Journal of Experimental Psychology: General

Recursive Mentalizing and Common Knowledge in the Bystander Effect

Kyle A. Thomas, Julian De Freitas, Peter DeScioli, and Steven Pinker

Online First Publication, February 25, 2016. http://dx.doi.org/10.1037/xge0000153

Journal of Experimental Psychology: General

0096-3445/16/$12.00 http://dx.doi.org/10.1037/xge0000153

Recursive Mentalizing and Common Knowledge in the Bystander Effect

Kyle A. Thomas
Harvard University

Julian De Freitas
Harvard University, and University of Oxford

Peter DeScioli
Stony Brook University

Steven Pinker
Harvard University

The more potential helpers there are, the less likely any individual is to help. A traditional explanation for this *bystander effect* is that responsibility diffuses across the multiple bystanders, diluting the responsibility of each. We investigate an alternative, which combines the *volunteer's dilemma* (each bystander is best off if another responds) with recursive theory of mind (each infers what the others know about what he knows) to predict that actors will strategically shirk when they think others feel compelled to help. In 3 experiments, participants responded to a (fictional) person who needed help from at least 1 volunteer. Participants were in groups of 2 or 5 and had varying information about whether other group members knew that help was needed. As predicted, people's decision to help zigzagged with the depth of their asymmetric, recursive knowledge (e.g., "John knows that Michael knows that John knows help is needed"), and replicated the classic bystander effect when they had common knowledge (everyone knowing what everyone knows). The results demonstrate that the bystander effect may result not from a mere diffusion of responsibility but specifically from actors' strategic computations.

Keywords: bystander effect, diffusion of responsibility, volunteer's dilemma, common knowledge, theory of mind

Supplemental materials: http://dx.doi.org/10.1037/xge0000153.supp

Whether people are doing household chores, responding to dangerous emergencies, collaborating on research, or reacting to international crises, the more potential helpers who are available, the less likely any individual is to help. Psychologists have argued that this *bystander effect* occurs because responsibility diffuses across the entire set of potential helpers, making it proportionally less likely that each will intervene (Darley & Latané, 1968; Latané & Darley, 1968, 1970; Latané & Nida, 1981; see Fischer et al., 2011 for a review).

Imagine, for example, renting a house from an absentee landlord who needs a tenant to change the oil filter on the building's furnace, or else the filter will inevitably clog, and the whole building will be out of heat and hot water for several days until it is repaired. If you were the only tenant on the premises, you would feel you had little choice, but if other tenants were there, you would just as soon that one of them did it.

A little thought reveals, though, that it is not just the *existence* of other potential helpers that should affect your decision, but your state of *knowledge* about them, and vice versa. If you never saw or met your neighbors, you might still shoulder the task, not taking a chance that it would be left undone. But now suppose you overhear the landlord asking another tenant whether anyone has changed the oil filter. If you duck behind a wall, you can leave this other tenant with the sole responsibility to help.

Insights from game theory can make this intuition more precise. The sociologist Andreas Diekmann (1985, 1986) used a game-theoretic model called the *volunteer's dilemma* to argue that diffusion of responsibility results from strategic behavior. In the volunteer's dilemma, a group of people confront a problem that can be resolved by one volunteer's help, but helping has costs such as time, risk, and foregone opportunities (Diekmann, 1985, 1986). If no one helps, then everyone pays a cost greater than a given cost of helping. Figure 1 shows the payoffs as sums of money, though they could also be psychological, such as the effort of changing an oil filter and the inconvenience of being left without heat, or, as in the classic scenarios in bystander intervention research, the empathic satisfaction of learning that a needy person has been helped or the distress of learning that he has not. Thus, everyone wants someone to help, but prefers if possible that it not be them. Diekmann showed that strategic reasoning in this game leads to less helping when there are more players, resulting from the *mixed strategy equilibrium* in the game.

Kyle A. Thomas, Department of Psychology, Harvard University; Julian De Freitas, Department of Psychology, Harvard University, and Department of Experimental Psychology, University of Oxford; Peter DeScioli, Department of Political Science, Stony Brook University; Steven Pinker, Department of Psychology, Harvard University.

Kyle A. Thomas and Julian De Freitas are co-first authors. We thank Patrick Mair for statistical assistance.

Correspondence concerning this article should be addressed to Kyle A. Thomas, Department of Psychology, Harvard University, William James Hall 964, 33 Kirkland Street, Cambridge, MA 02138. E-mail: kylethomas@post.harvard.edu

		Other Actor(s)	
		Help	Shirk
Actor	Help	$0.50, $0.50	$0.50, $1.00
	Shirk	$1.00, $0.50	$0, $0

Figure 1. Payoff matrix in the volunteer's dilemma. The focal actor's decisions are represented in the two rows, while the decision(s) of the other actor(s) are represented in the two columns. The four cells show the payoffs for each possible combination of decisions, with the focal actor's payoff given on the left, and the payoff of the other actor(s) given on the right.

In game theory, players can try to outwit their opponents by using mixed strategies that employ probabilistic choices. For example, poker players, soccer goalies, and military generals try to outguess their opponents and to avoid being predictable themselves. Players in the volunteer's dilemma are similarly in an outguessing standoff and could benefit by making probabilistic choices. At the mixed strategy equilibrium, players choose a probability of helping such that each player expects the same payoff whether they help or shirk. There is a unique nonzero probability of helping at which this indifference occurs (depending on the game's payoffs and the number of other players), and this outcome is stable because no individual can improve her payoffs by changing her probability of helping. Relevant to the bystander effect, the probability of helping at equilibrium decreases as the number of players increases. Diekmann's (1986, 1993) experiments with the volunteer's dilemma support the hypothesis that reduced helping stems from strategic decisions, rather than only dividing responsibility among group members.

