Patrick, B. C. (1996). Attachment and intimacy: Component processes. In E. T.

Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*

(pp. 523-563). New York: Guilford Press.

Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck & D. F. Hay

(Eds.), *Handbook of personal relationships: Theory, research, and interventions* (pp. 367-

389). Chichester, England: Wiley.

Renner, B. (2006). Curiosity about people: The development of a social curiosity measure in

adults. *Journal of Personality Assessment*, *87*(3), 305-316.

Robinson, J. D., & Heritage, J. (2006). Physicians' opening questions and patients' satisfaction.

Patient Education and Counseling, *60*(3), 279-285.

Rogers, C. R. (1951). *Client-centered therapy*. Boston, MA: Houghton-Mifflin.

Rogers, C. R. (1955). *Active listening*. Chicago, IL: University of Chicago, Industrial Relations

Center.

Rogers, C. R., & Farson, R. E. (2007). Active listening. In S. D. Ferguson & S. Ferguson (Eds.),

Organizational communication (2nd ed., pp. 319–334). Piscataway, NJ: Transaction.

Rogers, T. & Norton, M.I. (2011). The artful dodger: Answering the wrong question the right

way. *Journal of Experimental Psychology: Applied*, *17*(2), 139-147.

Rogers, T., Zeckhauser, R., Gino, F., Norton, M., & Schweitzer, M. (2016). Artful paltering: The

risks and rewards of using truthful statements to mislead others. *Journal of Personality

and Social Psychology*. In press.

Rosenfeld, H. M. (1966). Approval-seeking and approval-inducing functions of verbal and non-verbal responses in the dyad. *Journal of Personality and Social Psychology*, *4*(6), 597-605.

Schegloff, E. A. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis* (Vol. 1). Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289-327.

Sedikides, C., Campbell, W. K., Reader, G. D., & Elliot, A. J. (1999). The relationship closeness induction task. *Representative Research in Social Psychology*, *23*, 1-4.

Sezer, O., Gino, F., & Norton, M.I. (2015). Humblebragging: A distinct – and ineffective – self-presentation strategy. Harvard Business School Working Paper No. 15-080.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420-428.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science*, *17*(1), 57-60.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359-1366.

Sprecher, S. (1998). Insiders' perspectives on reasons for attraction to a close other. *Social Psychology Quarterly*, *61*, 287–300.

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology*, *49*(5), 860-866.

Stanley, S. M., Bradbury, T. N., & Markman, H. J. (2000). Structural flaws in the bridge from basic research on marriage to interventions for couples. *Journal of Marriage and the Family*, *62*, 256–264.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological),* 111-147.

Surra, C. A., & Longstreth, M. (1990). Similarity of outcomes, interdependence, and conflict in dating relationships. *Journal of Personality and Social Psychology*, *59*(3), 501-516.

Swann Jr., W. B., & Gill, M. J. (1997). Confidence and accuracy in person perception: Do we know what we think we know about our relationship partners? *Journal of Personality and Social Psychology*, *73*(4), 747-757.

Swann Jr., W. B., Stephenson, B., & Pittman, T. S. (1981). Curiosity and control: On the determinants of the search for social knowledge. *Journal of Personality and Social Psychology, 40*(4)*,* 635–642.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* 267-288.

Tingley, D., Yamamoto, T., Hirose, K., Keele L. & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software, 59*(5)*,* 1-38.

Van der Meij, H. (1987). Assumptions of information-seeking questions. *Questioning Exchange, 1*(2), 111-117.

Vangelisti, A. L., Knapp, M. L., & Daly, J. A. (1990). Conversational narcissism. *Communications Monographs*, *57*(4), 251-274.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model

selection. *BMC bioinformatics, 7(1),* 91.

Weger, H., Jr., Bell, G. R. C., Minei, E. M., & Robinson, M. C. (2014) The relative effectiveness

of active listening in initial interactions, *International Journal of Listening*, *28*(1), 13-31.

