

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


The psychology of coordination and common knowledge.

Citation

Thomas, Kyle A., Peter DeScioli, Omar Sultan Haque, and Steven Pinker. 2014. “The Psychology of Coordination and Common Knowledge.” Journal of Personality and Social Psychology 107 (4): 657–676. doi:10.1037/a0037037.

Published version

https://doi.org/10.1037/a0037037

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:14330738

Terms of use

This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles (OAP), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

The Psychology of Common Knowledge and Coordination

Kyle A. Thomas[1], Peter DeScioli[1,2], Omar Sultan Haque[1], Steven Pinker[1]

[1]Department of Psychology, Harvard University

[2]Department of Political Science, Stony Brook University

Author Note

Kyle A. Thomas, Omar Sultan Haque, & Steven Pinker, Department of Psychology, Harvard University; Peter DeScioli, Department of Psychology, Harvard University and Department of Political Science, Stony Brook University.

We thank Moshe Hoffman for providing feedback on the manuscript, and Cheng Li, Pooja Ami Patel, Natalie Aharon, and Sara Paul for helping with data collection and analysis. This work was first presented at the 23rd annual meeting of the Human Behavior and Evolution Society in Albuquerque, New Mexico.

Correspondence concerning this article should be addressed to Kyle Thomas, Department of Psychology, Harvard University, William James Hall 964, 33 Kirkland Street, Cambridge, MA, 02138. Email: kathomas@fas.harvard.edu

Abstract

Research on human cooperation has concentrated on the puzzle of altruism, in which one actor incurs a cost to benefit another, and the psychology of reciprocity, which evolved to solve this problem. We examine the complementary puzzle of mutualism, in which actors can benefit each other simultaneously, and the psychology of coordination, which ensures such benefits. Coordination is facilitated by *common knowledge*—the recursive belief state in which A knows X, B knows X, A knows that B knows X, B knows that A knows X, ad infinitum. We test whether people are sensitive to common knowledge when deciding whether to engage in risky coordination. Participants decided between working alone for a certain profit and working together for a potentially higher profit that they would receive only if their partner made the same choice. Results showed that more participants attempted risky coordination when they and their prospective partner had common knowledge of the payoffs (broadcasted over a loudspeaker) than when they had only shared knowledge (conveyed to both by a messenger) or primary knowledge (revealed to each partner separately). These results confirm the hypothesis that people represent common knowledge as a distinct cognitive category that licenses them to coordinate with others for mutual gain. We discuss how this hypothesis can provide a unified explanation for diverse phenomena in human social life, including recursive mentalizing, performative speech acts, public assemblies and protests, and self-conscious emotional expressions.

Keywords: common knowledge, coordination, theory of mind, cooperation, mutualism, stag hunt

The Psychology of Common Knowledge and Coordination

A strange and ethereal protest took place in Belarus during the summer of 2011, consisting solely of protesters' phones ringing simultaneously. Police swarmed the event, recorded who was there, and made aggressive arrests (Barry, 2011). What were the protesters trying to accomplish? And why were the police concerned with such a seemingly trivial event?

People interact in a variety of situations in which they need to coordinate their actions to achieve common goals, such as opposing unfair governments, capturing gains in trade, agreeing on the use of standard symbols and protocols, and countless everyday activities such as scheduling meetings, contributing to potluck dinners, and carrying two ends of a heavy object. Because it is costly to engage in a coordinated activity when no one else does so, attempts to coordinate can be risky when it is unclear what other people will do. In repressive regimes a single protestor risks prosecution and violence, a risk which can be mitigated only by overwhelming numbers of people successfully coordinating their actions: If one protestor shows up he gets shot, if a million show up they may send the dictator packing. In these situations, even modest displays of synchrony, such as simultaneous phone rings, can set the stage for larger-scale coordination. However, even when it's clear that other people want to work together, coordination can be a challenge. Exactly how, for instance, do thousands of would-be protestors converge on a single time and place to voice their concerns?

Coordination problems are a subtopic in the psychology of cooperation. Though cooperation has become a burgeoning area in psychology, economics, and evolutionary biology, research and theory have concentrated on the subtype of cooperation that is *altruistic* (in the biological sense): A cooperator confers a benefit on a partner at a cost to himself. Altruistic cooperation has received the lion's share of attention because it raises the evolutionary puzzle of

how a behavior that harms the actor could be selected for. The paradox is often captured in the game-theoretic scenario of the Prisoner's Dilemma, and the challenge to the psychologist is in characterizing the cognitive abilities and emotional motives that allow humans to surmount it. These include the ability to recognize individuals and detect cheaters, and a suite of emotions that police reciprocation, such as sympathy, anger, gratitude, forgiveness, guilt, and trust (Trivers, 1971; Cosmides & Tooby, 1992, 2005).

Coordination, in contrast, is *mutualistic*: Each cooperator confers a benefit on the other while simultaneously conferring a benefit on himself or herself. Despite this convergence of interests, coordination, too, poses an evolutionary challenge. The challenge is not motivational though, but epistemological: accurately representing the other actor's state of knowledge. The epistemological problem results from the difficulty of converging on a single solution when more than one is available. For instance, two friends both benefit if they meet at Starbucks, or at Peet's, but for this to happen each friend has to know that the other knows which location they have agreed upon.

If this problem can be resolved, the incentives of the game pose no further obstacle, and can even help guide optimal behavior rather than hinder it (Lewis, 1969; Schelling, 1960; Skyrms, 2004). The paradigm game-theoretic model of a coordination problem is the Stag Hunt, first introduced by the philosopher Jean-Jacques Rousseau (Rousseau, 1754/1984; Skyrms, 2004). In the Stag Hunt, two hunters can set out in the morning either to hunt stag together (a large payoff) or to hunt rabbits separately (a small payoff); a single hunter cannot fell a stag and will return empty-handed (a high opportunity cost). To attain the highest payoff, each hunter must not only know that stag offers higher payoffs, but they must also know that the other hunter knows the payoffs, know that the other hunter knows that they know the payoffs, and so on.

Yet despite this epistemological problem, humans are adept at achieving coordination. Protestors meet up in Tahrir Square at 5pm on Friday, different suppliers produce the parts for a complex product, allied battalions converge on an enemy, diners use the bread plate to the left, coworkers in a building settle on an informal name for a meeting space. Given a long evolutionary history of group living, human cognition may have been shaped by natural selection to solve coordination problems (Tooby & Cosmides, 2010; Tooby, Cosmides, & Price, 2006). If game theorists are correct that common knowledge is needed for coordination, then humans might have cognitive mechanisms for recognizing it.

This paper attempts to begin to redress the imbalance in the literature on the psychology of cooperation by exploring the epistemological challenges and the possible cognitive and motivational adaptations surrounding the problem of mutualistic coordination.[1] We focus on a special kind of representation called *common knowledge* (sometimes called *mutual knowledge* or *common ground*; Clark & Marshall, 1981; Clark, 1996; Lewis, 1969; Pinker, 2007; Rubinstein, 1989; Schelling, 1960; Smith, 1982). Common knowledge is defined as an infinite string of embedded levels of mutual knowledge, i.e., Michael knows *X*; Lisa knows *X*; Michael knows that Lisa knows *X*; Lisa knows that Michael knows *X*; Michael knows that Lisa knows that Michael knows *X*; ad infinitum.

The infinite levels of knowledge required for common knowledge may seem to present a different kind of epistemological problem, namely that a finite mind cannot represent an infinite set of nested propositions. However, people need not represent each level of knowledge explicitly, but could simply represent a recursive formula that entails all levels of knowledge,

[1] A PsychInfo search reveals that in the years 1992-2013, 1,264 papers listed "altruism" as a major subject heading or keyword, whereas only 54 listed "mutualism" (and most of these were for studies of nonhuman animals). There were 321 references to the Prisoners' Dilemma, but only 3 to the Stag Hunt.

such as *Y* = "Everyone knows *X,* and everyone knows *Y*", or even just a single symbol that indicates the state of common knowledge itself (Clark, 1996; Pinker, 2007). This formula or symbol, moreover, can be activated in people's minds by any salient public signal which reliably causes the knowledge, such as a message broadcasted on a loudspeaker: Everyone who receives the signal knows that everyone else has received it, and can deduce that everyone else can deduce that, ad infinitum (Aumann, 1976).