The second key idea from game theory for understanding the bystander effect (and other social dilemmas) is that a player's best move depends on whether all players have *common knowledge* of the situation, meaning that all individuals know that all other individuals know about the need to help. When players have common knowledge, they occupy symmetric positions in the outguessing standoff, and each player's best move can be playing a mixed strategy. But there are several patterns of asymmetric knowledge that can also occur in groups, in which the optimal strategy depends on what they and the other bystanders know. These include *private knowledge,* when an individual possesses knowledge without knowing whether anyone else possesses it; *secondary knowledge,* when an individual possesses knowledge and also knows that another agent possesses it; *tertiary knowledge,* when an individual possesses knowledge and also knows that another agent possesses it, and that the other agent knows that the individual possesses the knowledge; *quaternary knowledge,* when an individual possesses knowledge and also knows that another agent possesses it, and that the other agent knows that the individual both possesses the knowledge and knows that the agent also possesses it.

In these cases with asymmetric knowledge, the pattern of knowledge determines which choice is the agent's best move, and the number of bystanders is less important than it is with common knowledge. For example, a tenant with private knowledge that the furnace's oil filter needs to be changed should do the chore since he is unsure anyone else knows about it. In contrast, a tenant with secondary knowledge that another tenant knows about the filter can leave him to do the chore because this other tenant only has private knowledge. These situations differ from when two tenants find out about the filter at the same time while making eye contact, generating common knowledge.

Despite the straightforward relevance of strategic analysis to understanding the bystander effect, it has played little role in the voluminous literature on the phenomenon.[1] Nor has the role of potential helpers' state of knowledge been considered important; most analyses of the bystander effect argue that diffusion of responsibility does not depend on what other people know, but only on the number of potential helpers. In their classic review, Latané & Nida (1981) claimed that, "the knowledge that others are present and available to respond, *even if the individual cannot see or be seen by them,* allows the shifting of some of the responsibility for helping to them" (p. 309, emphasis added).

The neglect of actors' knowledge in understanding the bystander effect is a significant gap, because in other areas of psychology, there has been enormous interest in understanding how people represent the mental states of others (mentalizing, mind-reading, or "theory of mind"; Baron-Cohen, 1995; Frith & Frith, 2003; Wimmer & Perner, 1983; for recent reviews, see Apperly & Butterfill, 2009; Saxe & Young, 2013). More recently, a key role has been given to *recursive* mentalizing (people's knowledge of what other people know), particularly people's sensitivity to common knowledge, which has been argued to explain diverse social phenomena such as economic cooperation, innuendo and indirect speech, public ceremonies and rituals, self-conscious emotions, and political collective action (Chwe, 2001; Lee & Pinker, 2010; Pinker, 2007; Pinker, Nowak, & Lee, 2008; Thomas, DeScioli, Haque, & Pinker, 2014).

One reason for the neglect of knowledge and its strategic implications in understanding the bystander effect is that previous research has not systematically varied potential helpers' state of knowledge. In previous experiments, group members almost invariably had common knowledge of the situation: all individuals knew that help was needed, all knew that the other individuals knew it, and so on (e.g., Darley & Latané, 1968; Latané & Darley, 1968, 1970; Latané & Nida, 1981). Common knowledge is exactly the state that sets up the volunteer's dilemma and predicts the associated failure of individuals to help. Indeed, the original inspiration for research on the bystander effect was a news report (largely apocryphal, it turns out) of the 1964 murder of a woman named Kitty Genovese in an apartment courtyard in full view of a large number of unresponsive witnesses (Manning, Levine, & Collins, 2007). This kind of physical arrangement—a public space that allows many people to see an event while seeing each other see the event—is a textbook example of the generation of common knowledge.

The only exception to the ubiquity of common knowledge in bystander research is a study by Barron and Yechiam (2002), which found that individuals are more likely to respond to an e-mail help request when they are the only one who receives the message compared with when they are carbon copied (cc'd) to-

[1] For example, the most recent review of the literature (Fischer et al., 2011) does not cite Diekmann, and mentions the volunteer's dilemma only in passing.

gether with other recipients. For present purposes, the single-recipient condition corresponds to private knowledge and the cc condition to common knowledge. In this paper, we extend this observation while going beyond this binary contrast, investigating how additional levels of recursive knowledge affect helping behavior.

Using Economic Games to Study the Bystander Effect

Following previous research (Diekmann, 1985, 1986), we use economic games to study the bystander effect. Participants interact in a volunteer's dilemma in which they decide whether to help or shirk, and they receive monetary payoffs depending on everyone's choices. Only the relative payoffs are relevant: Players earn the most money when they shirk and others help, but if others shirk then a player earns more money by helping. These payoffs capture the essence of a traditional bystander situation—multiple people could take the initiative to help but each person prefers that someone else assume the burden.

As in other research using economic games, our studies differ from real-world scenarios in which a bystander's help is needed, and from many situations studied in previous research, such as staged emergencies. For instance, when housemates help with chores, or nations send disaster relief, these helpers do not have simple monetary payoffs. Of course their decisions do have tangible consequences (i.e., payoffs), but these consequences are rarely monetary sums.

Even so, monetary payoffs offer key methodological advantages (Camerer, 2003; Kagel & Roth, 1995; Smith, 1982). Game theory models people's motivations by using numerical payoffs that summarize how they value different outcomes, whether those valuations are in fact based on money, altruism, reputation, emotions, or other motives. This generality allows game theory to bridge the social sciences because it uses a single framework to model motives as diverse as interpersonal helping, romantic love, economic exchange, and the pursuit of political power (e.g., Frank, 1988; Gintis, 2000). Testing these models experimentally requires well-defined payoffs; thus, using monetary payoffs to precisely recreate fundamental social interactions provides a key innovation for testing game-theoretic predictions in the lab (Camerer, 2003).