Wilson, W., & Henzlik, W. (1986). Reciprocity of liking following face-to-face encounters with

attractive and unattractive others. *Psychological Reports*, *59*, 599–609.

Table 1

Examples of conversational turns containing each question type (from Study 1 data). We show the question-asker's turn in bold font and their partner's previous turn in italics.

Question Type	Example
Follow Up	*I'm planning a trip to Canada.* **Oh, cool. Have you ever been there before?**
Full Switch	*I am working at a dry cleaners.* **What do you like doing for fun?**
Partial Switch	*Not super outdoorsy, but not opposed to a hike or something once in awhile.* **Have you been to the beach much in Boston?**
Mirror	*What did you have for breakfast?* **I had eggs and fruit. How about you?**
Introductory	*hello!* **Hey, how's it going?**
Rhetorical	*What's the craziest event you've been to?* **Yesterday I followed a marching band around. Where were they going? It's a mystery.**

Table 2

Multiple regressions of question type rates from Study 1 and Study 2 on two dependent

measures: the asker's responsiveness to their partner, and how much that partner likes the asker.

	Responsiveness	**Partner Liking**
Follow-Up Question Rate	1.92*** *(.346)*	.887* *(.364)*
Full Switch Question Rate	-2.83*** *(.405)*	-.285 *(.426)*
Mirror Question Rate	.159 *(.468)*	-.026 *(.492)*
Introductory Question Rate	.982 *(1.11)*	-.803 *(1.17)*
Sample Size	**368**	**368**

Legend (two-tailed): ^= $p<0.10$; *= $p<0.05$; ** = $p<0.01$; *** = $p< 0.005$

Table 3

Standardized hierarchical logistic regression models predicting speed daters' willingness to go on a second date in Study 3.

	1	**2**	**3**	**4**	**5**	**6**	**7**
Total Question Count	.097* (.045)						
Total Question Rate		.079^ (.045)		.137*** (.047)		.150* (.061)	
Full Switch Question Rate			-.118 (.086)		.083 (.088)		.126 (.105)
Follow-Up Question Rate			.184* (.086)		.235** (.089)		.287** (.104)
Introductory Question Rate			.041 (.047)		.090^ (.048)		.062 (.055)
Mirror Question Rate			-.001 (.068)		-.060 (.096)		-.049 (.061)
Gender (1=Male)				.850*** (.094)	.888*** (.096)		
Rater Fixed Effects	NO	NO	NO	NO	NO	YES	YES
Dyad Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Dependent Variable	wants 2nd date	wants 2nd date	wants 2nd date	wants 2nd date	wants 2nd date	wants 2nd date	wants 2nd date
Observations	1961	1961	1961	1961	1961	1961	1961
pseudo-R^2	.002	.001	.002	.032	.035	.135	.137