Nor is it necessary that that the commonly entertained propositions be known with absolute certainty. Coordination may be achieved with the weaker notion of *common belief*, in which two agents each believe that a proposition is likely to be true with probability at least *p*, each believes that the other believes it with probability at least *p*, and so on (Monderer & Samet, 1989). For any situation with a stag-hunt payoff structure, there is a minimum level of *p* whose value depends on the relative advantage of coordination over acting alone, for which it is rational for agents with *common p-belief* to choose to coordinate (Dalkiran, Hoffman, Paturi, Ricketts, & Vattani, 2012). In the rest of this paper, we will use the term *common knowledge* broadly, to include "sufficiently high common *p*-belief".

Common knowledge can be contrasted with what we will refer to as *shared knowledge,* any string of embedded levels of knowledge that falls short of infinity, and with *primary knowledge,* knowledge that individuals possess without knowing whether anyone else possesses it. Common knowledge is intimately connected with the logical problem of coordination; in theory, coordination can be irrational without it. With the help of three experiments in which participants are given the opportunity to engage in a simple form of economic cooperation, we examine the extent to which people really do depend on common knowledge and other forms of knowledge to achieve coordination.

The Game Theory of Coordination and Common Knowledge

Research in game theory on coordination games shows why shared knowledge may be insufficient for coordination. Technically, coordination games are situations of interdependent decision-making that have multiple equilibria. Conceptually, they are situations in which two or more people each make a decision, with the potential to achieve mutual benefits only if their decisions are consistent (Lewis, 1969; Schelling, 1960). The rendezvous example is a coordination game because both friends benefit from choosing the same location, but that location could be either Starbucks or Peet's. To choose among multiple solutions an individual must take into account what she expects the other actor to do. However, what another actor is likely to do is in turn dependent upon his expectations of what she will do, leading to interdependent expectations that generate an infinite recursion of embedded beliefs.

A classic paper demonstrated the importance of common knowledge for maximizing payoffs from a coordination game, and showed how anything less than the infinite levels of knowledge that common knowledge entails may be insufficient (Rubinstein, 1989). Rubinstein's model showed that under a specific, restrictive set of assumptions any level of knowledge short of common knowledge is no better than no knowledge at all. Subsequent work has suggested that this conclusion was too strong, and that shared knowledge or less-than-certain beliefs can enable coordination better than primary knowledge (Binmore & Samuelson, 2001; Dalkiran, Hoffman, Paturi, Ricketts, & Vattani, 2012; Monderer & Samet, 1989). However, even in these models, common knowledge has a privileged role to play in facilitating coordination, in part because it avoids a second-order coordination problem presented by shared knowledge. With shared knowledge people must decide how many levels of shared knowledge is enough to attempt coordination: How can individuals be certain that everyone requires the same number of levels of

shared knowledge to attempt risky coordination? In short, all of these models demonstrate that common knowledge provides the most effective and reliable path to coordination.

The problem of coordination and common knowledge has been examined by many disciplines, including political science (Ostrom, 1990), philosophy (Hume, 1739-1740/1969; Rousseau, 1754/1984; Lewis, 1969; Skyrms, 2004), economics (Chwe, 2001; Geanakoplos, 1992); linguistics (Clark, 1992, 1996; Smith, 1982), sociology (Willer, Kubuwara, & Macy, 2009), and even computer science (Alberucci & Jäger, 2005). Yet despite the fact that common knowledge is fundamentally a psychological phenomenon, little is known about the psychology of common knowledge (some notable exceptions include Chaudhuri, Schotter, & Sopher, 2009; Lee & Pinker, 2010). We briefly review two literatures (experimental economics and theory of mind) that are indirectly relevant to the phenomenon before outlining our own research questions.

Experimental Economics: Coordination Using Salient Focal Points

A few experiments have examined whether people are better at solving coordination problems than classical game theory suggests. They focus on Schelling's (1960) concept of a *focal point,* an option that stands out of a set of possible choices as uniquely salient, encouraging everyone to converge upon it as a single choice. Schelling suggested that in practice people may rely on focal points to solve coordination problems because they generate common knowledge of a single solution (Schelling, 1960; Sugden, 1995). Mehta, Starmer, and Sugden (1994a, 1994b) examined people's play in coordination games and their ability to converge on focal points (what they called "Schelling salience"). Participants responded to questions with many possible answers (e.g., "Write down any positive number," and "Name any flower"). In one group, participants were paid to answer any way they wanted. In another, they were paid based on how

well their answers matched with those of another randomly chosen participant. Participants were far more successful at coordinating answers when they were trying to do so than when they answered as they wished. This suggests that people can meet the challenge of coordination by identifying it as a problem distinct from the primary demands of a task. Though the finding, by itself, cannot distinguish whether people used shared knowledge or common knowledge to improve their coordination, recent unpublished studies suggest that people really do use common knowledge in these tasks (Bardsley, Mehta, Starmer, & Sugden, 2008; Chartier, Abele, Stasser, & Shriver, 2012).

Theory of Mind Research: Representing Shared Knowledge

Most existing research on knowledge about other people's knowledge falls in the area known as Theory of Mind, intuitive psychology, mind-reading, or mentalizing, all terms for the mental representation of other people's mental states (Baron Cohen, 1995; Frith & Frith, 2003; Wimmer & Perner, 1983; for recent reviews, see Apperly & Butterfill, 2009; Saxe & Young, in press). Developmental psychologists have found that by 6-7 months children are able to use implicit representations of attention, desires, goals, and intentions to guide their behavior (Hamlin, Hallinan, & Woodward, 2008). By fifteen months, children can (implicitly) differentiate their own knowledge from another person's knowledge; for example, infants are surprised when someone seeks out an object in a spot where it was moved when the person was absent (Onishi & Baillargeon, 2008). By 3-5 years, children show an ability to explicitly represent others' mental states in the false-belief task (Callaghan et al., 2005; Wellman et al., 2001). By 6-7 years, children are able to represent two levels of shared knowledge, as evidenced by their ability to understand that someone else can have false beliefs (Perner & Wimmer, 1985). By adulthood, people can correctly answer questions about fourth-order levels of shared

knowledge (e.g., Bob knows that Carol knows that Ted knows that Alice knows X), but they tend to fail questions about fifth-order knowledge (Kinderman, Dunbar, & Bentall, 1998), possibly because this exceeds the capacity of short-term verbal memory (Cowan, 2000).

Although people are capable of representing other people's mental states, they do not always do so effectively. Both adults and children tend to assume that their knowledge is shared by other people. This shortcoming is evident in the well-documented failure of three-year-olds to pass a false-belief task, and is also seen in adults in work on the *curse of knowledge* (Birch & Bloom, 2003, 2007; Camerer, Lowenstein, & Weber, 1989).

Since coordination depends on the ability to anticipate other people's actions, and since people's actions depend on their mental states, one would expect mentalizing ability to facilitate coordination. Indeed, Curry & Jones Chesters (2012) found that people who are better at employing theory of mind are also better at coordinating their answers with other people on questions with many possible responses. Yet, characterizations of theory of mind almost always use shared knowledge as the paradigm case, and shared knowledge is in general insufficient to solve coordination problems. Imagine, for example, that Sally and Ann are trying to find each other at a fairground. They previously discussed meeting at the funhouse or the carousel but never came to an agreement. Where should Sally go to meet Ann? Sally can represent Ann's knowledge of the two locations, and her desire to meet at the same location, and vice-versa. Yet even if Ann thought it would be best to meet at the funhouse, and Sally knew that Ann thought so, but Ann worried that Sally thought it would be best to meet at the carousel, Ann might go to the carousel while Sally went to the funhouse. No matter how many nested levels of knowledge Sally represents, she will not know where to look for Ann. Coordination games pose a problem

that is typically not raised in the literature on theory of mind: How does one read the mind of a mind reader?

The Present Research

In these experiments we examine the cognitive processes underlying coordination. Participants interact with partners in a role-playing scenario that involves a symmetric coordination game, with payoffs that instantiate a Stag Hunt. In the game, participants must decide either to work alone, which offers a small but certain profit, or to try to work with a partner, which offers the potential to make more money but only if their partner makes the same choice: If they choose to work together but their partner does not, then they receive nothing. We test whether people differentiate between shared and common knowledge in deciding whether to try to work together, whether shared knowledge and common knowledge have distinct cognitive representations, and whether people use workarounds to a lack of common knowledge when attempting to coordinate their actions.

