WILEY

Perceiving commitments: When we both know that you are counting on me

Francesca Bonalumi[1] | **John Michael[1,2]** | **Christophe Heintz[1]**

[1]Department of Cognitive Science, Central European University, Budapest, Hungary

[2]Division of Psychology, University of Stirling & Department of Cognitive Science, Central European University, Budapest, Hungary

Correspondence
Francesca Bonalumi, Department of Cognitive Science, Central European University, Nador u. 9, 1051 Budapest, Hungary.
Email: bonalumi_francesca@phd.ceu.edu

Can commitments be generated without promises or gestures conventionally interpreted as such? We hypothesized that people believe that commitments are in place when one agent has led a recipient to rely on her to do something, even without a commissive speech act or any action conventionalized as such, and this is mutual knowledge. To probe this, we presented participants with online vignettes describing everyday situations in which a recipient's expectations were frustrated by one's behavior. Our results show that moral judgments differed significantly according to whether the recipient's reliance was mutually known, irrespective of whether this was verbally acknowledged.

KEYWORDS
commitment, expectations, moral judgment, mutual knowledge, reliance

1 | INTRODUCTION

Imagine that you and your friend Kate are planning to meet at the gym to work out together at 6 p.m. At 5.30 p.m. you discover that some other friends are meeting at the very same time for drinks, and you would prefer to join them, but you also feel you cannot let your friend Kate down. Indeed, she expects to meet you there. She is counting on you. We are often confronted with such choices in everyday life, and our decisions typically involve the feeling that we are committed. We also often find ourselves in situations like that in which Kate finds herself: expecting, counting on, or relying on someone to do something. Commitments are important in

a wide variety of social and nonsocial contexts: We are committed to our partners, our social groups, our jobs, our individual and our shared goals, our values, and even ourselves. Although there are likely to be many similarities across these situations, the current set of studies is restricted to instances of interpersonal commitment—that is, to those commitments that are made by one individual to another individual (cf., Clark, 2006).

In the philosophical literature, commitment is usually treated as a relation among one committed agent, one agent to whom the commitment has been made, and an action which the committed agent is obligated to perform in virtue of having given her assurance to the second agent that she would do so (Michael, Sebanz & Knoblich, 2016a; cf., Searle, 1969; Scanlon, 1998). Moreover, commitment is treated in this literature as a binary notion: Either the aforementioned conditions have been fulfilled (and there is a commitment) or they have not (and there is no commitment). More recently, in the psychological literature, Michael et al. (2016a) have proposed to treat commitment as a graded phenomenon: One agent can be more or less motivated to perform an action that a second agent is relying on, and may feel more or less guilty if she does not perform the action. To capture this, they introduce the notion of a "sense of commitment," which admits of degrees. In the current paper, we adopt this non-binary conception, as we are interested in people's psychological attitudes about commitment rather than in commitment in the normative sense.

We present empirical results that show what it takes for people to perceive that a commitment has been made. We thus investigate the social conditions that lead people facing standard situations to perceive that a commitment has been made. The act of promising is the canonical way to generate a commitment, and philosophers have analyzed the conditions under which a promise is performed and possesses a normative power that commits a speaker to a certain course of action. Speech act theorists claim that this normative power arises when the speaker performs a commissive speech act, that is, a speech act that indicates the speaker's intention to incur a moral obligation to perform (or omit) a particular action, or that a convention dictates that the given speech act has been performed in such a way and under such circumstances that such obligations have arisen; for instance, stating "I will do it" or nodding after a request are the kind of speech acts (verbal or not) that in the right circumstances are conventionally interpreted as promises (Austin, 1962; Searle, 1969). This raises the question—which is our focus here—what the *right circumstances* are under which people perceive there to be a commitment even in the absence of a commissive speech act.

Several philosophers have pointed out the role of common or mutual knowledge in making a commitment. For instance, Gilbert (1990, 2006) provides examples where commitments arise from common knowledge of joint goals in the absence of commissive speech acts. MacCormick and Raz (1972) and Scanlon (1998), however, argue that commitments can be formed with neither conventional norms (as when performing a commissive speech act) nor shared goals. What matters, they say, is that one agent leads another agent to form expectations about her future behavior and to rely on this behavior. In the current set of studies, we test whether people perceive that a commitment is in place when reliance is mutually known. Our findings indicate that the accounts offered by MacCormick and Raz, as well as by Scanlon, nicely reflect people's judgments when asked to evaluate ecologically valid scenarios.

There is much debate around the notion of "common knowledge." Schelling (1980) and Lewis (1969) point out that coordination games can be solved by assumptions of recursive common knowledge between agents, and Schiffer (1972) defines common knowledge as a hierarchy of propositions that pose strong inferential demands (I know that you know that I know that you know, etc.). However, many acknowledge that agents cannot entertain infinite recursive

epistemic states, and several deflationary accounts provide more plausible psychological implementations, such as the availability of the given information in the common ground (Lewis, 1978; Sperber & Wilson, 1986/1995, Chapter 8; Vanderschraaf & Sillari, 2014; Carpenter & Liebal, 2011). Following these cognitively realistic accounts, we understand mutual knowledge in the minimal sense of availability of the information in the common ground, and not as recursive higher-order knowledge. In our experiments, we describe scenarios in which some degree of mutual knowledge of one's reliance is present, that is, in which the agents at least know that one agent will rely on the other agent's behavior.

Building upon MacCormick and Raz, Scanlon, as well as on Michael et al. (2016a) and Michael, Sebanz and Knoblich (2016b), we hypothesize that people have a sense that an agent—the "sender"—is committed to performing X (to believe that the sender is committed, to attribute blame and to experience negative emotions if the sender does not perform X), if the following conditions are met: (a) the sender has led a second agent (the recipient) to rely on her to do something; and (b) this is mutually known by the two agents. We operationalize the notion of reliance as the recipient changing her course of action based on her expectations of X occurring.[1] The phenomenon of reliance is often expressed by the recipient with idiomatic expressions such as "I am counting on you," which make explicit the fact that certain expectations are in place, and that the recipient will act accordingly.

In our studies, we consider instances in which it is mutually known that one action of an agent (the sender) has led a second agent (the recipient) to expect her to perform an action X, independently of whether the sender has verbally acknowledged those expectations. In order to investigate whether a recipient's reliance (when it is mutually known) is one factor determining whether a sense of commitment arises, we thus implemented four studies in which we presented participants with scenarios where a sender fails to do X, and manipulated mutual knowledge and the means by which the recipient's expectations were raised: either via an explicit speech act, or through nonverbal events. On the basis of our hypothesis, we predicted that participants' attitudes about whether a commitment has been violated (and about the extent to which the commitment violation warrants blame and more reputational consequences) would depend on whether the recipient's reliance was mutually known, whereas the means by which mutual knowledge has been created would not significantly impact participants' evaluations.

2 | STUDY 1

The first study we conducted was designed to test the hypothesis that mutual knowledge of reliance is a sufficient condition for triggering commitment. To this end, we presented participants with vignettes describing everyday situations in which a sender failed to fulfill the expectations of the recipient. We measured the perception of a commitment being in place by prompting a normative judgment about the sender's behavior (normative question), by asking whether the situation triggered a feeling of annoyance (affective question), and by probing to what extent the participant herself would be willing to interact with the sender in the future (partner choice question).

---

[1]It is worth noting that, contrary to some accounts of reliance that do not involve expectations (see, for example, Alonso, 2016), the notion of reliance with which we are working requires the recipient's expectations to be in place.