The critical issue is not whether experimental games are identical to everyday interactions but whether they create situations that fall within the domain of the theory under examination. Theories of social behavior generally do not provide principled exclusions for situations with monetary stakes, making these theories amenable to testing with economic games. For example, prominent theories of altruism, bargaining, group cooperation, and punishment include within their domain situations involving both monetary and nonmonetary consequences, implying that these theories can be tested in incentivized experiments. Indeed, experiments using the prisoner's dilemma, dictator game, ultimatum game, and public goods game have become common in psychology and have contributed many insights to our understanding of social behavior (e.g., Camerer, 2003; Kagel & Roth, 1995).

In the present case, theories about the bystander effect require a situation with multiple helpers whose decisions affect one another; the volunteer's dilemma provides a close fit to this domain, allowing incentive-controlled tests of bystander theories. Specifically, we seek to understand bystanders' decisions when they have asymmetric knowledge that help is needed. We use monetary payoffs to create a bystander situation with an uncertain need for help and then vary participants' knowledge about whether help is needed. We test whether different knowledge states affect participants' propensity to help, which is predicted by the strategic model, but not by the traditional diffusion of responsibility theory. Moreover, this experimental design precludes additional potential contributors to bystander inhibition, such as evaluation apprehension and social influence, allowing for a clean comparison of these competing predictions.

The Present Research

Participants interacted in groups of two or five in an incentivized volunteer's dilemma depicted in Figure 1, described as a role-playing scenario. Participants were told that the scenario was fictional, but the cash payoffs were real. In the scenario, the participants were merchants who rented stalls in a market from Mr. Smith, and they earned money each day they sold their goods at the market. The merchants were obligated to help Mr. Smith get supplies when he needed them, but only one volunteer's help was needed. If a participant helped Mr. Smith, it would cost them half their daily earnings, but if no one decided to help, then Mr. Smith would fine everyone a day's earnings, so that they all earned nothing.

Participants were told that Mr. Smith only occasionally needed help. Hence, the need for help was initially uncertain. This uncertainty allowed us to manipulate participants' levels of knowledge about what their partners knew about the need for help. All participants then learned that on that day Mr. Smith needed help, either from a loudspeaker that all merchants could hear—creating *common knowledge* among participants that help was needed—or from a messenger who delivered the information to them individually.

We manipulated asymmetric, recursive knowledge by varying the circumstances surrounding the delivery of the message. For *private knowledge,* participants simply received the message from the messenger, with no information about whether other people received the message too. For *secondary knowledge,* participants also knew that the message was delivered to their partner(s), but that their partner(s) did not know they knew this. For *tertiary knowledge,* participants knew their partner(s) knew they had received the message, but nothing more. For *quaternary knowledge* (Experiment 3), participants knew their partner(s) knew that they were aware that their partner(s) had received the message. Participants then decided whether they wanted to help Mr. Smith, and lose half their earnings for sure, or not help him, in which case they would collect their full earnings if someone else decided to help, or earn nothing if no one else did.

If the decision to help is strategic, then different knowledge levels will yield different probabilities of helping; we refer to this as the *knowledge-level hypothesis.* A second prediction is that individuals in the private knowledge condition will show the highest probabilities of helping in both group sizes, since they might be the only one in the group who knows that help is needed.

A third and less obvious prediction is that participants will zigzag between high and low probabilities of helping as levels of knowledge increase. This prediction is based on the following strategic logic. With private knowledge, a person will have a high probability of helping, since they might be the only one who knows help is needed. With secondary knowledge, a person will help with low probability, because they expect their counterpart (who has private knowledge) to help with high probability. And with tertiary knowledge, a person will help with high probability, because they expect their counterpart (who has secondary knowledge) to help with low probability. Theoretically, this zigzag pattern continues indefinitely for quaternary, quinary, senary, and higher levels of shared knowledge, because at each level a player helps with high or low probability depending on whether they expect players at the next level down to help with high or low probability.

A fourth prediction is that participants in the common knowledge condition will show the same group-size effect observed in previous research. Because all players know they have the same information, they have symmetric positions in an outguessing standoff, unlike the asymmetric positions they occupy with all lower levels of knowledge. Hence, participants will help with a higher probability in the 2-player condition than in the 5-player condition; moreover, these probabilities of helping will be intermediate between the high and low probabilities predicted for the lower levels of asymmetric, recursive knowledge.

In contrast, if the bystander effect results from a psychological process that does not take other bystanders' beliefs into account, there should be no differences in rates of helping across different knowledge levels. In other words, if it is indeed true that all that matters is one's knowledge that other bystanders are present (Latané & Nida, 1981), then rates of helping across the different knowledge levels should be similar, and responsibility should be equally diffused across all levels of beliefs. We refer to this as the *pure diffusion hypothesis*.

Experiment 1 compares helping rates for private and common knowledge, simply to test whether helping is sensitive to mental state information. Experiment 2 introduces two intermediate levels of shared knowledge, to test whether rates of helping strategically track the payoffs that go with the prevailing level of knowledge. Experiment 3 replicates the results from the previous experiments, but with a scenario involving more naturalistic cues to generate the different kinds of knowledge, and adds one more level of knowledge to confirm that the patterns of helping rates observed in Experiments 1 and 2 were not an artifact of testing a restricted range of knowledge conditions.

Experiment 1

Participants

We recruited 200 participants from the United States through Amazon's Mechanical Turk, an online labor crowdsourcing platform (see Buhrmester, Kwang, & Gosling, 2011; Goodman, Cryder, & Cheema, 2013; Ipeirotis, 2010; Paolacci, Chandler, & Ipeirotis, 2010; Berinsky, Huber, & Lenz, 2012). This sample size provides sufficient power (.8) to detect large effects (n per cell > 32 for $w > .5$), based on the large effect sizes observed in previous research on coordination decisions (Thomas et al., 2014). Following standard practice in such studies, we excluded from the analysis participants who gave any incorrect answers to comprehension questions about the scenario's payoff structure (these questions are provided in the online supplementary materials), yielding a final sample of $N = 159$ ($M_{\text{age}} = 30.5$, 28% female).