The game involves two merchants, a butcher and a baker, who decide each day whether to work independently to sell chicken wings and dinner rolls, respectively, or to work together to sell complete hot dogs, for which they earn more (Figure 1). No one will buy just the buns or just the hot dog meat, so they risk earning nothing if they fail to coordinate their actions. Moreover, participants are told that sometimes the hot dogs can earn them both more money than working independently, but sometimes hot dogs earn less money, so the merchants need common knowledge of higher profits to coordinate. But, their only means of communication with each other is an unreliable messenger boy.

To appreciate the need for common knowledge in this scenario, consider what happens on a given day that the baker sends a message to the butcher telling him to bring hot dogs. The

butcher sends a confirmation to let the baker know he received the message. The baker receives the confirmation, but realizes that the butcher cannot be sure whether the messenger delivered the confirmation. So, the baker sends a confirmation of the confirmation. Upon receipt of this message, the butcher realizes that yet another confirmation is required. In fact, no finite number of successful confirmations can help the hapless merchants because they can never be sure that the most recent confirmation message was delivered by the unreliable messenger boy, and neither knows how many messages might be sufficient for the other merchant to bring his ingredient for the hot dogs (embodying the second order coordination problem presented by multiple shared knowledge solutions).[2] Common knowledge is therefore needed to reliably solve the merchants' problem.

To test whether people tacitly appreciate this requirement, we manipulated what they knew about their partner's knowledge about the payoffs—whether knowledge of the payoffs was private, shared, or common. The game-theoretic analysis of coordination suggests the *common knowledge recognition hypothesis*: In coordination environments, people strategically differentiate between shared knowledge and common knowledge. This hypothesis predicts that participants will try to work together more frequently when they have common knowledge of the payoffs than when they have shared or primary knowledge.

Alternatively, people may not represent common knowledge as a distinct state. The only major distinction affecting their coordination decisions would then be the difference between primary and shared knowledge (as suggested by the theory of mind literature). We call this the *shared knowledge hypothesis.*

---

[2] Rubinstein (1989) describes a related scenario called "the electronic mail game" which he used to prove that no finite set of messages guarantees coordination with uncertain communication.

Finally, the literature on the curse of knowledge raises the possibility that people do not reliably use either shared or common knowledge to solve coordination problems. If people attribute their own knowledge to other people, then distinctions among levels of knowledge would be irrelevant. This *curse of knowledge hypothesis* predicts that participants will try to work together with the same frequency across all knowledge levels.

Knowledge-level representation

If people do distinguish between common and shared knowledge, then this raises the further question of how they represent the distinction. One possibility is that these knowledge states have a single cognitive format and the distinction between them is simply quantitative, with common knowledge being represented as an upper limit of shared knowledge. Alternatively, shared and common knowledge may have distinct representations, which would make the distinction qualitative and categorical.

These possibilities can potentially be distinguished by the pattern of classification errors people make when reporting their level of knowledge. Research on people's theory of mind capabilities suggests that shared knowledge becomes more difficult to represent as the levels of knowledge increase. If people entertain a single kind of representation (the *single-representation hypothesis*), then the most errors will be observed in the common knowledge condition (the maximum number of shared knowledge levels), with fewer errors made as the levels of shared knowledge decrease. If, in contrast, common knowledge has its own representation, it need not contain multiple levels of embedded knowledge; it could consist of a single mental symbol which means, "We have common knowledge." Thus, errors will increase only with the number of levels of shared knowledge, whereas errors on common knowledge will be few (similar to the error rate for secondary knowledge which also requires only a single level of representation).

The *distinct-representation hypothesis* makes the further prediction that errors will be systematic, respecting the boundary between the two kinds of knowledge: Different levels of shared knowledge will be mistaken for each other, but not for common knowledge.

Sensitivity to costs and benefits

To further characterize the decision processes behind coordination, we vary the game payoffs to test people's sensitivity to costs and benefits. Rational choice theory (e.g., Becker, 1976) indicates that people should decide to work together if the expected value—the amount earned for successful coordination multiplied by the probability that they think their partner will do the same—is greater than the amount earned for working alone. The *rational actor hypothesis* predicts that as the ratio of costs to benefits increases, people will be less likely to try to work together.

As we shall see, the requirements for successful coordination by rational actors in a coordination game are sometimes formidable, depending on precise knowledge of the distribution of beliefs and values among a large number of other actors. Given this complexity, it is possible that under a wide range of cases people may focus only on states of knowledge, in the hope that they can discern the common knowledge (or at least a sufficiently high degree of common *p*-belief) that would lead to successful coordination regardless of the details of the payoff structure in force. This *knowledge-level heuristic hypothesis* predicts that participants' decisions will not closely track the cost-benefit ratio of different payoffs.

Other motivations for coordination

Coordination decisions may be influenced by factors other than public knowledge. Since coordination requires elusive knowledge that another person has made the same choice as oneself, people may use their own decision-making processes to simulate how their partners will

think and behave (Gallese & Goldman, 1998), particularly when they view themselves as similar to their partner (Mitchell, Macrae, & Banaji, 2006; Tamir & Mitchell, 2013). According to this *perceived similarity hypothesis*, the more similar people perceive themselves to be to their partners, the more likely they will be to predict that their partner will choose as they do, and thus to attempt risky coordination.[3]

People may also be motivated to coordinate by altruistic or other-regarding preferences. Altruistic and reputational motives have been well documented in social psychology and experimental economics (e.g., Haley & Fessler, 2005; Messick & McClintock, 1968; Milinski, Semmann, & Krambeck, 2002; Van Lange, 1999). The Big Five personality trait of Agreeableness is associated with altruism, pro-sociality, friendliness, and generosity (Goldberg, 1992; Graziano & Tobin, 2002; Roccas, Sagiv, Schwartz, & Knafo, 2002), and has specifically been associated with altruistic motivations towards non-relatives and strangers (Graziano, Habashi, Sheese, & Tobin, 2007). The *altruistic motives hypothesis* predicts that people who are higher on Agreeableness will be more likely to try to work together.

Finally, some people may simply be willing to accept the potential cost of discoordination in the hope that they can earn more money through high-payoff coordination. The Big-Five personality trait of Openness is associated with risk-seeking (Nicholson, Soane, Fenton-O'Creevy, & Willman, 2005) and in particular with the seeking of chances for gains (Lauriola & Levin, 2001). The decision to work together is as a social gamble where one can bet a certain payout to win an additional increment of profit. The *risk-seeking hypothesis* predicts

[3]This is related to the concept of superrationality, in which rational actors decide to cooperate in a Prisoner's Dilemma because they each assume that both they and their partner rationally see the wisdom of mutual cooperation, and know that the other sees it, knows that the other knows that they know that the other sees it, and so on (see Colman, 2003; Fischer, 2009; Hofstadter, 1985). The same logic can be applied to coordination games with symmetrical payoff structures.

that people who are higher on Openness will be more likely to take the bet by trying to work together.

We report three experiments designed to test these hypotheses. In Experiments 1 and 2, we test the effects of knowledge level and cost-benefit structure on coordination decisions involving, respectively, one and three partners. In Experiment 3 we investigate how shared and common knowledge are cognitively represented and test the three social motivation hypotheses.

Experiment 1

Experiment 1 implements the butcher-baker coordination game explained above. Each participant interacts with a partner, playing the role of either the butcher or the baker. They read that they could work either alone or with the partner; the amount they could earn for working alone was constant, but the amount earned for working together would vary from day to day and might be less than or greater than the amount they could make by working alone. They were then told that on the day of the actual decision facing them the payoff for working together was greater than the payoff for working alone, and were then given one of four signals about what their potential partner knew about the payoff, which we varied in a between-subject design.

In the *primary knowledge* condition, a participant was told he or she could earn 10 cents more for working with the partner, but were not given information about what the partner knew. In the *secondary knowledge* condition, the participant was told that their partner also knew about this payoff. In the *tertiary knowledge* condition, they were told that their partner knew the payoff *and* knew that the participant himself or herself knew the payoff. In the *common knowledge* condition, the payoff was presented as public information, commonly known between the two participants.

To see whether coordination decisions were sensitive to costs and benefits, we manipulated, between subjects, the amount the participants could earn by working alone and together, yielding four payoff structures (see Figure 1): \$1.00/\$1.10, \$2.00/\$2.10, \$5.00/\$5.10, and \$10.00/\$10.10 (hereafter referred to as the \$1, \$2, \$5, and \$10 payoff conditions). We chose small payoffs for coordination (and thus relatively high opportunity costs for discoordination) to counter the typical demand characteristics of experimental games, which tend to encourage cooperative actions (Pederson, Kurzban, & McCullough, 2013).