2.1 | Methods

2.1.1 | Participants

We implemented a between-subjects design on an online platform (SurveyMonkey). Since previous online studies conducted in our lab indicated that nonpaid participants present high rates of incomplete and invalid surveys, we opted for a large sample size. A power analysis using G*Power 3.1 (Faul, Erdfelder, Lang & Buchner, 2007) indicated that a total sample size of 308 participants would be needed to detect a medium effect size ($f = 0.25$) with a predicted statistical power of 98% using a one-way ANOVA with alpha at .05. Since we planned to run nonparametric tests, we added 15% to our desired sample (Lehmann, 2006). We anticipated that about 25% of participants would not complete the experiment and answer the control questions correctly. Participants were 536 adults, recruited via social media, e-mail, and word of mouth. Data was discarded from participants who did not complete the survey ($N = 118$) or failed one or more control questions ($N = 49$), and also from participants who reported being younger than 18 years old ($N = 6$). This left a total sample size of 364 participants (173 females; $M_{age} = 25.80$ years, $SD = 6.95$)—129 in the *no mutual knowledge* condition, 128 in the *implicit mutual knowledge* condition, and 107 in the *explicit mutual knowledge* condition. The sample was composed of 53.6% North Americans, 29.2% Europeans, and 17.2% participants from other regions.

Here and elsewhere, the methods used were in accordance with the international ethical requirements of psychological research and approved by the EPKEB (United Ethical Review Committee for Research in Psychology) in Hungary. All participants gave their informed consent by ticking a box prior to the experiment.

2.1.2 | Design and procedure

Participants were asked to read different hypothetical situations in which a sender violates a recipient's expectations. They were presented with one scenario, in which the agents' expectations either are or are not mutually known, and in which the sender acknowledges these expectations either verbally or only implicitly.

Participants were randomly assigned to one of three conditions: *explicit mutual knowledge*, *implicit mutual knowledge*, or *no mutual knowledge*. In the *implicit mutual knowledge* condition, the scenario reads as follows:

> Beth and Ashley are two friends who are planning to go to the seaside for the weekend. Ashley insists on leaving as early as possible because she would like to reach the beach before noon and have lunch there. She offers to pick Beth up at 7 a.m. Beth would rather leave at 9 a.m. and have lunch on the way because she hates waking up early. Each of them keeps insisting on her own preference, and they wind up getting mad at each other. The conversation on Friday night ends with Beth telling Ashley "I will wait for you at 9 a.m.", and Ashley telling Beth "I will pick you up at 7 a.m.!".
> The same evening Beth goes out to a pub with another friend, who tells her about a nice bistro on the seaside. She then realizes that it could be nice to leave at 7 a.m. after all, and reach the seaside in order to have lunch at this bistro. She

sends a message to inform Ashley that she wants to leave early, as Ashley had suggested, and that she will therefore be waiting for Ashley at 7 a.m. When Beth checks her messaging app, she can see that Ashley read the message a couple of minutes after she (Beth) sent it. In the morning, Beth wakes up early and is ready to go at 7 a.m. As it happens, when Ashley's alarm rings, she decides to turn it off and sleep a bit longer. Ashley arrives at Beth's place at 9 a.m.

In the *explicit mutual knowledge* condition, the vignette differed insofar as Ashley replied with a message saying that she would come at 7 a.m., and in the *no mutual knowledge* condition the vignette differed in that Ashley did not receive the message.[2] After reading one of the vignettes, participants were asked to respond to questions about the moral and cooperative character of the agent who changed her course of action (the sender). We hold the sense of commitment to be on a continuous scale rather than a yes/no phenomenon, so we opted for the use of scales as opposed to binary questions. The questions were the following:

Control Question 1: "At what time did Ashley tell Beth that she would pick Beth up?" ("At 7 a.m."; "at 9 a.m."; "at 11 a.m.").

Control Question 2: "At what time did Beth want Ashley to pick her up before she (Beth) learned about the bistro?" ("At 7 a.m."; "at 9 a.m."; "at 11 a.m.").

Normative Question: "How wrongly do you think Ashley behaved?" ("Very wrongly"; "A bit wrongly"; "Not particularly wrongly"; "Not at all wrongly").

Partner Choice Question: "If you imagine yourself in Beth's situation, would you feel like going on another trip with Ashley in a couple of weeks?" ("Very much"; "A bit"; "Not particularly"; "Not at all").

Affective Question: "If you imagine yourself in Beth's situation, would you feel frustrated/upset/angry towards Ashley?" ("Very much"; "A bit"; "Not particularly"; "Not at all").

Control Question 3: "On the basis of the information that you have, which of the following statements is the most accurate?" ("Ashley did not receive Beth's message about leaving earlier"; "Ashley responded to Beth's message about leaving earlier"; "According to Beth's messaging application, she read the message but did not respond").

The normative question was designed to trigger participants' explicit normative judgments about the sender.[3] We predicted that they would evaluate the sender to having misbehaved more often in the two *mutual knowledge* conditions than in the *no mutual knowledge* condition, as participants' judgments as to whether a commitment has been violated would depend on whether the recipient's expectations were mutually known. We further predicted that the *explicit* versus *implicit mutual knowledge* conditions would lead to no significant difference in the answers to the questions, as the means by which mutual knowledge was created should be irrelevant for such judgments.

The purpose of the affective and partner choice questions was to control for any mismatch between normative criteria for commitment and a subtler feeling of commitment or emotional

---

[2]See https://osf.io/h3mjp/ for the full vignettes.

[3]According to our hypothesis, Ashley is here the sender and Beth the receiver (despite the fact that in the story, Beth is the person who sent the message to Ashley).

disappointment that is not affected by such considerations, as reported by Michael et al. (2016b). The affective question was designed to tap participants' emotional reactions to the violation described. We predicted that they would indicate a higher level of frustration in the two *mutual knowledge* conditions than in the *no mutual knowledge* condition, with the additional prediction that there would be no significant difference between the *explicit* and *implicit mutual knowledge* conditions. We reasoned that the violation of a commitment would lead to a negative emotional reaction, and thus the same factors influencing a normative evaluation of the agent's deed would impact participants' levels of frustration.

The partner choice question was designed to probe whether people might engage in a partner choice strategy following the violation of a commitment. We predicted that they would more likely indicate a lower willingness to interact with the sender in the future in the two *mutual knowledge* conditions than in the *no mutual knowledge* condition, with the additional prediction that there would be no significant difference between the *explicit* and the *implicit mutual knowledge* conditions. We reasoned that participants would rather avoid interacting with commitment violators, and that the same factors influencing a normative evaluation of the agent's deed would therefore impact participants' partner choices.

The control questions were designed to check whether the participant had read the story with sufficient care to register the information required in order to answer the target questions. Control Question 3 was particularly important insofar as it was devised to probe whether participants had understood the critical manipulation. The control and the target questions were presented to the participants in a randomized order, except for the third control question, which was always presented last, since being forced to make a judgment about the epistemic states of the agents could have an effect on the other judgments. Data from those who failed to answer any of the control questions correctly was discarded from the final sample.

2.2 | Results

To test these hypotheses, we ran a series of Kruskal–Wallis nonparametric tests, and a series of post hoc tests. Given that our measures involve ordinal scales, we opted for using appropriate nonparametric rather than metric tests (Liddell & Kruschke, 2018). Here and elsewhere, the analyses were performed using IBM SPSS v.25. In accordance with our predictions, a Kruskal–Wallis H test showed that there was a statistically significant difference in the responses to the normative question, $\chi^2(2) = 108$, $p < .001$, $\eta^2 = 0.29$ (large effect size), with a mean rank rate of 110.43 for the *no mutual knowledge* condition, a mean rank rate of 212.52 for the *implicit mutual knowledge* condition, and a mean rank rate of 233.48 for the *explicit mutual knowledge* condition. In order to determine which condition(s) were responsible for this difference, we ran a series of post hoc pairwise comparison tests showing that responses were significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$) and in the *implicit mutual knowledge* condition ($p < .001$). However, no significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .320$) (see Figure 1). Here and elsewhere, significance values have been adjusted by the Bonferroni correction. This confirms the hypothesis that the levels of perceived commitment were higher in conditions in which the expectations were mutually known by the agents.