Procedure

Participants were randomly assigned to one of four conditions in a 2 (Knowledge: private vs. common) × 2 (Group size: five vs. two) between-subjects design.

Participants read that they would interact in groups of two or five people (counterbalanced, between-subjects) from Mechanical Turk. Participants read that they would earn 40 cents for completing the experiment and could earn more money based on the decisions they and their partner(s) made. They then read the fictional scenario described above, in which they and their partner(s) assumed the role of merchants who would earn $1.00 today for selling their goods at the market. They all worked for the same stall owner, Mr. Smith, who might require help. If he needed help and a participant decided to help, that participant earned only 50 cents because they lost half of the day's earnings while helping. Importantly, Mr. Smith only needed one of the merchants from their group to help, but if no one helped he would fine them all $1.00, and thus they would all earn nothing. These payoffs are typical for workers employed on Mechanical Turk (see, Horton, Rand, & Zeckhauser, 2011).

Participants then read that if Mr. Smith needed help they would receive this information either from a loudspeaker in the market that they could all hear, or from a messenger who delivers messages to them individually and privately. Then, they learned that Mr. Smith needed help in one of these two ways. In the *private knowledge* condition, they read: "The Messenger Boy has come to tell you that Mr. Smith needs help today. The Messenger Boy says that he has not seen the other merchant(s) yet, so you don't know if he (they) will also get the message." In the *common knowledge* condition, they read: "An announcement has been made on the loudspeaker that Mr. Smith needs help today, so both (all) of you know he needs help," which made it commonly known by all merchants.

Participants decided whether they wanted to help or not, explained their decision, and answered a series of comprehension questions about what they would earn based on the decision they and the other merchant(s) made and about what they and their partner(s) knew during the interaction. They then answered some basic demographic questions and were debriefed. After completing the study, participants were randomly assigned to groups according to knowledge level and group conditions in order to calculate their payoffs based on the decisions they and their other group members made. All participants were then paid their earnings through Mechanical Turk's bonus payment functionality.

All stimuli are provided in the online supplementary materials.

Results and Discussion

Figure 2 shows rates of helping. In the common knowledge condition, participants were less likely to help when there were five bystanders (45%) than two bystanders (73%), $\chi^2(1, N = 81) = 6.66$, $p = .010$, $\phi = .29$. This observation replicates the traditional bystander effect, which has almost invariably been studied under conditions in which the potential helpers have common knowledge. As such, it increases confidence that this exper-

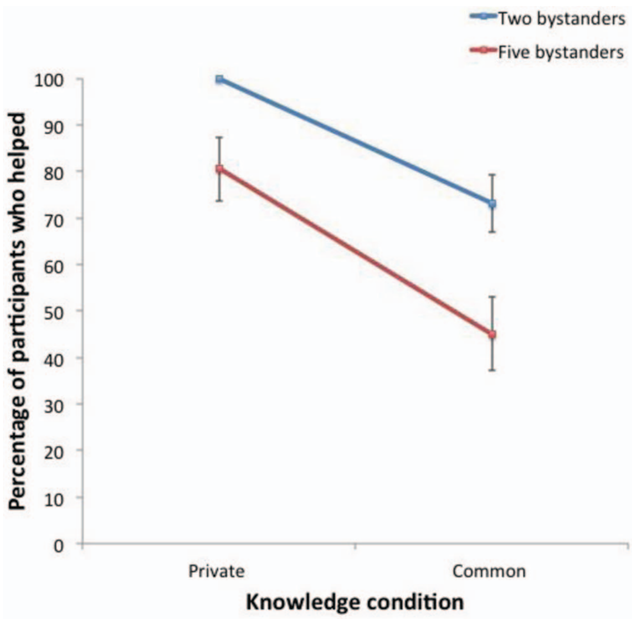

Figure 2. Percentage of participants who decided to help in Experiment 1, organized by knowledge condition, and group size. Error bars represent standard error. See the online article for the color version of this figure.

imental paradigm elicits the same diffusion of responsibility that has been observed in a variety of other situations, including in-person emergencies (e.g., Darley & Latané, 1968; Latané & Darley, 1968, 1970; Latané & Nida, 1981), in-person nonemergencies (e.g., Freeman, Walker, Borden, & Latané, 1975; Hurley & Allen, 1974; Levy et al., 1972), electronic interactions (e.g., Barron & Yechiam, 2002; Blair, Thompson, & Wuensch, 2005; Markey, 2000; Voelpel, Eckhoff, & Förster, 2008), and experimental economic games (Diekmann, 1986, 1993). In the private knowledge condition, participants were also less likely to help when there were five bystanders (80%) than two bystanders (100%), $\chi^2(1, N = 78) = 8.05$, $p = .006$, $\phi = .32$.

We next test the hypothesis that helping depends on knowledge-level. Participants were more likely to help when they had private knowledge than when they had common knowledge, both when there were two bystanders (100% vs. 73%), $\chi^2(1, N = 78) = 11.56$, $p = .001$, $\phi = .39$, and five bystanders (80% vs. 45%), $\chi^2(1, N = 81) = 10.93$, $p = .001$, $\phi = .37$.

These results show that a bystander's decision to help depends not only on what they know privately but also on what they know about other people's knowledge. When knowledge was private, helping rates were high even when there were five bystanders. This shows that the classic bystander effect depends on both the number of bystanders and common knowledge of the situation.

Experiment 2

Experiment 2 used the same methods as Experiment 1 but included two additional levels of knowledge, secondary and tertiary, to test whether the probability of helping zigzags as knowledge levels increase. The knowledge-level hypothesis predicts that within each group size, individuals with private and tertiary knowledge will show high probabilities of helping, individuals with secondary knowledge will show low probabilities of helping, and individuals with common knowledge will show intermediate probabilities of helping, according to the mixed strategy equilibrium. Specifically, with common knowledge, we expect the traditional bystander effect: A greater number of bystanders will make individuals less likely to help (Diekmann, 1985, 1986).