QUESTION-ASKING

Legend (two-tailed): ^= $p<0.10$; *= $p<0.05$; ** = $p<0.01$; *** = $p< 0.005$
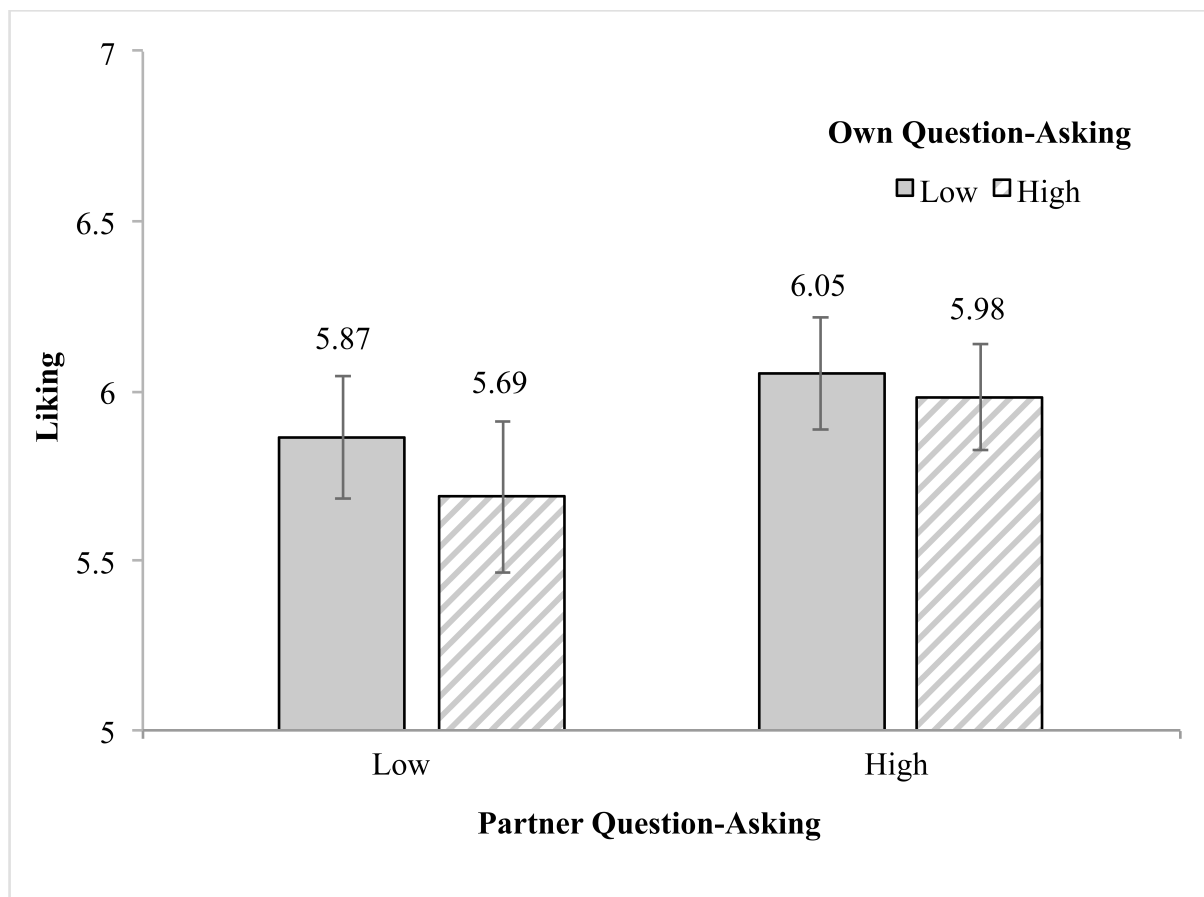
Figure 1