In accordance with our predictions, the responses to the affective question showed the same pattern as for the normative question: A Kruskal–Wallis test revealed that the responses were significantly different in the three conditions, $\chi^2(2) = 83.5$, $p < .001$, $\eta^2 = 0.26$ (large effect size), with
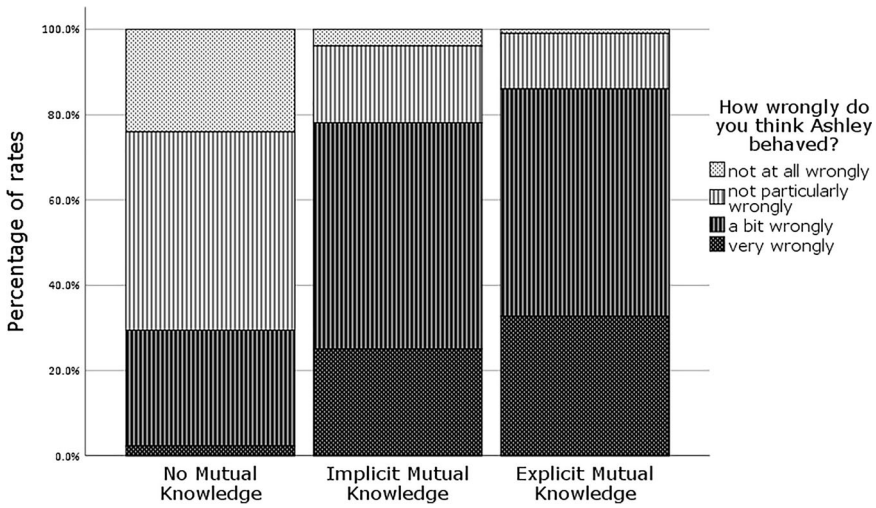
FIGURE 1 Level of perceived commitment--normative question. The responses to the normative question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 363$, $\chi^2(2) = 108$, $p < .001$, $\eta^2 = 0.29$

a mean rank rate of 119.38 for the *no mutual knowledge* condition, a mean rank rate of 207.80 for the *implicit mutual knowledge* condition, and a mean rank rate of 226.75 for the *explicit mutual knowledge* condition. Again, a series of post hoc pairwise comparison tests showed that responses are significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$) and in the *implicit mutual knowledge* condition ($p < .001$). No significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .773$) (see Figure 2). The responses to the affective question predictably correlated with the responses to the normative question, $r_s(364) = .53$, $p < .001$.
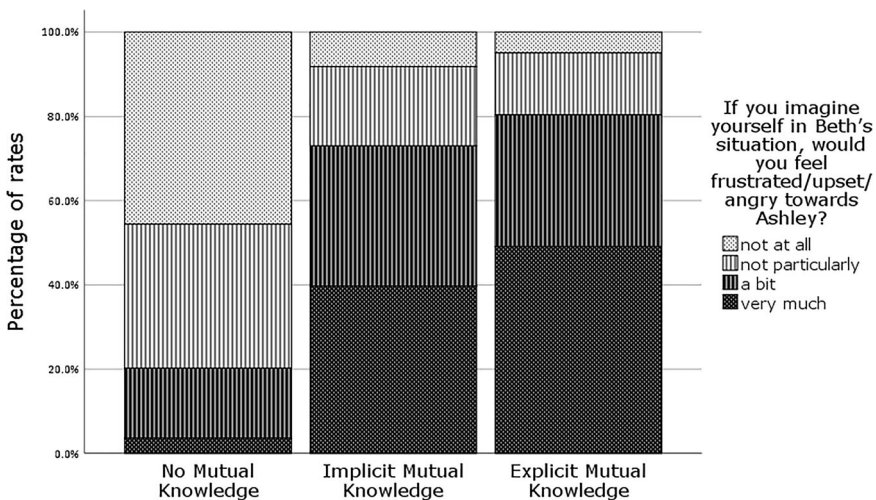

FIGURE 2 Level of perceived commitment--affective question. The responses to the affective question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 363$, $\chi^2(2) = 83.5$, $p < .001$, $\eta^2 = 0.26$
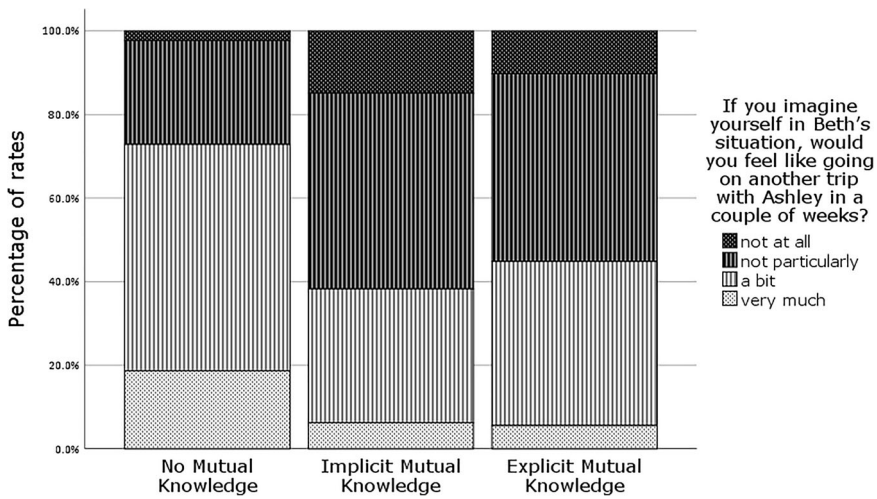
FIGURE 3 Level of perceived commitment--partner choice question. The responses to the partner choice question are significantly higher in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 363$, $\chi^2(2) = 40.4$, $p < .001$, $\eta^2 = 0.11$

The pattern presented above is confirmed for the partner choice question: The responses, tapping participant's willingness to interact again with the sender, were significantly different in the three conditions, Kruskal–Wallis Test, $\chi^2(2) = 40.4$, $p < .001$, $\eta^2 = 0.11$ (medium effect size), with a mean rank rate of 226.01 for the *no mutual knowledge* condition, a mean rank rate of 152.71 for the *implicit mutual knowledge* condition, and a mean rank rate of 165.68 for the *explicit mutual knowledge* condition. To check that the critical difference lay between the *no mutual knowledge* condition and the others, we ran a series of post hoc tests that showed that the responses are significantly higher in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$) and in the *implicit mutual knowledge* condition ($p < .001$). However, no significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .936$; see Figure 3). These results rule out the possibility that participants, while responding to the normative question, were already engaging in some partner choice strategy or implicit disapproval without genuinely evaluating their partner's behavior as morally wrong.

The responses to the partner choice question correlated both with the responses to the normative question, $r_s = .422$, $p < .001$, and with the responses to the affective question, $r_s = .459$, $p < .001$.

2.3 | Discussion

The results corroborated our predictions. Participants evaluated the sender more severely in cases in which the sender had led the recipient to rely on her (and this reliance was mutually known), irrespective of how mutual knowledge had been formed (i.e., whether the sender performed a speech act or not).

Participants' willingness to engage in an unspecified future interaction with the sender was influenced by this factor, but not as strongly as their affective response or their normative

evaluation of the sender. There are several possible explanations of this. People take into account several types of information when reasoning about whether one is a desirable partner, information that spans from her competence in a relevant domain (e.g., whether Thomas a good tennis player, if I have to team for a tennis tournament) to her benevolence and willingness to cooperate (e.g., whether Thomas is moved by benevolent intentions) (Fiske, Cuddy & Glick, 2007; Heintz, Karabegovic & Molnar, 2016). Violating previous commitments is surely among the latter considerations, but it is reasonable to assume that in the scenario there were other implicit commitments in place between the two agents in addition to the one that was violated, commitments that maybe weight more, as the ones entailed by being friends, and an assumption of reliability due to the (inferred) history of the friendship. When asked about future potential interactions, participants might have taken these factors into account. Furthermore, after responding to the normative question, participants might have been satisfied with having attributed blame to the violator, and therefore considered that an additional precaution would be redundant.