Participants

Exactly 1,200 participants from the United States, equally distributed across conditions, were recruited through Amazon's Mechanical Turk service. We chose a sample size that provides sufficient power (.8) to detect medium effect sizes (n per cell > 88 for $w > .3$) because of the medium effect sizes observed in Experiment 1. Participants who failed comprehension questions about payoffs were excluded from analyses, yielding a final sample of $N = 914$ ($M_{age} = 32.9$, 40% female).

Procedure

The procedure was the same as in Experiment 1, with two additional knowledge conditions in each group size condition, yielding a 4 (Knowledge: private, secondary, tertiary, common) × 2 (Group size: two vs. five) between-subjects design. In the *secondary knowledge* condition, participants read: "The Messenger Boy has come to tell you that Mr. Smith needs help today. The Messenger Boy says that he stopped by the other stall (four stalls) before coming to see you. He tells you that the other (four) merchant(s) knows (know) that Mr. Smith needs help today. However, he says that he forgot to mention to the other merchant(s) that he was coming to see you, so the other merchant(s) is (are) not aware that you know Mr. Smith needs help today." In the *tertiary knowledge* condition, participants read: "The Messenger Boy has come to tell you that Mr. Smith needs help today. The Messenger Boy mentions that he is also heading over to see the other (four) merchant(s), and will let them know Mr. Smith needs help today. The messenger boy will also tell the other (four) merchant(s) that he just came from your stall and told you about Mr. Smith's request. However, the messenger boy will not inform the other merchant(s) that he told you he would be heading over there. So, while the other merchant(s) is (are) aware that you know Mr. Smith needs help today, they are not aware that you know that they know that." All other aspects of the procedure were the same as in Experiment 1.

Results and Discussion

Figure 3 shows rates of helping. In the common knowledge condition, participants were less likely to help when there were five bystanders (44%) than two bystanders (77%), replicating the classic diffusion of responsibility effect, $\chi^2(1, N = 222) = 25.06$, $p < .001$, $\phi = .34$. In the private knowledge condition, participants were also less likely to help when there were five bystanders (75%) than two bystanders (91%), $\chi^2(1, N = 227) = 10.64$, $p = .001$, $\phi = .22$. This difference in helping rates was also observed in the tertiary knowledge condition (63% vs. 76%), $\chi^2(1, N = 229) = 4.74$, $p = .029$, $\phi = .14$, but not in the secondary knowledge condition (46% vs. 40%), $\chi^2(1, N = 236) = 0.78$, $p = .377$, $\phi = .06$.

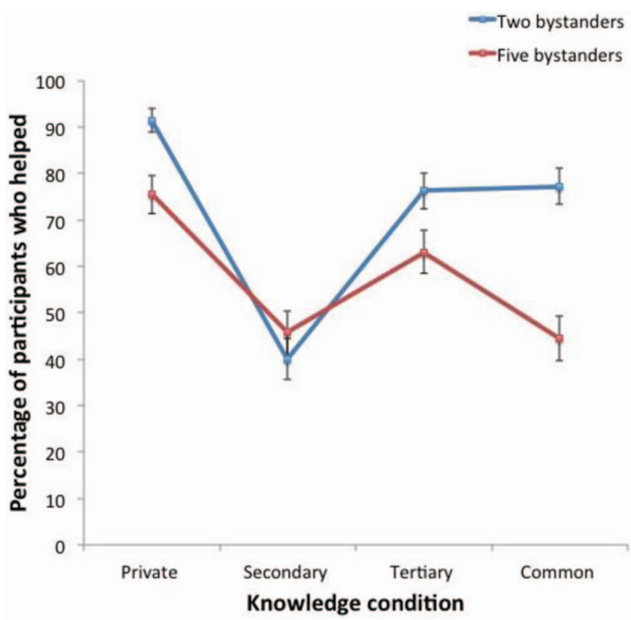

Figure 3. Percentage of participants who decided to help in Experiment 2, organized by knowledge condition, and group size. Error bars represent standard error. See the online article for the color version of this figure.

We next test the hypothesis that helping depends on knowledge. The knowledge-level hypothesis predicts a zigzag pattern of helping probability across the asymmetric levels of knowledge (private, secondary, tertiary). To test this prediction, we analyzed the probability of helping using logistic regression models with both linear and quadratic effects for knowledge-level (excluding common knowledge). In the two-bystander condition, the linear and quadratic effects were significant (all z values > 8.0, p values $<$ .001), consistent with the predicted zigzag pattern. Moreover, a likelihood ratio test showed that the quadratic model outperformed a purely linear model, $\chi^2(1, N = 355) = 74.07$, $p < .001$. Similarly, in the five-bystander condition, the linear and quadratic effects were significant (all z values > 4.2, p values $< .001$), with the quadratic model outperforming a purely linear model, $\chi^2(1, N = 337) = 18.05$, $p < .001$.

This pattern of results shows that participants responded to the strategic implications of different levels of knowledge, not just to their sheer complexity. Participants with private knowledge tend to help, those with secondary knowledge expect their partners to help and so choose to refrain, and those with tertiary knowledge expect their partners (who have secondary knowledge) to refrain and so choose to help. This zigzag pattern of help shows that participants engaged in strategic mental state reasoning, rather than simply increasing or decreasing their probability of helping as knowledge levels increased. Finally, participants with common knowledge showed the classic bystander effect, consistent with the game's mixed strategy equilibrium: Participants in groups of two were more likely to help than participants in groups of five.