The effect of question-asking on liking in Study 1. In each pair, one person was randomly assigned to receive either few or many questions from the question-asker. Error bars represent 95% CI for the group means.
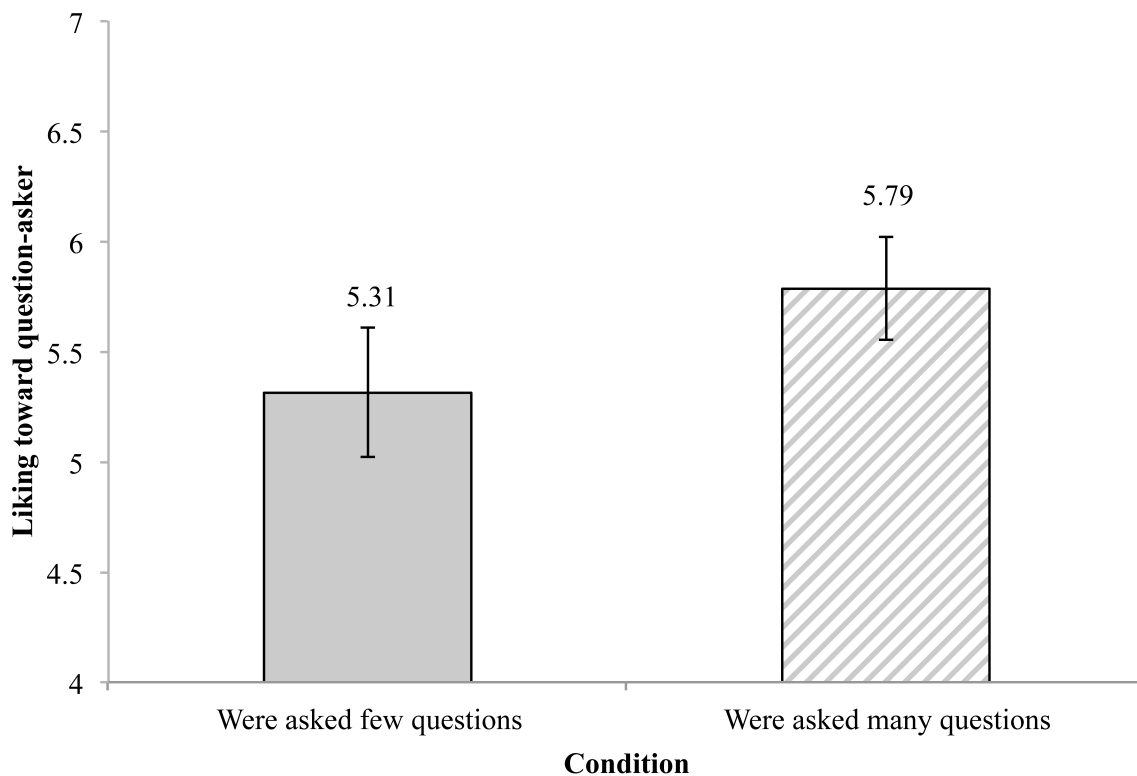
Figure 2

The effect of question-asking on liking in Study 2A. Each participant was given question-asking instructions, such that each person's own instructions and partner's instructions were independently manipulated in a 2x2 design. Error bars represent 95% CI for the group means.

Figure 3

The effect of question-asking on dating outcomes in Study 3. Each point is one speed

dater. The X axis represents the average number of questions they asked per date (as a percentage

of conversational turns) while the Y axis shows each person's speed dating success rate (i.e. get-

ting second-date requests from their partners). The success rate is adjusted for partner fixed ef-

fects, which accounts for the wide gender disparity in the raw success rates. The regression line

represents the HLM fit on the adjusted data (error bands indicate 95% CI of regression slope).
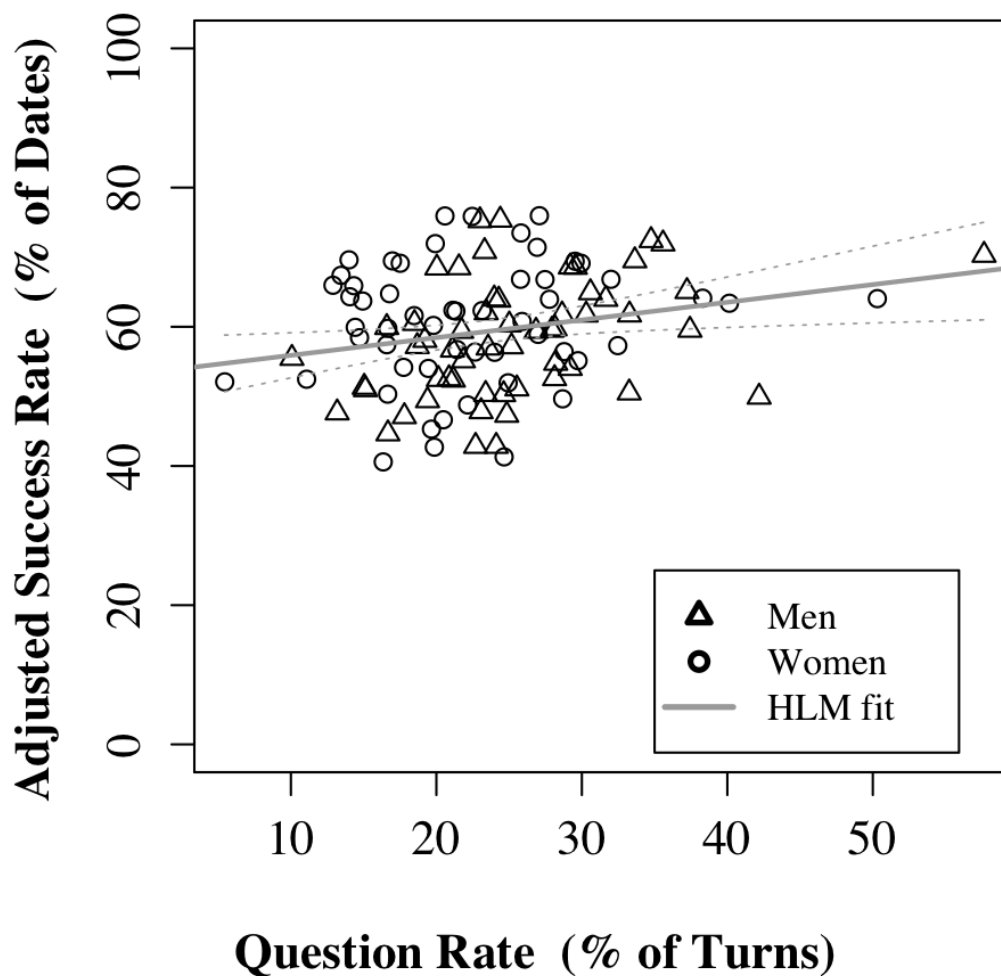
QUESTION-ASKING