Given that vignettes may be open to a broad range of interpretations, and in light of the inherent noisiness of online data collection, we designed Study 2 to replicate the results of Study 1 using different vignettes.

3 | STUDY 2

Study 2 was designed to implement two different scenarios. Before analyzing the data, we ran a preliminary test to check whether the different scenario presented to the participants influenced their responses. An independent-measure Mann–Whitney U test revealed that the responses to the normative question were significantly different between the two scenarios, Mann–Whitney: $N = 204$, $U = 2,846.500$, $p < .001$. The responses to the affective question were also found to be significantly different, Mann–Whitney: $N = 204$, $U = 3,539.000$, $p < .001$, as well as the responses to the partner choice question, Mann–Whitney: $N = 204$, $U = 3,424.000$, $p < .001$. These results persuaded us to run additional tests separately and to consider the two scenarios as different studies. We therefore considered the data from the one scenario as Study 2, and the data from the other scenario as Study 3.

Compared with Study 1, in Study 2 we modified an element that might plausibly be relevant to participants' interpretation of the situation, namely the nature of the relationship between the two agents—that is, in Study 1 the two agents were friends, whereas in Study 2 they were colleagues. We implemented mutual knowledge in a similar fashion, that is, via an automatic in-built function of a communication device. This limits the plausible deniability for the sender of not having been exposed to the relevant information.

3.1 | Methods

3.1.1 | Participants

We used Amazon M-Turk to implement a web-based paradigm with a between-subjects design. In anticipation of an effect size similar to what was observed in Study 1, a power analysis using G*Power 3.1 indicated that a total sample size of 231 participants would be needed to detect the expected effect size ($f = 0.22$) (derived from a predicted statistical power of 85% using a one-way

ANOVA with alpha at .05). We added 15% to our desired sample, thus we aimed for a sample size of 265 participants. In total, 265 adults completed the experiment, each of whom was rewarded with $ 0.45. Data was discarded from participants who did not complete the survey ($N = 11$) or failed one or more control questions ($N = 48$), and technical errors ($N = 2$) leaving a total of 204 participants in the final data set. Hundred and twenty-three participants were assigned to Study 2 (76 females; $M_{age} = 40.67$ years, $SD = 12.91$), 52 in the *no mutual knowledge* condition, 40 in the *implicit mutual knowledge* condition, and 31 in the *explicit mutual knowledge* condition. As participants were recruited via Amazon M-Turk, the sample was composed entirely by North Americans.

3.1.2 | Design and procedure

As a replication of Study 1, we followed the very same procedure: Participants were again randomly assigned to one of three between-subjects conditions (*explicit mutual knowledge*, *implicit mutual knowledge*, *no mutual knowledge*). In the *implicit mutual knowledge* condition, the scenario reads as follows:

> Betty is a researcher and she is about to attend a workshop in New York along with her team. She is now at the airport, waiting to board her flight. Her colleague and co-presenter Ann will be flying directly to New York from her hometown and meeting Betty and the rest of the team at the workshop. While thinking about her presentation at the boarding gate, Betty realizes that it would be a good idea to include an analysis that Ann did a year earlier. This would help them to impress the team leader at the workshop. So, Betty sends an e-mail to Ann, asking her to bring this material to New York.
>
> When Betty arrives in New York, the night before the workshop, she checks her e-mail inbox. She sees that she has received a read receipt from Ann's account, confirming that she (Ann) read the e-mail a couple of minutes after Betty sent it.
>
> As it happens, Ann did not bring her hard-drive with the earlier analysis to New York. So, she and Betty do not have this material at the workshop, and do not manage to impress their team leader with their results.

In the *explicit mutual knowledge* condition, the vignette differs insofar as the sender gives a verbal explicit reassurance to the recipient, whereas in the *no mutual knowledge* condition the vignette differs as the sender did not receive the information that the recipient was relying on her.[4]

The target questions were the same as in Study 1, with minor adjustments related to the activity in which the characters were intending to engage. We again controlled for participants' understanding of the text by asking three control questions, the last of which being particularly important because it reveals whether participants understood the critical manipulation.

The questions were presented to participants in a randomized order, except for Control Question 3, which was always presented last, since we determined that could influence responses to the other questions. Responses from those who failed to answer the control questions correctly were discarded from the final sample.

---

[4]See https://osf.io/h3mjp/ for the full vignettes.

3.2 | Results

The results are in line with those of Study 1. A Kruskal–Wallis Test showed that the responses to the normative question were significantly different in the three conditions, $\chi^2(2) = 25.8$, $p < .001$, $\eta^2 = 0.21$ (medium effect size), with a mean rank rate of 44.36 for the *no mutual knowledge* condition, a mean rank rate of 71.06 for the *implicit mutual knowledge* condition, and a mean rank rate of 79.90 for the *explicit mutual knowledge* condition. A series of post hoc pairwise comparisons tests showed that the responses were significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$) and in the *implicit mutual knowledge* condition ($p = .001$). No significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition, ($p = .820$; see Figure 4).
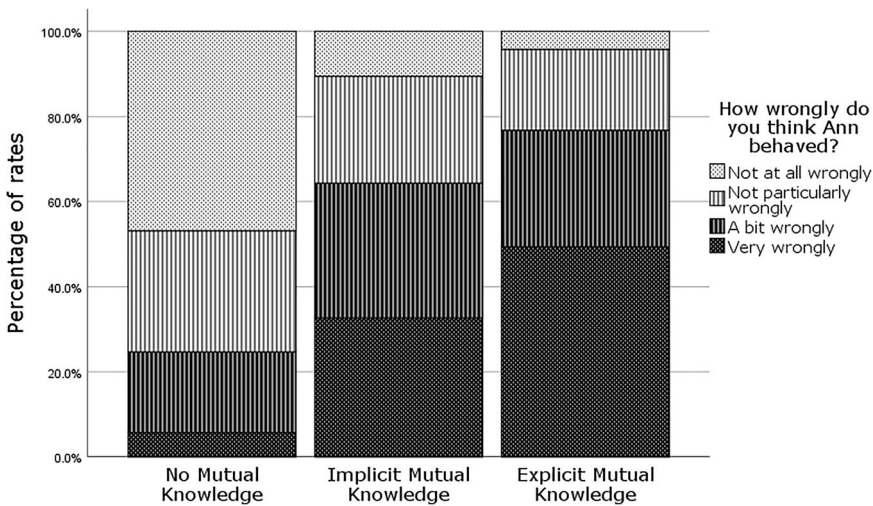

FIGURE 4 Level of perceived commitment--normative question. The responses to the normative question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 123$, $\chi^2(2) = 25.8$, $p < .001$, $\eta^2 = 0.21$

Consistently with the previous findings, the responses to the affective question showed a similar pattern compared to the normative question: A Kruskal–Wallis Test showed that the responses were significantly different in the three conditions, $\chi^2(2) = 7.5$, $p = .024$, $\eta^2 = 0.06$ (small effect size), with a mean rank rate of 53.30 for the *no mutual knowledge* condition, a mean rank rate of 64.56 for the *implicit mutual knowledge* condition, and a mean rank rate of 73.29 for the *explicit mutual knowledge* condition. However, a series of post hoc pairwise comparison tests showed that the responses were significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p = .023$) but no significantly lower than in the *implicit mutual knowledge* condition ($p = .315$). As predicted, no significant difference is found between the *implicit mutual knowledge* condition (and the *explicit mutual knowledge* condition; $p = .808$; see Figure 5). The responses to the affective question predictably correlated with the responses to the normative question, $r_s(123) = .584$, $p < .001$.