These interpretations are, however, limited by the fact that we tested only two intermediate levels of knowledge between private and common knowledge, and one may wonder whether this pattern would continue across higher levels of shared knowledge. Experiment 3 was designed to replicate and extend these results by adding another level of knowledge—quaternary knowledge—in order to confirm that the observed pattern is genuinely strategic rather than responding to some specific difference between secondary and tertiary knowledge.

Experiment 3

Experiment 3 was designed to conceptually replicate the results of Experiments 1 and 2 with more naturalistic scenarios, and included one additional knowledge level. The participants ascertained the different knowledge levels through the kinds of body language and gaze direction cues that people normally encounter when assessing who knows what in a natural situation, as opposed to being told about knowledge levels explicitly in the scenarios.

Experiment 3 also tested the effects of quaternary knowledge. The knowledge-level hypothesis predicts that the probability of helping when bystanders have quaternary knowledge should resemble the probability of helping with secondary knowledge. This would rule out the possibility that the zigzag pattern observed in Experiment 2 was due to people treating tertiary knowledge as a vague category in between secondary and common knowledge, as opposed to strategically calibrating their behavior to the prevailing knowledge level.

Participants

Exactly 1,400 participants from the United States, equally distributed across conditions, were recruited through Amazon's Mechanical Turk service. We chose a sample size that provides sufficient power (.8) to detect medium effect sizes (n per cell > 88 for $w > .3$) because of the effect sizes observed in Experiments 1 and 2. Participants who failed comprehension questions about payoffs were excluded from analyses, yielding a final sample of $N = 1005$ ($M_{age} = 34.6$, 42% female).

Procedure

In order to convey knowledge levels more naturalistically, we told participants that the messenger delivered his messages in an easily identifiable pink envelope, rather than explicitly telling them who knew what, as in Experiments 1 and 2.

In the *private knowledge* condition, participants read: "A messenger comes by with a message in a pink envelope. You open the envelope, read that Mr. Smith needs help today, and throw the note and envelope away before anyone else sees it. The messenger says that he hasn't seen the other merchant(s) today, so it is unlikely that they know Mr. Smith needs help."

In the *secondary knowledge* condition, they read: "A messenger comes by with a message in a pink envelope. You open the envelope, read that Mr. Smith needs help today, and throw the note and envelope away before anyone else sees it. A few minutes later, as you are walking around the market, you pass by the other merchant's stall (all four of the other merchants' stalls). You see the messenger's pink envelope on their stall counter (all four of their stall counters), so you know they got the message as well. However, you don't get a chance to speak with (any of) them, so they don't know (none of them know) you've seen their envelope."

In the *tertiary knowledge* condition, they read: "A messenger comes by with a message in a pink envelope. Right then you get swamped by a large group of customers, and just leave the envelope sitting on your counter while you attend to them. As you are helping your customers, you see the other merchant walk by and see the pink envelope on your counter (each of the other four merchants walk by one at a time, and all four see the pink envelope on your counter). So the other merchant (each of the four other merchants) must know that you received the message that Mr. Smith needs help. However, the other merchant doesn't notice (none of the other four merchants notice) that you saw them pass by, and is (they are all) gone by the time you finish with your customers."

In the *quaternary knowledge* condition, they read: "A messenger comes by with a message in a pink envelope. You open the envelope, read that Mr. Smith needs help today, and throw the note and envelope away before anyone else sees it. Later that day, as you are walking with a colleague from one stall to the next, you see the messenger's pink envelope on the other merchant's counter (each of the other four merchants' counters), so you know that they (all of them) got the message as well. (In each case), the other merchant sees you look directly at the envelope but when they try to get your attention, you look away and start speaking to your colleague. So the other merchant sees (all four merchants see) you look at their envelope, but thinks (each thinks) that you didn't notice that they saw you."

All other aspects of the procedure were the same as in Experiments 1 and 2.

Results and Discussion

Figure 4 shows rates of helping. Replicating the previous two experiments and the classical bystander effect, in the common

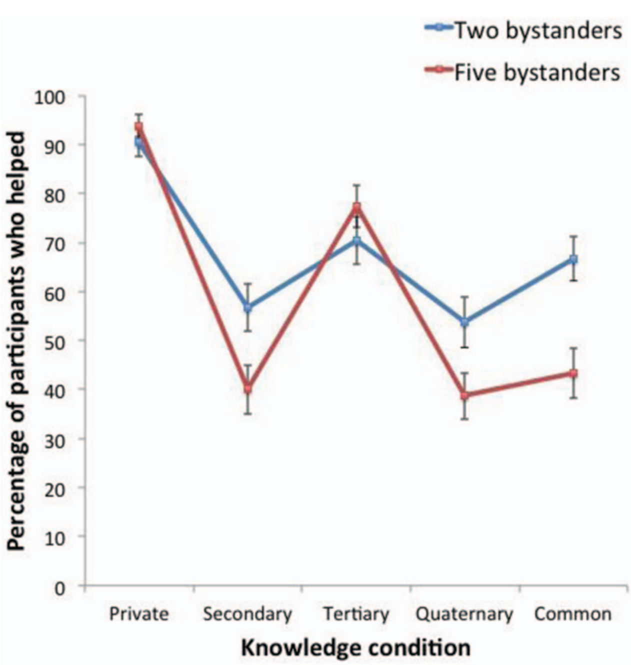

Figure 4. Percentage of participants who decided to help in Experiment 3, organized by knowledge condition, and group size. Error bars represent standard error. See the online article for the color version of this figure.

knowledge condition participants were less likely to help when there were five bystanders (44%) than when there were two bystanders (77%), $\chi^2(1, N = 208) = 11.46$, $p = .001$, $\phi = .24$. However, unlike in the other two experiments, in the private knowledge condition participants were no more likely to help when there were two bystanders (90%) than five bystanders (94%), $\chi^2(1, N = 216) = 0.84$, $p = .359$, $\phi = .06$. Furthermore, unlike Experiment 2, participants were less likely to help when there were five bystanders (40%) than two bystanders (57%) in the secondary knowledge condition, $\chi^2(1, N = 199) = 5.56$, $p = .018$, $\phi = .17$, and there was no significant difference in helping rates in the tertiary knowledge condition (70% vs. 77%), $\chi^2(1, N = 181) = 1.14$, $p = .285$, $\phi = .08$. In the quaternary knowledge condition, participants were less likely to help when there were five bystanders (39%) than two bystanders (54%), $\chi^2(1, N = 201) = 7.14$, $p = .008$, $\phi = .18$.