<u>Appendix A: Measures from Studies 1-2</u>

Here, we report all measures that were collected during the studies. We discuss the findings from our main measures in the manuscript. Other measures reported below were not connected to the main results, but we report them here for methodological transparency.

<u>Pairing Check (Studies 1, 2A)</u>

Were you paired with another participant? Note: If there were technical issues with the chat, it will not affect your payment.

· Yes, I was joined in the chat with another participant.
· No, no one ever joined me in the chat.

<u>Chat Type (Studies 1, 2A)</u>

If you were paired with a participant, how would you best describe your interaction with the other participant? Note: Your responses will not affect your payment.

· The other participant and I had a full conversation.
· The other participant only said a few lines, then didn't respond.
· I only said a few lines, then didn't respond.
· The other participant only said a few lines, then left the chat.
· The other participant never responded.
· I never responded to the other participant.

<u>Reading Check (Study 2B)</u>

Were you able to read the conversation between User 1 and User 2? (Note: If you were not able to read the conversation, you will still be paid for participating in the study.) (Yes/No)

The following interpersonal perception measures are on 1-7 scales unless noted otherwise.

<u>APQ Self (Study 1)</u>

Please, tell us how much you enjoy doing each of the following activities. That is, give us your own, personal opinion for each question.

1 (Dislike Extremely) – 9 (Like Extremely)

Reading
Watching TV

QUESTION-ASKING

Partying
Seeing live music
Watching sports
Playing board games
Cooking
Working out
Cleaning

APQ Partner (Study 1)

We asked your partner to tell us how much they "enjoy doing each of the following activities."
Now we'd like you to guess how your partner responded. That is, what answer did they just give
to these exact questions?

(Same items as APQ Self)

Liking (Studies 1, 2A)

My partner is likeable.
I liked my partner.
I would enjoy spending time with my partner.
I dislike my partner. *(reverse-scored)*

Predicted Liking (Studies 1, 2A)

My partner thinks I'm likeable.
My partner liked me.
My partner would enjoy spending time with me.
My partner dislikes me. *(reverse-scored)*

Ratings of Liking (Study 2B)

Liking of User 1

User 1 is likeable.
I like User 1.
I would enjoy spending time with User 1.
I dislike User 1.

Liking of User 2

QUESTION-ASKING

User 2 is likeable

I like User 2

I would enjoy spending time with User 2.