The responses to the partner choice question confirmed the results found in Study 1: The responses were significantly different in the three conditions, Kruskal–Wallis Test, $\chi^2(2) = 20.3$,
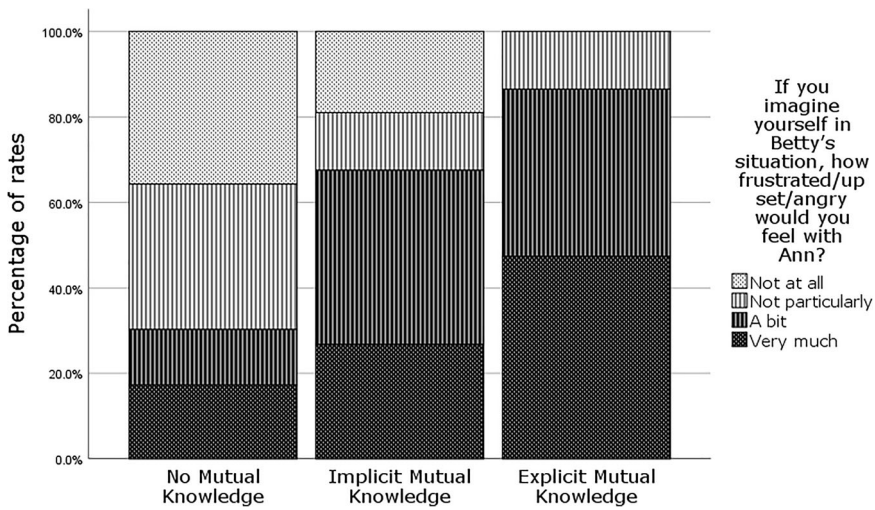
FIGURE 5 Level of perceived commitment--affective question. The responses to the affective question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 123$, $\chi^2(2) = 7.5$, $p = .024$, $\eta^2 = 0.06$

$p < .001$, $\eta^2 = 0.17$ (medium effect size), with a mean rank rate of 77.84 for the *no mutual knowledge* condition, a mean rank rate of 52.15 for the *implicit mutual knowledge* condition, and a mean rank rate of 48.15 for the *explicit mutual knowledge* condition. Again, a series of post hoc pairwise comparisons tests were run. The results showed that the responses were significantly higher in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$) and in the *implicit mutual knowledge* condition ($p = .001$). Consistently with our hypothesis, no significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = 1.000$; see Figure 6).
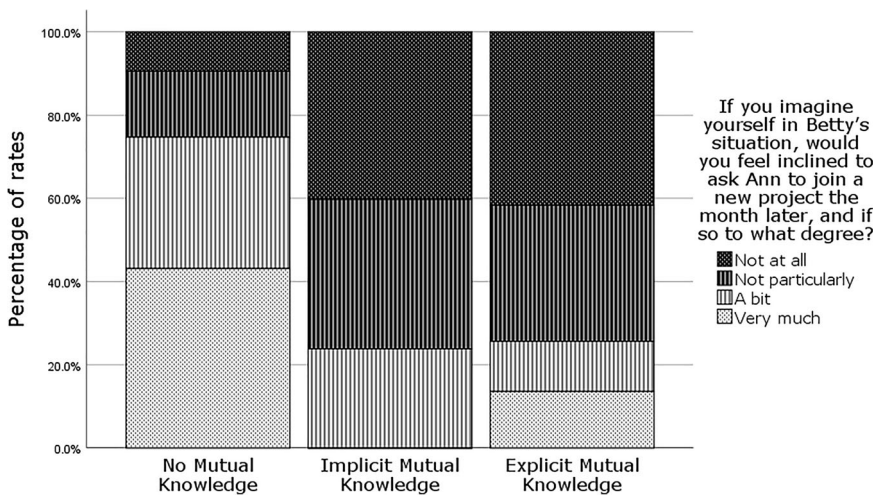

FIGURE 6 Level of perceived commitment--partner choice question. The responses to the partner choice question are significantly higher in the No Mutual Knowledge condition than in the Implicit Mutual Knowledge and the Explicit Mutual Knowledge condition, Kruskal–Wallis Test, $N = 123$, $\chi^2(2) = 20.3$, $p < .001$, $\eta^2 = 0.17$

The responses to the partner choice question correlated significantly with the responses to the normative question, r_s (123) = .556, $p < .001$, and with the responses to the affective question, r_s (123) = .553, $p < .001$.

3.3 | Discussion

The results of this second study confirmed our previous findings. The variation in the narrative, as well as the kind of relationship between the two agents, did not affect the pattern we observed previously.

4 | STUDY 3

In Studies 1 and 2 mutual knowledge resulted from a technological device. We designed Study 3 to probe whether commitment can also arise when minimal cues of mutual knowledge are present, such as when it results from a joint attentional process. Participants read descriptions of what we intended to be evidence of mutual knowledge: eye contact, joint attention to a relevant stimulus, and ostensive silence (as suggested by Carpenter & Liebal, 2011). Furthermore, in Studies 1 and 2 the *no mutual knowledge* conditions present the following structure: The sender does not lead the recipient to rely on X, and no mutual knowledge about the recipient's reliance is present. To more directly test our claim that a sense of commitment is critically influenced also by the fact that it is mutually known by the agents that the sender had raised the recipient's expectations, Study 3 implemented a situation in which the sender always led the recipient to rely on X: In the *no mutual knowledge* condition, this is unknown to the sender, while this is mutually known by the agents in both the *implicit mutual knowledge* and in the *explicit mutual knowledge* conditions.

4.1 | Methods

4.1.1 | Participants

Participants were recruited together with participants for Study 2. From the original dataset, 81 participants were assigned to Study 3 (44 females; M_{age} = 37.48 years, SD = 10.88), 20 in the *no mutual knowledge* condition, 23 in the *implicit mutual knowledge* condition, and 38 in the *explicit mutual knowledge* condition. As participants were recruited via Amazon M-Turk, the sample was composed entirely by North Americans.

4.1.2 | Design and procedure

In the *implicit mutual knowledge* condition, the scenario reads as follows:

> Jenny and Lisa are two colleagues who work at the same office and get along well. This coming Friday evening, there is an office party taking place in the office lounge. Jenny thinks that it would be a good idea to attend the party, but

she usually feels very awkward at such events. Everyone in the office, included Lisa, knows that Jenny always attends parties like this if Lisa, who is very chatty and easygoing, also attends. On Friday morning, Jenny and Lisa are talking with their boss about the party in the evening. Since Lisa was carrying a couple of bottles of wine to the lounge, Jenny inferred that she was intending to go to the party. So she says to both Lisa and their boss that she will be at the party and that she is looking forward to tasting Lisa's wine. Lisa smiles to her, and the boss replies that he is happy that she (Jenny) will be attending. However, on Friday afternoon Lisa gets a call from a friend whom she has not seen for a long time. Lisa then decides not to go to the party. Jenny is very bored and does not particularly like any of the people at the party. She wishes that she had spent the evening somewhere else.

The procedure was identical to the one of Study 2, and the target and control questions were the same as in Study 1, with minor corrections related to the activity the characters would engage.

4.2 | Results

The results show very different patterns. A Kruskal–Wallis test revealed that the responses to the normative question were significantly different in the three conditions, $\chi^2(2) = 6.05$, $p = .048$, $\eta^2 = 0.08$ (small effect size). A series of post hoc tests showed no significant differences between each of the three conditions: A marginally significant difference was found between the *no mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .078$); a nonsignificant difference between the *no mutual knowledge* condition and the *implicit mutual knowledge* condition ($p = 1.000$); and a nonsignificant difference between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .224$).

In contrast to the previous findings, the responses to the affective question were not significantly different in the three conditions, Kruskal–Wallis test, $\chi^2(2) = 1.325$, $p = .516$. The responses to the affective question correlated significantly with responses to the normative question, $r_s(81) = .573$, $p < .001$.

And again, the responses to the partner choice question were not significantly different in the three conditions, Kruskal–Wallis test, $\chi^2(2) = 3.865$, $p = .145$. The responses to the partner choice question correlated significantly both with the responses to the normative question, $r_s(81) = .281$, $p = .011$, and with the responses to the affective question, $r_s(81) = .358$, $p < .001$.