We next test the hypothesis that helping depends on knowledge. To test for the predicted zigzag pattern, we analyzed helping by using logistic regression with linear, quadratic, and cubic effects of asymmetric knowledge levels (private, secondary, tertiary, and quaternary). In the two-bystander condition, the linear, quadratic, and cubic effects were significant (all z values > 3.8, p values $< .001$), confirming a significant zigzag pattern across the four levels. Furthermore, likelihood ratio tests showed that the cubic model outperformed both the quadratic model, $\chi^2(1, N = 391) = 15.68$, $p < .001$, and a purely linear model, $\chi^2(1, N = 391) = 21.84$, $p < .001$. Similarly, in the five-bystander condition, the linear, quadratic, and cubic effects were significant (all z values > 7.5, p values $< .001$), with the cubic model outperforming both the quadratic model, $\chi^2(1, N = 406) = 65.13$, $p < .001$, and a purely linear model, $\chi^2(1, N = 406) = 70.24$, $p < .001$.

This pattern of results again shows that participants responded strategically to different levels of knowledge. As in Experiment 2, participants with private knowledge tended to help, those with secondary knowledge refrained, and those with tertiary knowledge (whose partners had secondary knowledge, and thus could be expected to refrain from helping) also tended to help. Adding to this pattern, Experiment 3 showed that those with quaternary knowledge (whose partners had tertiary knowledge, and thus could be expected to help) refrained from helping. The continuation of this zigzag pattern to a further level of knowledge provides robust support for the hypothesis that participants engaged in strategic mental state reasoning when making their decisions.

However, another aspect of these results was inconsistent with those of Experiments 1 and 2. Unlike those experiments, there was no difference here in the private knowledge condition between helping rates with five bystanders and with two bystanders. Furthermore, unlike Experiment 2, there was a group-size effect in the secondary knowledge condition but not in the tertiary knowledge condition. In contrast to these inconsistent effects with lower, asymmetric levels of knowledge, when participants had common knowledge the difference in helping rates between the five-bystander and two-bystander conditions was not only consistent across experiments but also consistently showed the largest difference in helping rates across the two group sizes in all three experiments. In other words, the classic bystander effect is robust only when bystanders have common knowledge that help is needed.

General Discussion and Conclusion

These experiments reveal how people's decisions to help can be influenced not only by the number of other bystanders but also by the level of knowledge that bystanders have about each other. Furthermore, the zigzag pattern in probabilities of helping shows how people can strategically respond to asymmetric, recursive knowledge about the need for help.

These results expose a previously unrecognized but important variable in decisions to help, namely knowledge about what other bystanders know, which is predicted by strategic models but is not well explained by the traditional diffusion of responsibility theory. Probabilities of helping were markedly different depending on the level of knowledge participants had, and the signature bystander effect was most consistent and apparent when participants had common knowledge that help was needed. Minimally, these observations add a new psychological mechanism—recursive mentalizing—to our understanding bystander decisions. Furthermore, the fact that a strategic analysis correctly predicted these novel and nuanced effects *together with* the group-size effects predicted by the diffusion of responsibility theory (Diekmann, 1985, 1986) suggests that the strategic model can provide a unified explanation for bystander inhibition.

According to the strategic account, responsibility does not simply diffuse indiscriminately across bystanders; hence, the term "diffusion of responsibility" is something of a misnomer. Instead, bystanders' decisions take others' mental states into account, leading people to help or not based on their beliefs about what other bystanders know. These results thus demonstrate that the bystander effect has a major strategic component, rather than being an irrational anomaly in our moral psychology, and that it is driven by mentalizing (indeed, recursive mentalizing) about the knowledge of the other bystanders, rather than only being sensitive to their presence or absence.

Of course most behavior, including bystander helping, is caused by multiple mechanisms. Strategic reasoning operates in concert with other mechanisms, including *social influence,* in which other bystanders' reactions influence how one interprets an ambiguous situation (Latané & Nida, 1981), *evaluation apprehension,* in which an individual is reluctant to act out of a fear of being judged for responding in an inappropriate manner (Latané & Nida, 1981), and *confusion of responsibility,* in which an individual worries about being mistaken for a perpetrator (Cacioppo, Petty, & Losch, 1986). Our results offer no insight into these contributors to bystanders' decisions because the experimental design precludes them: participants could not observe other bystanders' decisions; they were anonymous; and helping could not be associated with harming a victim. However, our methods for manipulating bystanders' knowledge levels could be employed to test these mechanisms in future research. For example, private information about how others react to an ambiguous situation should be sufficient for informing one's interpretation of the situation; additional levels of shared knowledge should have no effect on this process. Similarly, knowing that other bystanders will not know if one intervenes should be sufficient to rule out any concern that one might be mistaken for a perpetrator.

Though the method of economic games has many advantages, it might also have had unintended effects on participants' behavior. Conceivably, presenting participants with explicit monetary payoffs might have elicited strategic reasoning that they would not use if the payoffs had been denominated in money, food, time, effort, reputation, or some other currency of psychological valuation. One indication that strategic reasoning cuts across all these currencies is that in the common knowledge condition of all three experiments, we replicated the group-size effect, which has been found in previous studies using diverse scenarios and methods. The generality of strategic reasoning can be tested directly in future work by employing volunteer's dilemmas that use well controlled nonmonetary outcomes, such as minutes spent on a tedious task or points awarded in a public competition.