Method

Participants. We used Amazon Mechanical Turk to recruit 1600 participants (100 per condition) from the United States to complete a short study for a small payment. After we excluded participants who gave incorrect answers to comprehension questions about the game's payoff structure (see Procedure), the final sample consisted of 1033 participants (58% female) with a mean age of 32.8 years (*SD* = 15.0).

Procedure. Participants read instructions explaining that they would earn a minimum of 50 cents, which they could augment based on their decisions in their interaction with another participant on Mechanical Turk. They were told that one of them would play a butcher and the other a baker. Each could either work alone for a sure profit (the butcher could make chicken wings, the baker dinner rolls) or attempt to work with their partner, the butcher making hot dogs, the baker the buns. By choosing to work together, they were told, the participant can earn a profit, but *only if the participant's partner also chooses to work together*; if either decides to collaborate but the partner does not, that person doesn't earn anything, because they cannot sell a bun without a hot dog or vice versa. Participants then read that they would earn a certain amount (\$1, \$2, \$5, or \$10) if they decided to work alone, but that the hot dog price varied from day to

day, and thus the earnings for working together might be more than or less than this sure profit. Finally, they read that the information about hot dog earnings might be conveyed to them by a messenger boy (displayed on their screen in a private box that only they could see), or by a loudspeaker (displayed on their screen in a public box that the other participant could see on his or her screen as well).

The participant then clicked a button to reveal the day's information about the price of hot dogs and hence the potential profit for collaborating; in each case it was ten cents more than each would earn by working alone. In the second between-subjects manipulation, participants received one of the following pieces of information (presented here from the perspective of the baker):

1. *Primary knowledge*—In the private box the participant read, "The Messenger Boy has not seen the Butcher today, so he cannot tell you anything about what the Butcher knows." The public box stated that the loudspeaker was silent.
2. *Secondary knowledge*—In the private box the participant read, "The Messenger Boy says he stopped by the butcher shop before coming to your bakery. He tells you that the Butcher knows what today's hot dog price is. However, he says that he forgot to mention to the Butcher that he was coming to see you, so the Butcher is not aware that you know today's hot dog price." The public box stated that the loudspeaker was silent.
3. *Tertiary knowledge*—In the private box the participant read, "The Messenger Boy mentions that he is heading over to the butcher shop, and will let the Butcher know today's price as well. The Messenger Boy will also tell the Butcher that he just came from your bakery and told you the price. However, the Messenger Boy will not inform the Butcher that he told you he would be heading over there. So, while the Butcher is aware

that you know today's price, he is not aware that you know that he knows that." The public box stated that the loudspeaker was silent.

4. *Common knowledge*—In the public box the participant read, "The loudspeaker broadcast the market price of [today's price] (of which you could earn [earnings for working together])." In the private box the participant read, "The messenger boy did not come by. Because the market price was broadcast on the loudspeaker, the Butcher knows [today's price], and he knows that you know this information as well."

The participant then made a decision to work alone or with their partner and indicated it with the keyboard. They were then asked to explain how they made the decision, and were given two sets of comprehension questions. The first contained three questions about the profits under various combinations of decisions (e.g., "If you chose to make hot dogs but the butcher did not, then how much would you have made?"), which allowed us to exclude participants who did not understand the game's payoff structure. The second contained four questions about what they and their partner knew, e.g., "Does the butcher know the price of hot dogs today? (yes/no/can't tell)".

Finally, participants filled out a brief demographic questionnaire, submitted the task, and received the base rate payment for completion. Offline, we randomly paired up participants to implement the conditions described in the scenario, calculated the profits determined by the two participants' decisions, and paid them that additional amount.

Results and discussion

Figure 2 shows that with all four payoffs, the percentage of participants who tried to work together was significantly affected by their state of knowledge (first row of Table 1). Planned comparisons across adjacent knowledge conditions (i.e., primary-secondary, secondary-tertiary,

and tertiary-common) are shown in Table 1 for all payoff conditions. In all four payoff conditions, more participants tried to work together with common knowledge than with tertiary knowledge. In three out of four payoff conditions, more participants tried to work together with secondary knowledge than with primary knowledge (the difference was only marginally significant in the $5 payoff condition). Coordination rates were the same with secondary and tertiary knowledge, except with the $5 payoff, for which the rate with secondary knowledge was anomalously low.

These results are consistent with the Common Knowledge Recognition hypothesis: Participants were more likely to try to work together with common knowledge than with any other state of knowledge. The results were inconsistent with a strong Curse of Knowledge hypothesis, because the likelihood of working together differed across knowledge conditions. In line with the Shared Knowledge hypothesis, few of the participants tried to work together with primary knowledge while more tried to work together with secondary knowledge. However, only slightly more participants tried to work together with tertiary knowledge, while far more participants tried to work together with common knowledge. The pattern is consistent with the hypothesis that people maintain a dual representation in which shared and common knowledge are thought of as qualitatively distinct.

These results were inconsistent with a strict Rational Actor hypothesis because the proportion of participants who decided to try to work together in each knowledge condition varied little across the payoff conditions, even as the ratio between the cost of the forgone profit from working alone to the additional benefit from working together increased 10-fold (Primary: $\chi^2(3, N = 261) = 3.13$, $p = .373$, $\phi = .11$; Secondary: $\chi^2(3, N = 259) = 7.66$, $p = .054$, $\phi = .17$; Tertiary: $\chi^2(3, N = 260) = 6.81$, $p = .078$, $\phi = .16$; Common: $\chi^2(3, N = 253) = .76$, $p = .859$, $\phi =$

.05). A rational actor would expect that as this cost-benefit ratio increases, the other rational actor would take the increased opportunity cost into account, and the probability that they would try to work together should correspondingly decrease, creating a positive feedback loop that would drive each of them to work alone. The fact that the proportion of people who tried to coordinate with common knowledge was invariant across payoffs contradicts the idea that coordination decisions were based on maximizing the expected payoff, and is instead consistent with the Knowledge-level Heuristic hypothesis.

Another test of the Rational Actor hypothesis may be obtained by examining the actual payouts that the participants would earn given their collective pattern of choices. Inspection of the frequency of coordination attempts in the secondary and tertiary knowledge conditions reveals that this payoff is likely to be low: Participants who decided to try to work together failed to coordinate with their partners (and thus relinquished their sure profit from working alone) around 50% of the time. To assess the overall rationality of these choices we calculated *expected earnings* based on all possible matchups with the other participants (rather than the actual earnings from the matchups we arbitrarily arranged in order to calculate their payments). This consists of the sum of the proportion of participants who chose to work alone, multiplied by the smaller certain payoff, and the proportion of participants that would, on average, successfully work with a partner (when both they and the partner chose to coordinate, which is the product of the proportion of participants that tried to work together and this proportion minus one), multiplied by the higher risky payoff. The discoordination payoff, from cases in which they would choose to cooperate but their partner would not, was zero, eliminating this term from the calculation. Figure 3a shows that for all payoffs, efficiency was higher with primary and common knowledge than with either level of shared knowledge.

In sum, Experiment 1 shows that when people make coordination decisions, they differentiate between primary, shared, and common knowledge (though apparently not among different levels of shared knowledge). Moreover, the level of knowledge, and the special appeal of common knowledge, are far more salient to them than the expected value of the options: Increasing the cost-benefit ratio 10-fold had no observable impact on their choices.

Experiment 2

How general is the sensitivity to knowledge and insensitivity to payoffs observed in Experiment 1? Presumably if achieving coordination is difficult enough, and the stakes are high enough, then even with common knowledge people would opt to work alone; as an extreme example, imagine risking a sure payoff of $1000 for working alone for a chance at earning $1001 by coordinating with a million partners. To test the limits of common knowledge as a qualitative coordination heuristic, we designed Experiment 2 as a four-person coordination game in which *all four* partners had to decide to work together to achieve the benefits of coordination. Coordination on the higher-paying option of working together is much more difficult with four people, because the probability of success is equal to the probability that any one partner decides to work together cubed.