4.3 | Discussion

It seems that the changes we implemented in Study 3 influenced participants' responses. The results of Study 3, which were not predicted, could be explained in three different ways: (a) the way we implemented mutual knowledge in Study 3 may not have been clear to participants—this is partially confirmed by the fact that almost one third of our participants ($N = 32$, 27.6%) failed the comprehension question about the epistemic stance of the sender, thus undermining the reliability of the correct answers; (b) the cues of joint attention we described, that is, eye contact, are not by themselves sufficient cues to mutual knowledge, contrary to previous evidence (Siposova, Tomasello & Carpenter, 2018; Thomas, DeScioli & Haque, 2014); (c) the study

on its own lacked the power needed to detect a small effect size; or (d) the fact that the recipient's reliance is the only factor influencing a sense of commitment, provided that these expectations were raised by the sender but irrespective of whether this is mutually known.

We believe that both (a) and (c) are likely explanations. Thus, we ran an additional study to address these concerns. Having only one type of vignette, we maintained a higher sample size to assure that the test would have enough statistical power, and we decided to present the story with a different modality rather than a verbal vignette.

5 | STUDY 4

Given that the inconclusive results of Study 3 might have been due to the manner in which we implemented the manipulation, we decided to replicate the study with a different design. We therefore implemented a different story, in which mutual knowledge was established by a joint attentional process rather than by a technological device. We also chose a different modality rather than a verbal narration of hypothetical events, namely a photo-story, with real people acting out a script. This particular type of design also has the advantage of increasing the plausibility of the scenario, which is now more likely to be interpreted as something the participants are witnessing rather than merely imagining, thus increasing the ecological validity.

5.1 | Methods

5.1.1 | Participants

We used Amazon M-Turk to implement a web-based paradigm with a between-subjects design. In view of the small effect sizes found in the previous studies, a power analysis using G*Power 3.1 indicated that a total sample size of 303 participants would be needed to detect the expected effect size ($f = 0.18$) (derived from a predicted statistical power of 80% using a one-way ANOVA with alpha at .05). We added 15% to our desired sample, thus we aimed to collect 348 participants. We included data from those participants who had already begun the experiment when M-Turk registered that this number had been reached. Our data set therefore comprised 370 adults, who were rewarded with $0.60 each. Data was discarded from participants who did not complete the survey ($N = 15$) or who failed one or more control question ($N = 117$), totaling 238 participants (121 females; $M_{age} = 38.30$ years, $SD = 12.26$)—93 in the *no mutual knowledge* condition, 64 in the *implicit mutual knowledge* condition, and 81 in the *explicit mutual knowledge* condition.

5.1.2 | Design and procedure

Participants were presented with the same basic scenario: For one group of participants the expectations of the agents were not mutually known, for a second group these expectations were mutually known because the sender acknowledged them explicitly, and for a third group these expectations were mutually known because the sender acknowledged them implicitly.

Participants were randomly assigned to one of the three between-subjects conditions (*explicit mutual knowledge*, *implicit mutual knowledge*, *no mutual knowledge*). The scenario was presented as a photo story, as depicted in Figure 7.

FIGURE 7 Participants were presented with photo stories which differed according to the three conditions

In the *explicit mutual knowledge* condition, the vignette differs insofar as the sender gives an explicit verbal reassurance to the recipient, whereas in the *no mutual knowledge* condition the vignette differs insofar as the sender was not exposed to the information.[5] The target questions were the same as in Study 1, with minor adjustments related to the activity the actors were

---

[5]See https://osf.io/h3mjp/ for the full vignettes.

engaged in. We controlled for participants' understanding of the text by asking two control questions. The second control question was particularly important because it revealed whether participants had understood the critical manipulation. Since being forced to make a judgment about the epistemic states of the agents could have an effect on responses to the other test questions, this question was always presented last and on a different page. Except for the second control question, which was always presented last, the questions were presented to the participants in a randomized order. Data from those who failed to answer the control questions correctly were discarded from the final sample.

5.2 | Results

We predicted that responses to the normative question would be significantly higher in the *explicit mutual knowledge* and in *implicit mutual knowledge* conditions than in the *no mutual knowledge* condition. Critically for our hypothesis, we predicted that the rates would not be significantly different between the *explicit mutual knowledge* and the *implicit mutual knowledge* conditions. To test these hypotheses, we ran a Kruskal–Wallis nonparametric test and a series of post hoc tests per measure.

Consistently with the predictions, a Kruskal–Wallis test revealed that the responses to the normative question were significantly different in the three conditions, $\chi^2(2) = 34.1$, $p < .001$, $\eta^2 = 0.14$ (medium effect size), with a mean rank rate of 89.49 for the *no mutual knowledge* condition, a mean rank rate of 139.55 for the *implicit mutual knowledge* condition, and a mean rank rate of 138.11 for the *explicit mutual knowledge* condition). A series of post hoc tests showed that the responses were significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition, $p < .001$); and in the *implicit commitment* condition ($p < .001$). However, no significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = 1.000$; see Figure 8). This confirms the
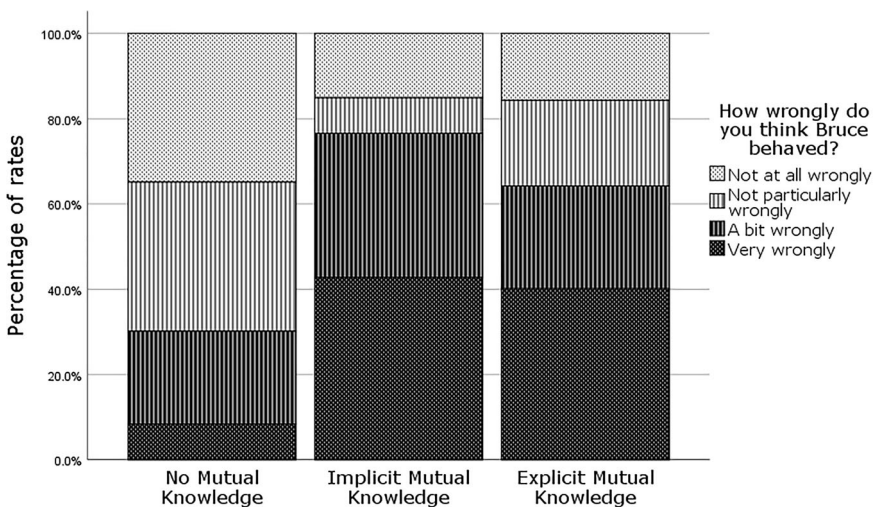

FIGURE 8 Level of perceived commitment--normative question. The responses to the normative question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 238$, $\chi^2(2) = 34.1$, $p < .001$, $\eta^2 = 0.14$

hypothesis that the levels of perceived commitment are higher in conditions in which the expectations are mutually known by the agents.

A Kruskal–Wallis test showed that the responses to the affective question were significantly different in the three conditions, $\chi^2(2) = 19.3$, $p < .001$, $\eta^2 = 0.08$ (small effect size), with a mean rank rate of 97.34 for the *no mutual knowledge* condition, a mean rank rate of 131.95 for the *implicit mutual knowledge* condition, and a mean rank rate of 135.10 for the *explicit mutual knowledge* condition). Consistently with the predictions, the responses showed the same pattern as for the normative question: A series of post hoc tests revealed that responses were significantly lower in the *no mutual knowledge* condition than in the *explicit mutual knowledge* condition ($p < .001$), and in the *implicit mutual knowledge* condition ($p = .002$). However, no significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = 1.000$; see Figure 9). The responses to the affective question were significantly correlated with the responses to the normative question, r_s (238) = .661, $p < .001$.