I dislike User 2.

Ratings of Predicted Liking (Study 2B)

Prediction of User 1's response

User 1 thinks User 2 is likeable.

User 1 liked User 2.

User 1 would enjoy spending time with User 2.

User 1 dislikes User 2.

Prediction of User 2's response

User 2 thinks User 1 is likeable.

User 2 liked User 1.

User 2 would enjoy spending time with User 1.

User 2 dislikes User 1.

Enjoyment (Study 1)

I enjoyed this conversation.

I thought this conversation was engaging.

I had an interesting conversation with this person.

Predicted Enjoyment (Study 1)

My partner enjoyed this conversation.

My partner thought this conversation was engaging.

My partner had an interesting conversation with me.

Perspective-Taking (Study 2A)

I understand this person's situation

I can imagine being in this person's place

I can easily imagine how things look from this person's perspective

Perceived Perspective-Taking (Study 2A)

This person understands my situation.
This person can imagine being in my place.
This person can easily imagine how things look from my perspective.

Empathic Concern (Study 2A)

I feel warm toward this person
I feel compassion for this person
I feel sympathetic toward this person

Perceived Empathic Concern (Study 2A)

This person feels warm toward me.
This person feels compassion for me.
This person feels sympathetic toward me.

Confidence (Study 1)

How well do you think you know what activities your partner would enjoy?
How well do you think your partner knows what activities you would enjoy?

Self-Other Similarity (Study 1)

How much do you think you have in common with your partner?
How similar do you think you and your partner are likely to be?

Estimated Questions (Studies 1, 2A)

During the conversation with your partner, how many questions did you ask? Please give your best estimate in the box below:

During the conversation with your partner, how many questions did your partner ask? Please give your best estimate in the box below:

Estimated Questions (Study 2B)

In the conversation you just read, how many questions did User 1 ask? Please give your best estimate in the box below:

In the conversation you just read, how many questions did User 2 ask? Please give your best estimate in the box below:

QUESTION-ASKING

Manipulation Check (Studies 1, 2A)

Were you given instructions to ask questions? (Yes/No)
How many questions were you instructed to ask?

Questions Wanted To Ask (Studies 1, 2A)

Did you want to ask more questions or fewer questions than you were instructed to ask? (More/ Fewer/About the same)

During the conversation with your partner, how many questions did you actually want to ask? Please enter a number in the box below:

Demographics (Studies 1, 2A, 2B)

What is your gender? (Male/Female)
What is your age (in years)?

Dictator Game (Study 2A)

You have been granted a $1.00 bonus (in addition to your $2.00 base pay for participating in this study).

You must now decide how much of the $1.00 to give to your partner, and how much to keep for yourself. You get to keep the amount that you do not give to your partner.

You can give some, all, or none of the $1.00 to your partner. Your partner will then keep the amount you give to them. (The payments are given after the survey is finished.)

Your partner will NOT know the amount of money you give until after the survey is finished. You will never see, meet, or interact with your partner in the future.

How much do you decide to give to your partner? Enter a number between $0.00 and $1.00 in the box below:

Appendix B: Follow-Up Question Detection Algorithm

In this section, we report our development of the follow-up question detector. This algorithm serves two purposes. First, we can classify question types automatically, so that we do not need manual coding for large-scale data (as in Study 3). Second, we can learn what features of follow-up questions distinguish them from other questions, providing greater insight than holistic manual codes. These distinctive features inform our academic understanding, and also provide a practical, prescriptive guide for those who want to ask more follow-up questions. We apply this question-type detector to our data from Study 3, but others may apply this detector to any conversational data more broadly. Please contact the authors for more information.

We follow a series of standard text analysis procedures (Jurafsky & Martin, 2009; Grimmer & Stewart, 2013), which are reported here in detail. First, we construct a training set from the manually-coded (human-coded) data in Studies 1 and 2. Second, we pre-process the question text in the training set to produce a high-dimensional feature space. Third, we use a machine learning algorithm to learn the distinctive features of each question type. Fourth, we apply the trained algorithm to out-of-sample data, to generate classifications for new questions. Finally, we estimate the accuracy of this procedure using nested cross-validation within the training set.