In fact the perceived probability of successful coordination may fall even faster than that. In addition to common knowledge of the payoffs, coordination also requires individuals to be confident in their partners' rationality. An irrational partner could prefer lower payoffs, choose blindly, or make some other unpredictable choice. With only two players, the chance of an irrational partner might be negligible, but this risk can be greater in larger groups. Since even a single irrational partner can be enough to torpedo coordination in a group, as the number of players goes up the likelihood of discoordination increases rapidly (everyone must be both

knowledgeable *and* rational, and believe everyone else is as well). For this reason the cost-benefit structure may become more salient to a participant as the number of other partners increases. Recall that a rational actor may choose to coordinate with less-than-perfect common knowledge (i.e., with common *p*-belief) as long as the probability of the other's belief exceeds a critical value which depends on the relative payoffs: The higher the opportunity cost, the higher that probability must be. Thus we may see a greater sensitivity to payoffs in a coordination game involving more people.

Methods

Participants. As in Experiment 1, 1600 participants were recruited from Amazon Mechanical Turk, evenly distributed across the sixteen combinations of four payoff and four knowledge conditions. After we excluded participants who did not understand the payoffs, the sample consisted of 1150 participants (48% female, M_{age} = 31.9, SD_{age} = 11.1).

Design and procedure. Participants were told they could work together to make "superburgers," which require a burger, a bun, cheese, and toppings from, respectively, a butcher, a baker, a cheese maker, and a produce vender. One participant was assigned to each of these four roles. As in Experiment 1, each participant also had the option to make a food item on his or her own for a sure profit. Participants were told that they would receive a profit for contributing to superburgers only if *all three* of the other merchants made the same choice, and would receive nothing otherwise. In each of the four knowledge conditions the participant's three partners were said to have the same level of knowledge. This was conveyed with identical instructions to those of Experiment 1, except that "the Butcher" or "the Baker" was replaced with "the other merchants." All other aspects of the procedure were the same as in Experiment 1.

Results and discussion

As in Experiment 1, players' state of knowledge affected their decision to work together (Figure 4). Table 2 shows that in all payoff conditions, significantly more participants tried to work together with common knowledge than with tertiary knowledge, and in three of the four payoff conditions, significantly more tried to work together with secondary than with primary knowledge. In none of the payoff conditions was there a significant difference between secondary and tertiary knowledge. This consistent lack of significant differences between the shared knowledge conditions suggests that people treat secondary and tertiary knowledge similarly.

Unlike Experiment 1, people showed some sensitivity to the payoff structure. In all knowledge conditions, increasing the relative costs of coordination failures brought down coordination rates. This is consistent with the observation that the minimum level of confidence in common knowledge (i.e., the minimum common *p*-belief) required for rational coordination rises more steeply with opportunity costs when the number of players (and hence the chance that at least one will be ignorant or irrational or both) increases.

Though participants' sensitivity to payoffs was more consistent with the Rational Actor hypothesis than in Experiment 1, another aspect of the results was not. Unlike what we obtained in Experiment 1, the most profitable knowledge condition with all four payoffs was Primary Knowledge (Figure 3b); Common Knowledge was *less* profitable than Shared Knowledge except with the least costly forgone payoff of $1. This shows an important limit to the advantages that people can obtain from common knowledge. When either the knowledge state or the rationality of all the necessary potential partners is less than perfect, coordination is difficult to achieve and hence poses a high risk of failure. In those cases even high rates of decisions to coordinate may not be enough to consummate successful coordination, and the temptation to coordinate

presented by common knowledge can actually *reduce* the coordinators' payoff. Yet more than half of the participants provided with common knowledge still opted for the risky higher payoff.

Experiment 3

Experiment 3 is a replication of one of the payoff conditions from Experiment 1 with additional components that allow us to test how shared and common knowledge are represented, and why people sometimes make what appears to be an irrational decision to cooperate with just shared knowledge.

At least since Miller & Nicely (1955), cognitive psychologists have used confusion matrices to test hypotheses about underlying mental representations, based on the assumption that confusable stimuli are likely to be represented similarly. A similar logic underlies the memory confusion paradigms commonly used in social psychology to reveal the dimensions of social categorization, such as the "Who said what?" paradigm (e.g., Klauer & Wegener, 1998; Lieberman, Oum, & Kurzban, 2008; Taylor, Fiske, Etcoff, & Ruderman, 1978). In our case, we use errors in responses to our questions about participants' comprehension about the level of knowledge as evidence of whether shared and common knowledge are represented in the same or in qualitatively distinct ways. Unfortunately, in the first two experiments these questions were so easy that all participants got them all correct. In this experiment, we made the questions more difficult in three ways: by putting them at the end of the survey, by adding a task before participants answered them, and by concealing the relevant information while they answered the questions (in the first two experiments, this information was visible on the screen).

Recall from the Introduction that there are several reasons that people may choose to coordinate even in the absence of common knowledge. One is that an actor may ascertain that she has similar values and biases to a potential partner, and thus that the partner is likely to

assess the situation in the same way that she does, including an assessment of whether she herself is likely to choose to coordinate with the partner. We thus manipulated whether the participants thought they were interacting with a partner who was similar or dissimilar to themselves in age, political orientation, tastes in music, and decision-making style. The other social motivations for coordination consist of personality traits that make the choice inherently appealing, including Agreeableness, which impels people to act in a pro-social manner, and Openness, whose risk-seeking component may impel people to gamble for a big payoff rather than accepting a smaller but surer payoff.

Methods

Participants. We recruited 800 participants from Mechanical Turk, evenly distributed across similarity and knowledge conditions. After eliminating people who failed the comprehension questions, we were left with 550 participants in the final analyses (approximately 46% male, M_{age} = 31.6, SD_{age} = 11.3).

Design and procedure. The design added three components to the $2 condition from Experiment 1.

Similarity manipulation. At the beginning of the experiment, participants answered four questions:

- "Do you prefer more intense kinds of music (e.g., rock or rap) or more mellow kinds of music (e.g., classical or jazz)?"
- "If you had to pick, would you say you are more liberal or more conservative?"
- "How old are you? [available answers: "I'm 35 years old or older," and "I'm younger than 35 years old"]
- "When making decisions do you tend to rely more on intuition or more on reason?"

In the Similar condition, participants were told that they would be matched with a partner who gave the same answers to three or more of these questions. In the Dissimilar condition, participants are told that they would be matched with a partner who gave the same answers to two or fewer of these questions. Participants were then asked to report how similar they perceived their partner to be to them, on a scale from 0% to 100%.

Big Five personality questionnaire. After participants read the role-playing scenario and made their decision, they were asked to fill out a standard 50-question survey that measured the Big Five personality traits (Goldberg, 1999).

Knowledge-level comprehension questions. Finally, the knowledge-level comprehension questions were administered on a separate page; when answering them, participants were unable to refer back to the initial instructions.

Results and discussion

Figure 5 shows that the Similarity manipulation made no systematic difference. Furthermore, while ratings of perceived similarity were higher in the Similar condition ($t(548) = 13.87$, $p < .001$), these subjective perceptions of similarity had no effect on participants' decisions (Wald $\chi^2(1, N = 550) = 0.01$, $p = .944$). We thus collapse across similarity in all other analyses.

Knowledge level had the same effect as in the first two experiments: More people tried to work together with common knowledge than with tertiary knowledge, $\chi^2(1, N = 270) = 22.28$, $p < .001$, $\phi = .29$, and more tried to work together with secondary than with primary knowledge, $\chi^2(1, N = 280) = 16.87$, $p < .001$, $\phi = .25$, but there was no difference between secondary and tertiary knowledge, $\chi^2(1, N = 284) = 0.72$, $p = .397$, $\phi = .05$.

Representations of shared and common knowledge. The confusion matrix for the questions about levels of knowledge is shown in Table 3. Participants made significantly more errors with tertiary knowledge than with any of other level of knowledge (planned comparisons: primary-tertiary, $\chi^2(1, N = 266) = 33.76$, $p < .001$, $\phi = .36$; secondary-tertiary, $\chi^2(1, N = 284) = 27.91$, $p < .001$, $\phi = .31$; common-tertiary, $\chi^2(1, N = 270) = 13.93$, $p < .001$, $\phi = .23$), and these errors consisted overwhelmingly of misremembering it as secondary knowledge (an error made by 23% of the participants in this condition). Error rates with common knowledge and with secondary knowledge were not significantly different ($\chi^2(1, N = 284) = 2.43$, $p = .119$, $\phi = .09$). None of the other off-diagonal confusions was as high as the one for mistaking tertiary for secondary knowledge. The next highest was 4% (mistaking tertiary knowledge for common knowledge), which was significantly different from the 23% rate for mistaking tertiary for secondary knowledge ($p < .001$).