On the other hand, the pattern presented by the partner choice questions was slightly different: The rates of willingness to interact again with the sender were significantly different in the three conditions, Kruskal–Wallis test: $N = 238$, $\chi^2(2) = 9.30$, $p = .010$, $\eta^2 = 0.04$ (small effect size), with a mean rank rate of 134.02 for the *no mutual knowledge* condition, a mean rank rate of 103.98 for the *implicit mutual knowledge* condition, and a mean rank rate of 115.10 for the *explicit mutual knowledge* condition. A series of post hoc pairwise comparisons tests revealed no significant difference between the *no mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .142$), but rates were significantly higher in the *no mutual knowledge* condition than in the *implicit mutual knowledge* condition ($p = .010$). No significant difference was found between the *implicit mutual knowledge* condition and the *explicit mutual knowledge* condition ($p = .867$) (see Figure 10).

The responses to the partner choice question were significantly correlated both with the responses to the normative question, r_s (238) = .415, $p < .001$, and with the responses to the affective question, r_s (238) = .367, $p < .001$.
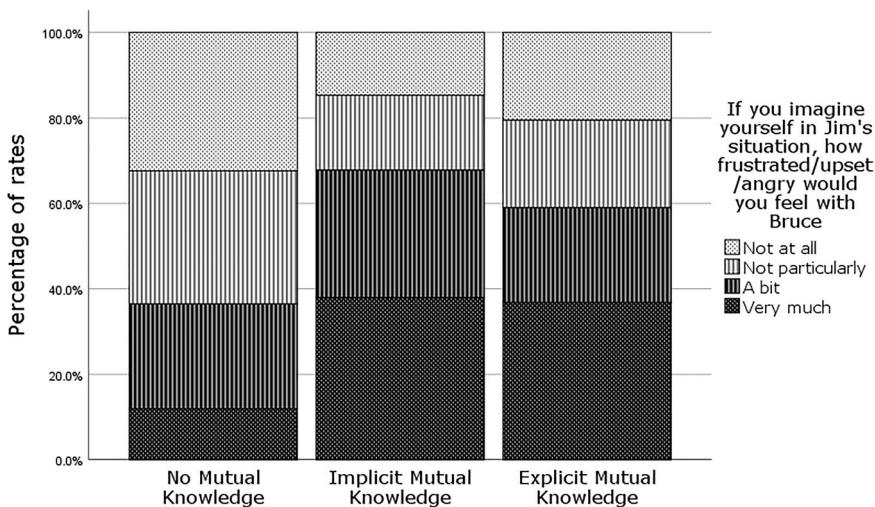

FIGURE 9 Level of perceived commitment--affective question. The responses to the affective question are significantly lower in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 238$, $\chi^2(2) = 19.3$, $p < .001$, $\eta^2 = 0.08$
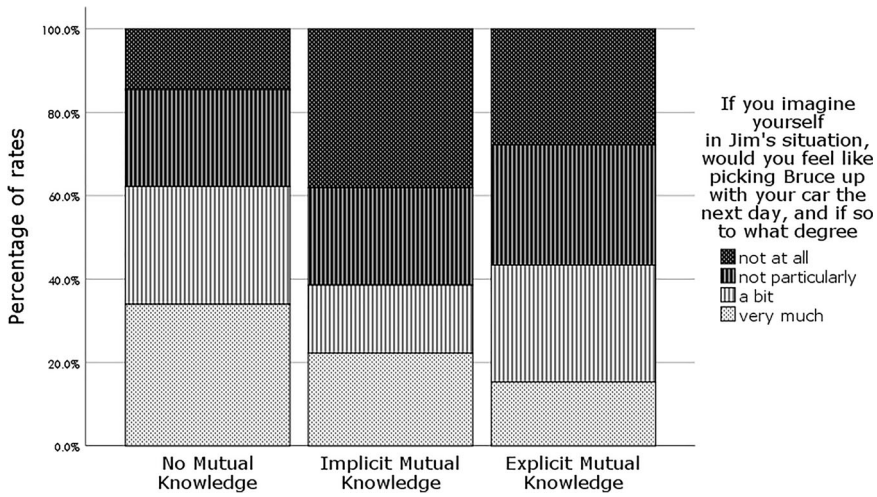
FIGURE 10 Level of perceived commitment--partner choice question. The responses to the partner choice question are significantly higher in the No Mutual Knowledge condition than in the Implicit and the Explicit Mutual Knowledge conditions, Kruskal–Wallis Test: $N = 238$, $\chi^2(2) = 9.30$, $p = .010$, $\eta^2 = 0.04$

5.3 | Discussion

The results of this study confirm the predictions and replicated the results found in Study 1 and Study 2. The fact that the recipient's expectations were raised by the sender enhances participant's sense of commitment only when this is mutually known by the agents. The new methodology also conveys the idea that when it is easier to track participants' epistemic states, eye contact is a sufficient trigger of mutual knowledge, as found by Thomas et al. (2014).

6 | GENERAL DISCUSSION

There are several ways in which a sender can lead a recipient to expect and rely on X, such as uttering a statement that constitutes a commissive speech act, performing an action, or simply omitting to prevent someone from having expectations. For instance, if your friend wants you to attend her party, and you both know that unless you say otherwise, she would expect you to attend, then your silence may be taken to signal your intention to attend, and may thereby generate a sense of commitment to attend. Such cases show that commitment can arise even when the sender does not utter a commissive speech act, such as a promise or an oath, although such acts are efficient means of making expectations mutually known. Likewise, it is not necessary that the sender explicitly acknowledge her recipient's expectations, nor that the sender *intended* to cause her recipient to expect X for a commitment to arise. For instance, if your dog notices that you are picking up a ball that had been lying on the floor, it is plausible that you will feel committed to playing fetch together, since your action, although unintended, has generated an expectation on the part of your dog that you will play fetch together (see Michael et al., 2016a). Thus, a sense of commitment can arise if the sender leads (voluntarily or not) the recipient to have expectations about her behavior, if the recipient relies on this expectation, and if this mutually known by them.

This is indeed what we found across a series of four studies. More precisely, we found evidence in support of the hypothesis that the perception of commitment is critically influenced by the extent to which the fact that a recipient has been led by a sender to expect her to do X (Studies 1 and 2), and that the recipient is going to rely on her to do X, is mutually known (Study 4). If it is mutually known that a recipient has been led by a sender to expect her to do X, and that the recipient is going to rely on her to doing X, but the sender does not do X, the recipient will hold her accountable. In line with this, the results from our studies indicate that participants evaluated the sender more severely when the recipient's reliance was mutually known than when it was not, irrespective of how their mutual knowledge had been established (i.e., whether or not the sender performed a speech act). It is worth noting that across the four studies the degree of certainty that the agents could have about whether the knowledge was mutual (i.e., whether there was first-order, second-order, or higher-order knowledge about the recipient's reliance) may well have differed—in Studies 3 and 4, in which mutual knowledge is implemented via cues of joint attention, the degree of certainty is greater than in Studies 1 and 2, in which mutual knowledge is implemented via in-built features of the technological device used. As much as deniability is reduced in these latter cases, some degree of uncertainty is still present. It is interesting to note, however, that even in these cases in which it is unclear whether knowledge is mutual or shared, people would often still negotiate in terms of what judgments would be made if knowledge were mutual (see Misyak & Chater, 2014).

Our findings are difficult to reconcile with the hypothesis, suggested by speech act theories, that commitments require speech acts indicating the intention of the speaker to incur a moral obligation to perform a particular action (or to refrain from doing so) (Austin, 1962; Searle, 1969). They are also difficult to reconcile with the conventionalist theories of promises, according to which promising is essentially a socially defined convention enabling coordination and trust within a group (Hume, 1739/1969; Rawls, 1955). While these views differ in important ways, they share at least one important feature—namely, they neglect the phenomenon of unconventional nonverbal commitment.