Finally, the critical role of asymmetric knowledge can inform efforts to encourage or discourage helping. For instance, people who e-mail requests for help can strategically use "cc" and "bcc" options to create the recursive knowledge states that promote helping. People who receive requests can conceal evidence that they got the message. Requesters who use physical letters can deliberately place them in conspicuous or inconspicuous locations to generate or avoid common knowledge. Witnesses to an emergency can feign ignorance to encourage other bystanders to help instead, or they can make it clear that they know another bystander has also witnessed the emergency. Charities can use recursive knowledge states to convey a donor's special obligation to help those in need. Governments can publicize or quiet political crises to affect citizens' ability to act collectively. In short, wherever help is needed, bystanders' knowledge and strategies can determine whether a group achieves its goals.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116,* 953–970. http://dx.doi.org/10.1037/a0016923

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press.

Barron, G., & Yechiam, E. (2002). Private e-mail requests and the diffusion of responsibility. *Computers in Human Behavior, 18,* 506–520. Retrieved from http://www.sciencedirect.com/science/article/pii/S0747563202000079#

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20,* 351–368. http://dx.doi.org/10.1093/pan/mpr057

Blair, C. A., Thompson, L. F., & Wuensch, K. L. (2005). Electronic helping behavior: The virtual presence of others makes a difference. *Basic and Applied Social Psychology, 27,* 171–178. http://dx.doi.org/10.1207/s15324834basp2702_8

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5. http://dx.doi.org/10.1177/1745691610393980

Cacioppo, J. T., Petty, R. E., & Losch, M. E. (1986). Attributions of responsibility of helping and doing harm: Evidence for confusion of responsibility. *Journal of Personality and Social Psychology, 50,* 100–105. http://dx.doi.org/10.1037/0022-3514.50.1.100

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Chwe, M. S. Y. (2001). *Rational ritual: Culture, coordination, and common knowledge*. Princeton, NJ: Princeton University Press.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8,* 377–383. http://dx.doi.org/10.1037/h0025589

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution, 29,* 605–610. http://dx.doi.org/10.1177/0022002785029004003

Diekmann, A. (1986). Volunteer's dilemma: A "social trap" without dominant strategy and some experimental results. In A. Diekmann & P. Mitter (Eds.), *Paradoxical effects of social behavior. Essays in honor of Anatol Rapoport* (pp. 187–197). Heidelberg: Physica-Verlag. http://dx.doi.org/10.1007/978-3-642-95874-8_13

Diekmann, A. (1993). Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence. *International Journal of Game Theory, 22,* 75–85. http://dx.doi.org/10.1007/BF01245571

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., . . . Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137,* 517–537. http://dx.doi.org/10.1037/a0023304

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York, NY: Norton.

Freeman, S., Walker, M. R., Borden, R., & Latané, B. (1975). Diffusion of responsibility and restaurant tipping: Cheaper by the bunch. *Personality and Social Psychology Bulletin, 1,* 584–587. http://dx.doi.org/10.1177/014616727500100407

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 358,* 459–473. http://dx.doi.org/10.1098/rstb.2002.1218

Gintis, H. (2000). *Game theory evolving: A problem-centered introduction to modeling strategic interaction*. Princeton, NJ: Princeton University Press.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26,* 213–224. http://dx.doi.org/10.1002/bdm.1753

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 14,* 399–425. http://dx.doi.org/10.1007/s10683-011-9273-9

Hurley, D., & Allen, B. P. (1974). The effect of the number of people present in a nonemergency situation. *The Journal of Social Psychology, 92,* 27–29. http://dx.doi.org/10.1080/00224545.1974.9923068

Ipeirotis, P. (2010, March 9). The new demographics of Mechanical Turk. Retrieved from http://www.behind-the-enemy-lines.com/2010/03/new-demographics-of-mechanical-turk.html

Kagel, J. H., & Roth, A. E. (Eds.). (1995). *The handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology, 10,* 215–221. http://dx.doi.org/10.1037/h0026570

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York, NY: Appleton-Century-Croft.

Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin, 89,* 308–324. http://dx.doi.org/10.1037/0033-2909.89.2.308

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review, 117,* 785–807. http://dx.doi.org/10.1037/a0019688

Levy, P., Lundgren, D., Ansel, M., Fell, D., Fink, B., & McGrath, J. E. (1972). Bystander effect in a demand-without-threat situation. *Journal of Personality and Social Psychology, 24,* 166–171. http://dx.doi.org/10.1037/h0033380

Manning, R., Levine, M., & Collins, A. (2007). The Kitty Genovese murder and the social psychology of helping: The parable of the 38 witnesses. *American Psychologist, 62,* 555–562. http://dx.doi.org/10.1037/0003-066X.62.6.555

Markey, P. M. (2000). Bystander intervention in computer-mediated communication. *Computers in Human Behavior, 16,* 183–188. http://dx.doi.org/10.1016/S0747-5632(99)00056-4

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York, NY: Viking.

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 833–838. http://dx.doi.org/10.1073/pnas.0707192105

Saxe, R., & Young, L. (2013). Theory of mind: How brains think about thoughts. In K. Ochsner & S. Kosslyn (Eds.), *The Handbook of Cognitive Neuroscience* (pp. 204–213). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199988709.013.0013

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review, 72,* 923–955. http://www.jstor.org/stable/1812014

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

Voelpel, S. C., Eckhoff, R. A., & Förster, J. (2008). David against Goliath? Group size and bystander effects in virtual knowledge sharing. *Human Relations, 61,* 271–295. http://dx.doi.org/10.1177/0018726707087787

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103–128. http://dx.doi.org/10.1016/0010-0277(83)90004-5

Received June 30, 2015
Revision received December 9, 2015
Accepted January 15, 2016 ■