These results show that higher levels of knowledge are increasingly difficult to represent, as suggested by the Theory of Mind literature, but only when the knowledge is merely shared; the highest level of all, common knowledge, is almost as easy to represent as the lowest level of shared knowledge. The confusion matrix thus suggests that shared and common knowledge have distinct cognitive representations, but that quantitatively different levels of shared knowledge do not.

Altruistic motives. Figure 6 shows that participants in the Shared knowledge conditions who tried to work together scored higher in Agreeableness than those who decided to work alone, a difference not observed in the Primary or Common knowledge conditions. Logistic regression, controlling for the main effect of knowledge condition, revealed a significant Knowledge × Agreeableness interaction in coordination attempts, Wald $\chi^2(3, N = 550) = 8.18$, p

= .043. Post-hoc *t*-tests with Agreeableness as the dependent variable and Decision (work alone vs. together) as an independent variable confirmed that the people who decided to work together with secondary and tertiary knowledge were significantly more agreeable ($t(147) = 2.25$, $p = .013$, and $t(133) = 1.89$, $p = .030$, respectively), but people who decided to work together with primary or common knowledge were not ($p > .60$). These results are consistent with the hypothesis that with shared knowledge, people may choose to coordinate with others out of a sense of altruism, perhaps as a signal to encourage coordination in possible future opportunities.

Risk-seeking. Figure 7 shows a similar pattern for the trait of Openness to Experience (Wald $\chi^2(3, N = 550) = 13.64$, $p = .003$). Coordinators were more Open than non-coordinators when they made their decision with secondary knowledge ($t(147) = 3.30$, $p < .001$), and when they made it with tertiary knowledge ($t(133) = 1.82$, $p = .036$), but not when they made the decision with primary or common knowledge ($p > .80$). These results are consistent with the hypothesis that people recognize that, as game theory predicts, attempting to coordinate with shared knowledge is risky, but attempting to coordinate with common knowledge is not; those who seek risks for gains may thus gamble in conditions of shared knowledge.

No differences were found for the other three personality factors (Extraversion, Conscientiousness, or Neuroticism); all *p*s > .2.

General Discussion

Humans have lived in large groups throughout their evolutionary history, providing many opportunities for mutually beneficial coordination. Game-theoretic models show that common knowledge has a privileged role in helping individuals solve coordination problems. Taken together, these observations suggest that humans evolved cognitive mechanisms for recognizing common knowledge and distinguishing it from shared knowledge. Our results support this

hypothesis. In all three experiments, and with every combination of payoffs, participants were more likely to attempt risky coordination with common knowledge than with shared knowledge. Moreover, coordination attempts with common knowledge did not closely track the cost-benefit ratio across payoff conditions, indicating that behavior was not driven by estimated probabilities of coordination but instead by a categorical recognition of a state of common knowledge.

In contrast to the marked distinctions participants made between common and shared knowledge, they made little distinction between different levels of shared knowledge. This is notable because the levels of shared knowledge tested here, secondary and tertiary, span almost the entire range of shared knowledge that people can readily represent, falling just one level short of the four-level maximum observed by Kinderman et al. (1998). The similar rates of coordination attempts with different levels of shared knowledge is echoed by the pattern of confusions in knowing which state of knowledge was present: People confused tertiary with secondary knowledge, but rarely confused common knowledge with shared knowledge or with primary knowledge, nor shared knowledge with primary knowledge. This is further reinforced by the finding that people who were higher on the personality traits of Agreeableness or Openness were more likely to attempt risky coordination in both shared knowledge conditions, and only in these shared knowledge conditions.

The cognitive difference between primary and shared knowledge has long been established in the literature on theory of mind (Apperly & Butterfill, 2009; Saxe & Young, in press). The present results suggest that there is also a fundamental difference between shared knowledge and common knowledge, with common knowledge being represented as a distinct conceptual category. In contrast to their sensitivity to their partner's knowledge state, participants were largely insensitive to the game's payoff structure, responding to changes in

expected returns only in an extreme case when successful coordination required three other people to make the same choice.

A puzzle in the results was why a fair number of participants chose to coordinate with shared knowledge. Recall that a participant with secondary knowledge was always matched with a partner with only primary knowledge (they were told that their partner knew the profit for working together, but that their partner was unaware that they also knew this profit). Hence, to coordinate with their partner, a participant should *act in the same way that they would act with primary knowledge*. But overall, our participants defied this logic, with more of them opting to work together with secondary knowledge than with primary knowledge, providing further support for the claim that participants used the different knowledge states heuristically, rather than rationally calculating their best move.

If they were not playing their best move in cost-benefit terms, why would people risk a certain payout for an uncertain small gain? The personality results indicate that participants may have chosen to try to work together to signal their cooperative character (Agreeableness), or because of a risk-seeking disposition (Openness). Participants did not leverage similarity in making their decisions, but the possibility remains that this manipulation was too cursory to be effective. The rewards for working alone and working together were abstract monetary payoffs, which differ only in magnitude and not in semantic content. However, in many real world situations the nature of the similarity between people is highly relevant to the kind of coordination they are considering. Two roommates, for example, who will have to endure or enjoy each other's music, ought to find similarity in musical tastes more relevant to the decision to live together than similarity in personality or politics.

The pattern of participant payouts provides clues as to why certain aspects of participants' behavior in these experiments were suboptimal. In an arbitrary social situation (such as the artificial scenario of interacting with strangers on the Internet in a contrived game), it is extraordinarily hard to predict whether coordination will be profitable, because it depends critically on small and unpredictable differences in the decisions of the other participants. In the case of the four-person game above (Experiment 3), coordination turned out to be the *least* profitable strategy, even when a majority of participants chose to coordinate, because even a relatively small number of non-coordinators was enough to scuttle coordination and its rewards. With people unable to predict at exactly which combination of probabilities and payoffs in a given situation this tipping point lies, they may focus predominantly on information that indicates other people's state of knowledge. Actors coordinate when they have evidence for common knowledge, and refrain from coordinating when they do not. For this heuristic to be advantageous in real life, people must have high quality information about common knowledge in ecologically typical environments, and other people's sensitivity to such information.

Broader implications: Common knowledge in social life

The finding that people use common knowledge in their decisions to coordinate their behavior, the evidence that common knowledge is a distinct cognitive category, and the suggestion that everyday social life provides reliable cues to common knowledge in opportunities for coordination, all imply that common knowledge has a strong presence in human life and in the phenomena studied by social psychology. This makes it surprising that the psychology of common knowledge has apparently had so little visibility either in psychology or in everyday life. If coordination is as important to social life as altruism, and if common

knowledge is as indispensable to coordination as reciprocity is to altruism, shouldn't we expect our language and our lives to be permeated with ideas of common knowledge?

We suggest that this is indeed the case, even if it has not been fully appreciated. Just as the logic of reciprocity makes us obsessed with concerns such as debt, favor, bargain, obligation, and so on, the logic of common knowledge makes us obsessed with concerns such as publicity, privacy, confidentiality, conventional wisdom, fame, celebrity, hypocrisy, taboo, tact, euphemism, piety, mock outrage, and political correctness. In other words, both psychology and everyday social life *have* been concerned with the manifestations of common knowledge, but they have not been treated as exemplars of a single principle. We suggest that an acknowledgement of the role of common knowledge in enabling coordination can unify and explain a variety of seemingly unrelated and puzzling phenomena. In particular, much of social life is affected by common-knowledge generators, and much of language and cognition is sensitive to the state of common knowledge.

The most obvious common-knowledge generator is *direct speech.* When one person says something to another in "plain language" or "in so many words," the content of the proposition is common knowledge. Lee and Pinker (2010) showed that when an experimental participant read a vignette in which one person issues an overt threat, bribe, or sexual come-on to another, the participant assumed that each party knows that the other knows that he knows (etc.) the relevant intention, whereas when the same proposition is proffered in an innuendo, even an obvious one, the participant assumes only that the parties privately know the content of the proposition (e.g., "Michael offered a bribe"), not the higher-order levels of knowledge (e.g., "Michael knows that the officer knows that he offered a bribe").

The generation of common knowledge may be the function of other deliberate and salient communicative acts. One example is *joint attention* (Scaife & Bruner, 1975; Tomasello, 1995), in which two people look back and forth at an object and at each other. Joint attention is thought to facilitate the acquisition of words, a classic example of a coordination equilibrium (see Lewis, 1969). Another class of common-knowledge generators consists of *performatives* (Austin, 1962; Searle, 1989) and the associated phenomenon of *public ceremonies*, in which the public utterance of a proposition (e.g., "I now pronounce you man and wife") ratifies a new coordination equilibrium such as a marriage, law, or court decision. Much of our moral psychology, including moral debate and condemnation, generates common knowledge of prohibited actions, which allows people to coordinate aggression toward wrongdoers (DeScioli, Bruening, & Kurzban, 2011; DeScioli & Kurzban, 2013). And, as we mentioned in the Introduction, challenges to power that require coordination among many actors are often effected by *public protests*, and increasingly their electronic equivalents.