Training Set

We approached the detection algorithm as a supervised learning problem. Specifically, the data from Study 1 and Study 2 were taken as ground truth examples of the question types we had asked our coders to classify. Across both datasets, we extracted all conversational turns that included a question (n=4,545), along with the hand-coded label as one of six types: introductory (n=251), mirror (n=865), full-switch (n=1252), follow-up (n=1,841), partial-switch (n=249), or

rhetorical (n=87). The partial-switch and rhetorical types were too rare here to make any reliable

inferences (and were not relevant to our hypotheses), so we decided to drop them.

The resulting training dataset of question turns (n=4,209) was composed of two mea-

sures. The outcome measure was a four-class multinomial label: introductory, mirror, switch, or

follow-up. The prediction measure was the contents of the text in each question turn. The goal of

the algorithm, then, would be to take any new conversational turn that included a question and

assign it one of the four labels from the training data. To accomplish this, we required an auto-

mated method for turning the unstructured text data into structured numeric data.

Text Processing

The text from each question turn was parsed according to the following steps. In order,

the text was converted to lowercase; then contractions were expanded; then punctuation was re-

moved (with the exception of question marks and exclamation points, which were treated as

though they were words). The remaining words were stemmed using the standard Porter stem-

mer, and then grouped into "ngrams" - groups of one, two, or three sequential word stems. For

example, "how are you doing?" would be parsed into twelve stemmed ngrams ["how", "are",

"you", "do", "?", "how are", "are you", "you do", "do ?" "how are you", "are you do", and "you

do ?"]. One exception to the standard natural language processing workflow is that stop words

were not removed, because we wanted to learn the syntactic structure of questions, rather than

extracting semantic content. In fact, 42% of all ngrams in our training set were stop words, in-

cluding all of the question words ("who", "what", "where", "when", "why", "how", and

"which").

The primary feature extraction method followed a "bag of words" approach, which simply counted the ngrams from each question, removing all information about the order in which the ngrams appear. To focus on the most prevalent features, ngrams which appeared in less than 1% of all questions were excluded. This process reduced the documents to a "feature count matrix", in which each question turn was assigned a row, while each ngram feature was assigned a column, and the value of each cell represented the number of times that ngram appeared in that document. This dataset is sparse – specifically, 96% of the cells are zero, since most questions only include a few of the 372 ngrams that were extracted during this process.

In addition to the ngram counts, we calculated six holistic features for each question turn. This included the word count of the entire turn, as well as whether multiple questions were asked in that turn. We also measured the distance of the turn from the beginning of the conversation, as a fraction of the total conversation length. Finally, we included variables indicating whether the question was preceded by a statement in the same turn, whether the question-asker's previous turn included a question, and the question target's previous turn included a question.

Classifier Algorithm

These steps processed the unstructured text into a high-dimensional set of features. So we needed an algorithm that could automatically determine which subset of features were the most useful for classifying question types. We use a common method, the LASSO, implemented in the *glmnet* package (Tibshirani, 1996; Friedman, Hastie, & Tibshirani, 2010). This algorithm estimates a multinomial logistic regression, and regularizes the effective feature space by imposing a constraint on the total absolute size of the coefficients across all features. The size of that constraint is determined empirically, by minimizing out-of-sample error via cross-validation within

QUESTION-ASKING

the training set. This process reduces many coefficients in the regression to exactly zero, leaving

a smaller set with non-zero coefficients in the model.

Although the model is too complex to report in full, we provide two tables for edification. In Table B1 we report the coefficients that map the holistic features to each question type, along with the average for the feature. The results are intuitive – compared to other question types, follow up questions are longer, asked later in the conversation, follow a previous question from the asker, and are somewhat more likely contain multiple questions in the same turn, but are not likely to be preceded by a statement in that turn.