In contrast, our findings show that Scanlon's account of commitment accurately describes the way people perceive commitment. Scanlon links commitment to the expectations and the reliance of a recipient: According to his theory of promises, the moral norm that we ought to keep our promises is grounded in the fact that promises generate expectations—that is, promising to do something creates in the recipient the expectation that the sender will do it (Scanlon, 1998, pp. 295–302). Our results are also consistent with MacCormick and Raz's claim that when one individual has intentionally led another to rely on her, she is then committed to living up to the other agent's expectation (1972), as well as with Gilbert's analysis, which accords a decisive role to common knowledge in the creation of joint commitments, and which does not require speech acts (1990, 2006). Furthermore, our findings also accommodate some theories of social norms that are grounded in reasonable expectations (Bicchieri, 2006; Sugden, 2000), and they are consistent with previous studies showing that people exhibit an aversion to disappointing others' expectations (Dana, Cain & Dawes, 2006; Ockenfels & Werner, 2012), provided that these expectations are not unreasonable (Heintz, Celse, Giardini & Max, 2015).

It must be noted that speech act theory, conventionalist accounts of promises and social norm theories are concerned with the normative components of commitment; that is, they do not directly address the issue of its psychological implementation. Thus, none of our findings directly refute these theories. On the other hand, our participants did engage in moral reasoning, which is better captured by an expectation-based explanation. Studies 1 and 2 implement

scenarios in which standardized technology-based signals are used as cues to acknowledge the recipient's expectations. One potential limitation of these scenarios is that these standardized technology-based signals could potentially be interpreted as conventionalized nonverbal speech acts (like nodding), at least in those groups in which they are commonly used. After all, these signals have the sole function of indicating to the users that the message has been received and read. Study 4, however, overcomes this limitation, and strengthens the claim that the perception of commitment is not tied to conventional rules or agreements.

Our findings also confirm the prediction that people's assessments of commitment violations influence their partner choices—even in cases in which the commitment was not generated by any speech act. This is important insofar as it highlights the reputational costs of violating commitments even in the absence of speech acts, and thereby also illuminates why agents are so often motivated to honor their commitments (with or without speech acts). In other words, in cases where the expectations of the recipient are mutual knowledge between the recipient and the sender, the sender may anticipate that she would face reputational costs if she did not fulfill these expectations (or at least warn the recipient before disappointing her expectations). Specifically, potential partners in the future may not be willing to rely on her, and may therefore choose not to interact with her. As a result, even in cases in which it would be in the sender's short-term interests not to honor the commitment, the long-term net effects may be negative. This is why mutual knowledge of a recipient's expectations about a sender's future actions—in particular in cases in which the expectations derive from an action performed by the sender—can be sufficient to generate a credible commitment. Interestingly, this point resonates with an observation which Hume made within the framework of contractualism: He noted that when an agent is expected to perform the action that expressed an intention to perform, the agent "subjects himself to the penalty of never being trusted again in case of failure" (Hume, T3.2.5.10). While confirming his intuition, our results show that this does not presuppose that commitment-keeping is a conventional practice; it is sufficient if the information flow within the group enables individuals to select their cooperators based on the reputations of the potential partners (Nowak & Sigmund, 1998).

Our findings also open up new avenues for further investigation. Our manipulation was designed to vary whether the recipient's reliance and expectations are mutually known. We implemented mutual knowledge with technology-based signals that limit the plausible deniability of one's knowledge of the partner's reliance, but also with minimal cues of joint attention. Some authors have claimed that eye contact is a potent cue of common knowledge, as eye contact can indicate to both parties that each is aware of the other attending a certain stimulus (in this case, the stimulus is the need of the recipient, and of her reliance on the sender's action; Siposova et al., 2018; see Carpenter & Liebal, 2011; Thomas et al., 2014). It would be important for future research to probe the effects of different ways of generating different levels of knowledge.

Moreover, in the scenarios implemented here, the presence or absence of mutual knowledge may also have influenced the degree to which participants attributed expectations to the recipient of the commitment. It is possible, for instance that where expectations are not mutually known, participants may have doubted whether the recipient really expected the sender to perform the action in question. It would be valuable for future studies to manipulate mutual knowledge independently of the strength of expectations.

ORCID

Francesca Bonalumi 🄳 https://orcid.org/0000-0001-5598-376X

REFERENCES

Alonso, F. M. (2016). Reasons for reliance. *Ethics*, *126*(2), 311–338. https://doi.org/10.1086/683536

Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Carpenter, M. & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seeman (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–182). Cambridge, MA: The MIT Press.

Clark, H. H. (2006). Social actions, social commitments. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 126–150). New York, NY: Bloomsbury.

Dana, J., Cain, D. M. & Dawes, R. M. R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193–201. https://doi.org/10.1016/j.obhdp.2005.10.001

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiske, S. T., Cuddy, A. J. C. & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, *15*, 1–14. https://doi.org/10.1111/j.1475-4975.1990.tb00202.x

Gilbert, M. (2006). *A theory of political obligation*. Oxford: Oxford University Press.

Heintz, C., Celse, J., Giardini, F. & Max, S. (2015). Facing expectations: Those that we prefer to fulfil and those that we disregard. *Judgment and Decision Making*, *10*(5), 442–455.

Heintz, C., Karabegovic, M. & Molnar, A. (2016). The co-evolution of honesty and strategic vigilance. *Frontiers in Psychology*, *7*, 1503. https://doi.org/10.3389/fpsyg.2016.01503

Hume, D. (1969/1739). *A treatise of human nature (Original work published 1739–1740, Vol. 26)*. London: Penguin.

Lehmann, E. L. (2006). *Nonparametrics: Statistical methods based on ranks*. New York, NY: Springer-Verlag.

Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

Lewis, D. K. (1978). Truth in fiction. *American Philosophical Quarterly*, *15*, 37–46.

Liddell, T. M. & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, *79*, 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

MacCormick, N. & Raz, J. (1972). Voluntary obligations and normative powers. *Proceedings of the Aristotelian Society. Supplementary Volumes*, *46*, 59–102. https://doi.org/10.1097/EDE.ObO13e31812e5535

Michael, J., Sebanz, N. & Knoblich, G. (2016a). The sense of commitment: A minimal approach. *Frontiers in Psychology*, *6*, 1968. https://doi.org/10.3389/fpsyg.2015.01968

Michael, J., Sebanz, N. & Knoblich, G. (2016b). Observing joint action: Coordination creates commitment. *Cognition*, *157*, 106–113. https://doi.org/10.1016/j.cognition.2016.08.024

Misyak, J. B. & Chater, N. (2014). Virtual bargaining: A theory of social decision-making. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1655), 20130487. https://doi.org/10.1098/rstb.2013.0487

Nowak, M. A. & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *437*, 573–577. https://doi.org/10.1038/31225

Ockenfels, A. & Werner, P. (2012). "Hiding behind a small cake" in a newspaper dictator game. *Journal of Economic Behavior & Organization*, *82*(1), 82–85.

Rawls, J. (1955). Two concepts of rules. *The Philosophical Review*, *64*(1), 3–32.

Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.

Schelling, T. C. (1980). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Schiffer, T. (1972). *Meaning*. Cambridge, MA: Harvard University Press.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Siposova, B., Tomasello, M. & Carpenter, M. (2018). Communicative eye contact signals a commitment to cooperate for young children. *Cognition*, *179*, 192–201. https://doi.org/10.1016/j.cognition.2018.06.010

Sperber, D. & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Oxford: Blackwell Publishing.

Sugden, R. (2000). The motivating power of expectations. In J. Nida-Rümelin & W. Spohn (Eds.), *Rationality, rules, and structure* (pp. 103–129). Dodrecht: Springer.

Thomas, K. A., DeScioli, P. & Haque, O. S. (2014). The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology*, *107*(4), 657–676. https://doi.org/10.1037/a0037037

Vanderschraaf, P. & Sillari, G. (2014). Common knowledge. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2014 ed.). Stanford, CA: Stanford University Retrieved from: http://plato.stanford.edu/archives/spr2014/entries/common-knowledge/