Common knowledge can also be conveyed nonverbally. Indeed, we propose that the nonverbal signals which accompany self-conscious emotions evolved with their peculiar anatomy and physiological configurations (Tracy & Matsumoto, 2008) precisely because those configurations are simultaneously salient to the expresser and the perceiver (see Provine, 1996, 2012, for discussion). The perceiver knows not only the intended mental state of the expresser but knows that the expresser knows it, that the expresser knows that the perceiver knows it, and so on. Among these nonverbal common-knowledge generators may be the following.

- *Eye contact* is a potent social signal of threats and sexual come-ons precisely because both parties commonly know that they are acknowledging each other's acknowledgment.

- *Blushing* is felt as a somatosensory sensation by the blusher at the same time as it is displayed as a change in skin color to the perceiver. The acute discomfort in blushing resides largely in the knowledge that the blusher knows he or she is blushing, knows that an onlooker knows it, that the onlooker knows that the blusher knows that they know, and so on.
- *Crying* has the same inside-outside salience: A distraught person looking at onlookers through tears cannot avoid the knowledge that others know his tearful state, know that he knows, and so on.
- *Laughter*, with its disruption of the respiration rhythms necessary for speech and its unignorable noise, is also mutually salient to expresser and perceiver.

If this analysis is correct, it predicts that the common knowledge generated by each of these displays is necessary to attain a mutually beneficial equilibrium in a coordination game. Pinker (2007), Pinker, Nowak, & Lee (2008), and Lee & Pinker (2010) suggest that the relevant game is the joint adoption of a Relational Model that consensually governs their interactions, such as communal sharing, authority ranking, equality matching, or market pricing (Fiske 1992, 2004). For example, two people can prosper if they agree to be friends and share things indiscriminately, or if they agree to transact business and one sells something to the other, but not if one believes they are friends and helps himself to a possession that the other is in the business of selling. In the case of expressions of self-conscious emotions, the game may consist of two parties agreeing that one of them has committed an unintended or regretted harmful act, or is in a vulnerable state, and thus that the second one needn't punish or ostracize him. This equilibrium leaves both of them better off than they would be if the second incurred the cost of punishing or ostracizing the first for a harm he would never repeat anyway (McCullough, 2008).

If this theory of nonverbal communication is correct, then expressions that are less likely to generate common knowledge (such as facial expressions which a person can express with little awareness he is expressing it) should not be yoked to an identifiable coordination game.

Finally, if common knowledge is a pervasive concern of social life, then it should leave a mark on language in the form of a *conceptual metaphor* (Lakoff & Johnson, 1980; Pinker, 2007): a family of idioms organized around a central image, such as ARGUMENT IS WAR or LOVE IS A JOURNEY. In the case of common knowledge, the central image alludes to the quintessential common-knowledge generator: COMMON KNOWLEDGE IS A CONSPICUOUS OBJECT OR SOUND. Thus we have a family of expressions which invoke a salient object or event to assert that some proposition or speech act is common knowledge (and hence compels acknowledgment and action by two or more parties):

The emperor's new clothes.

The elephant in the room.

It's out there; you can't take it back.

It's on the record; to go on record.

The bell can't be unrung (*also*: Some things once said cannot be unsaid).

That's a pretty big matzo ball hanging out there [when one person says "I love you" and the other doesn't reciprocate; from the television show *Seinfeld*]

A bald lie; a barefaced lie [*compare*: a veiled threat; a fig leaf]

To save face; to lose face.

That insult was in his face; he couldn't ignore it.

It's as plain as the nose on your face.

In recent decades, psychologists have recognized that cooperation is one of the hallmarks of the human species, and that its game-theoretic demands have shaped our emotions, our morality, our social relationships, and our language. Much has been learned about these domains of psychology from a focus on the problem of altruistic cooperation and the mechanisms of reciprocity. We hope that comparable insights are waiting to be discovered by psychologists as they investigate the problem of mutualistic cooperation, and the mechanisms of common knowledge are—as we might say—put out there.

References

Alberucci, L., & Jäger, G. (2005). About cut elimination for logics of common knowledge. *Annals of Pure and Applied Logic, 133*(1-3), 73-99. doi:10.1016/j.apal.2004.10.004

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953-970. doi:10.1037/a0016923

Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics, 4*(6), 1236-1239. doi:10.1214/aos/1176343654

Austin, J. L. (1962). *How to do things with words, 2nd ed.* J. O. Urmson & M. Sbisa (Eds.). Cambridge, MA: Harvard University Press.

Bardsley, N., Mehta, J., Starmer, C., & Sugden, R. (2008, December). *Explaining focal points: Cognitive hierarchy theory versus team reasoning.* Discussion paper, Centre for Decision Research and Experimental Economics. Retrieved from http://www.nottingham.ac.uk/cedex/documents/papers/2008-17.pdf

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

Barry, E. (2011, July 14). Sound of post-Soviet protest: Claps and beeps. *The New York Times*. Retrieved from http://www.nytimes.com/2011/07/15/world/Europe/15belarus.html?pagewanted=all

Becker, G. S. (1976). *The economic approach to human behavior.* Chicago, IL: The University of Chicago Press.

Binmore, K., & Samuelson, L. (2001). Coordinated action in the electronic mail game. *Games and Economic Behavior, 35,* 6-30. doi:10.1006/game.1998.0749

Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state

attribution. *Psychological Science, 14*, 283-286. doi:10.1111/1467-9280.03436

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*, 382-386. doi:10.1111/j.1467-9280.2007.01909

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning. *Psychological Science, 16*(5), 378-384. doi:10.1111/j.0956-7976.2005.01544.x

Camerer, C., Loewenstien, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy, 97*, 1232-1254. Retrieved from http://www.jstor.org/stable/1831894

Chartier, C., Abele, S., Stasser, G., & Shriver, E. (2012, January). *Secondary and Schelling salience in matching and mismatching coordination*. Poster session presented at the 13th annual Society for Personality and Social Psychology meeting, San Diego, CA.

Chaudhuri, A., Schotter, A., Sopher, B. (2009). Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. *The Economic Journal, 119*(534), 91-122. doi:10.1111/j.1468-0297.2008.02207.x

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

Clark, H. H. (1992). *Arenas of language use.* Chicago, IL: University of Chicago Press.

Clark, H. H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge, UK: Cambridge University Press.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26*, 139-198. doi:10.1017/S0140525X03000050

Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). Oxford, UK: Oxford University Press.

Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 584-627). Hoboken, NJ: Wiley.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-185. doi: 10.1017/S0140525X01003922

Curry, O., & Chesters, M. J. (2012). 'Putting ourselves in the other fellow's shoes': The role of 'Theory of Mind' in solving coordination problems. *Journal of Cognition and Culture, 12*, 147-159. doi:10.1163/156853712X633974

Dalkiran, N. A., Hoffman, M., Ramamohan, P., Ricketts, D., & Vattani, A. (2012). Common knowledge and state-dependent equilibria. In *Proceedings of Symposium on Algorithmic Game Theory* (pp. 84-95). doi:10.1007/978-3-642-33996-7_8

DeScioli, P., Bruening, R., & Kurzban. R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior*, 32, 204-215. doi: 10.1016/j.evolhumbehav.2011.01.003

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, 139, 477-496. doi:10.1037/a0029065

Fischer, I. (2009). Friend or foe: Subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General, 138*(3), 341-350. doi:10.1037/a0016073

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99,* 689-723. doi:10.1037/0033-295X.99.4.689

Fiske, A. P. (2004). Four modes of constituting relationships: Consubstantial assimilation; space, magnitude, time, and force; concrete procedures; abstract symbolism. In N. Haslam (Ed.), *Relational models theory: A contemporary overview* (pp. 61-146). Mahwah, NJ: Erlbaum.