Table B1: Model Coefficients for Holistic Features

Holistic Feature	Average	Follow-up	Switch	Introductory	Mirror
Word count of turn	16.9	0.220	0.000	0.000	0.000
Distance into the conversation	42.9%	0.248	-0.007	-1.399	0.007
Askee's last turn was question	49.4%	0.000	0.000	-0.454	0.619
Asker's last turn was question	53.8%	0.306	0.121	-0.363	-0.121
Multiple questions in turn	10.9%	0.077	-0.044	0.000	0.000
Pre-question statement in turn	36.0%	-0.108	0.000	0.000	0.376

In Table B2, we report distinctive ngram features for each question type. Again, the results are intuitive. Turns with follow-up questions involve appreciation of the previous statement ("nice", "cool", "wow"), or question phrases that prompt elaborations ("why...", "which..", "what kind...", "how do...", "where do..", "is it..."). Full-switch questions are open-ended, and focus on generic facts (e.g. "where [are you from/do you live]?", "what do you like..",

QUESTION-ASKING

"interests", "hobbies?"). Introductory questions (e.g. "hi, how are you today?") and mirror questions (e.g. "yes, i am great, and how about yourself?") are simpler categories, that follow conventional and distinctive scripts. While the full model contains far more complexity than this table alone can provide, it is reassuring to know that the features selected by the model align with basic intuition.

Table B2: Distinctive features for each question type

Follow-up	Full-switch	Introductory	Mirror
which	how old	how are you	how about
why	do you like	hello	what about
what kind	travel	your name	yourself?
cool	fun	how are	and
nice	do you live	hi how	i am
wow	interests	today?	and you
is it	hobbies?	what is	what about you
how do	you a student	go?	and
where do	weather	name?	no, i
want to	you from?	are you?	yes, i

Predictions

Once the model was built, any new test set of questions could be labelled according to this four-class scheme. To do so, the text in the test set would simply have to be processed in the same way as the text in the training set, tallying the presence of the same 372 ngrams in the new text (along with the six holistic features). Note that the prevalence filter was still applied with respect to the training set, not the test set – that is, only the 372 ngrams that were found in at

QUESTION-ASKING

least 1% of the training set were counted, no matter how rare or common they were in the test set.

For every new conversational turn, the algorithm provided probabilistic multinomial likelihoods across the four classes. As an example, the turn "I see. Why did you have to move?" was given the following likelihoods: introductory=0%; mirror=4%, switch=20%, follow-up=76%. For our analyses, we decided to reduce the output by adding up the mirror, introductory, and full switch likelihoods to an "other-type" likelihood, while keeping the follow-up likelihood as-is. This simplified the output of our model, and also more directly addressed our hypotheses, which concerned follow-up questions specifically, but made no predictions about other types of questions. We confirmed that the results of our analyses were identical across a wide range of other analytical choices - either by using the full set of four-class labels, or by translating the probabilistic model output into deterministic labels for each question.

Accuracy Validation

Before applying the algorithm to new unlabeled data, we wanted to estimate the algorithm's accuracy on the labeled data in the training set. We did this using a nested cross-validation procedure (Stone, 1974; Varma & Simon, 2006). The entire dataset was randomly divided into ten folds of equal size. To produce out-of-sample predictions for each fold, a classification model was trained and tuned on the other nine folds, and applied directly to the held-out data to estimate the likelihood that each held-out question was a follow-up question. To smooth out the random fluctuation, we performed this procedure five times, and averaged across all five likelihoods, to determine a final estimated likelihood for each question.

We computed accuracy using the area under the curve metric (AUC). This tests calculates the probability that any follow-up question will be assigned a higher follow-up likelihood than any non-follow-up question (with 50% as a baseline of random guessing). Across the entire training set, this validation exercise produce an accuracy level of 86.9% [95% CI: 85.9%, 88.0%]. We considered this a very high level of accuracy, considering that the true test of our model would be in a different conversational context (spoken vs. written).