Frith, U., & Frith, C. D. (2003). Developmental and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B, 358*, 459-473. doi: 10.1098/rstb.2002.1218

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493-501. doi:10.1016/S1364-6613(98)01262-5

Geanakoplos, J. (1992). Common knowledge. *The Journal of Economic Perspectives, 6*(4), 53-82. doi:10.1257/jep.6.4.53

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Dreary, F. De Fruyt, & F. Ostendord (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Goldberg, L. R. (1992). The development of markers of the Big Five factor structure. *Psychological Assessment, 4,* 26-42. doi:10.1037/1040-3590.4.1.26

Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness,

empathy, and helping: A person × situation perspective. *Journal of Personality and Social Psychology, 93*(4), 583-599. doi:10.1037/0022-3514.93.4.583

Graziano, W. G., & Tobin, R. M. (2002). Agreeableness: Dimension of personality or social desirability artifact? *Journal of Personality, 70*(5), 695-728. doi:10.1111/1467-6494.05021

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*, 245-256. doi:10.1016/j.evoljumbehav.2005.01.002

Hamlin, J. K., Hallinan, E. V., & Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Developmental Science, 11*(4), 487-494. doi: 10.1111/j.1467-7687.2008.00694.x

Hofstadter, D. (1985). *Metamagical themas: Questing for the essence of mind and pattern.* New York, NY: Basic Books.

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

Kinderman, P., Dunbar, R., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology, 89,* 191-204. doi:10.1111/j.2044-8295.1998.tb02680.x

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "who said what?" paradigm. *Journal of Personality and Social Psychology, 75*(5), 1155-1178. doi: 10.1037/0022-3514.75.5.1155

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago, IL: University of Chicago Press.

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences, 31*, 215-226. doi:10.1016/S0191-8869(00)00130-6

Lee, J. J., Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review, 117*(3), 785-807. doi:10.1037/a0019688

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

Lieberman, D., Oum, R., & Kurzban, R. (2008). The family of fundamental social categories includes kinship: Evidence from the memory confusion paradigm. *European Journal of Social Psychology, 38*(6), 998-1012. doi:10.1002/ejsp.528

McCullough, M. E. (2008). *Beyond revenge: The evolution of the forgiveness instinct.* San Francisco, CA: Jossey-Bass.

Mehta, J., Starmer, C., & Sugden, R. (1994a). Focal points in pure coordination games: An experimental investigation. *Theory and Decision, 36*(2), 163-185. doi:10.1007/BF01079211

Mehta, J., Starmer, C., & Sugden, R. (1994b). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review, 84*(3), 658-673. Retrieved from http://search.proquest.com.ezp-prod1.hul.harvard.edu/docview/1289709287?accountid=11311

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4*, 1-25. doi:10.1016/0022-1031(68)90046-2

Milinski, M., Semmann, D., Krambeck, H. (2002). Reputation helps solve the ‘tragedy of the

commons'. *Nature, 415*, 424-426. doi:10.1038/415424a

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some english consonants. *Journal of the Acoustical Society of America, 27*(2), 338-352. doi: 10.1121/1.1907526

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50*, 655-663. doi: 10.1016/j.neuron.2006.03.040

Monderer, D., & Samet, D. (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior, 1*(2), 170-190. doi:10.1016/0899-825689900171

Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and domain-specific risk taking. *Journal of Risk Research, 8*(2), 157-176. doi:10.1080/1366987032000123856

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255-258. doi:10.1126/science.1107621

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* Cambridge, UK: Cambridge University Press.

Pedersen, E. J., Kurzban, R., & McCullough, M. E. (2013). Do humans really punish altruistically? A closer look. *Proceedings of the Royal Society B: Biological Sciences, 280*(1758). doi:10.1098/rspb.2012.2723

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that…" attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437-471. doi:10.1016/0022-0965(85)90051-7

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, 105*(3), 833-838. doi:10.1073/pnas.0707192105

Provine, R. R. (1996). Laughter. *American Scientist, 84*(1), 38-45. Retrieved from http://www.jstor.org/stable/29775596

Provine, R. R. (2012). *Curious behavior: Yawning, laughing, hiccupping, and beyond.* Belnap Press.

Roccas, S., Sagiv, L., Schwartz, S. H., Knafo, A. (2002). The Big Five personality factors and personal values. *Personality and Social Psychology Bulletin, 28*(6), 789-801. doi:10.1177/0146167202289008

Rousseau, J. (1984). *A discourse on inequality.* (M. Cranston, Trans.). New York, NY: Penguin Books. (Original work published 1754).

Rubinstein, A. (1989). The electronic mail game: Strategic behavior under "almost common knowledge." *The American Economic Review, 79*(3), 385-391. Retrieved from http://www.jstor.org/stable/1806851

Saxe, R., & Young, L. (in press). Theory of mind: How brains think about thoughts. In K. Ochsner & S. Kosslyn (Eds.), *The Handbook of Cognitive Neuroscience*. Oxford, UK: Oxford University Press.

Scaife, M, & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*(5489), 265-266. doi:10.1038/253265a0

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Searle, J. (1989). How performatives work. *Linguistics and Philosophy, 12*(5), 535-558. doi: 10.1007/BF00627773

Smith, N. (1982). *Mutual knowledge.* Orlando, FL: Academic Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure.* Cambridge, UK: Cambridge University Press.

Sugden, R. (1995). A theory of focal points. *The Economic Journal, 105*(430), 533-550. Retrieved from http://www.jstor.org/stable/2235016

Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization in intergroup behavior. *European Journal of Social Psychology, 1*, 149-178. doi: 10.1002/ejsp.2420010202

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General, 142*(1), 151-162. doi:10.1037/a0028232

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778-793. doi:10.1037/0022-3514.36.7.778

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore, & P. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum.

Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. In H. Høgh-Olesen (Ed.), *Human morality & sociality: Evolutionary and comparative perspectives* (pp. 91-234). New York, NY: Palgrave MacMillan.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for *n*-person exchange: The evolutionary roots of organizational behavior. *Managerial and Decision Economics, 27*, 103-129. doi:10.1002/mde.1287

Tracy, J. L., & Matsumoto, D. (2008). The spontaneous expression of pride and shame: Evidence for biologically innate nonverbal displays. *Proceedings of the National Academy of Sciences, 105*(33), 11655-11660. doi:10.1073/pnas.0802686105

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35-57. Retrieved from http://www.jstor.org/stable/2822435

Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*, 337-349. doi:10.1037/0022-3514.77.2.337

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684. doi:10.1111/1467-8624.00304

Willer, R., Kuwabara, K., & Macy, M. (2009). The false enforcement of unpopular norms. *American Journal of Sociology, 115*(2), 451-490. doi:10.1086/599250

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. doi:10.1016/0010-0277(83)90004-5

Table 1

Comparison of Knowledge Levels in each Payoff Condition, Experiment 1

Knowledge Levels	$1.00 Payoff			$2.00 Payoff			$5.00 Payoff			$10.00 Payoff		
	χ^2	n	ϕ	χ^2	n	ϕ	χ^2	n	ϕ	χ^2	n	ϕ
All Levels	63.70***	276	.48	62.79***	272	.48	56.71***	236	.49	70.41***	249	.53
1^o vs. 2^o	21.26***	147	.38	23.27***	134	.42	2.66	121	.15	19.06***	118	.40
2^o vs. 3^o	0.69	141	.07	0.63	139	.07	10.88***	113	.31	0.01	126	.01
3^o vs. CK	8.84**	129	.26	7.12**	138	.26	4.02*	115	.19	23.05***	131	.42

Note. Chi-square tests for the proportions of subjects who tried to work together across all knowledge levels, and adjacent knowledge levels by payoff condition. We compare primary (1^o), secondary (2^o), tertiary (3^o), and common knowledge (CK). The comparisons across all knowledge levels have three degrees of freedom, and the comparisons across adjacent knowledge levels have one degree of freedom.

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 2

Comparison of Knowledge Levels in each Payoff Condition, Experiment 2

Knowledge Levels	$1.00 Payoff			$2.00 Payoff			$5.00 Payoff			$10.00 Payoff		
	χ^2	n	ϕ	χ^2	n	ϕ	χ^2	n	ϕ	χ^2	n	ϕ
All Levels	68.24***	281	.49	45.32***	297	.39	46.36***	283	.41	24.72***	289	.29
1^o vs. 2^o	13.59***	142	.31	14.95***	147	.32	1.77	148	.11	11.54***	140	.29
2^o vs. 3^o	2.50	132	.11	0.69	153	.07	2.26	137	.13	0.03	149	.01
3^o vs. CK	10.14**	139	.27	6.21*	150	.20	12.39***	135	.30	4.23*	149	.17

Note. Chi-square tests for the proportions of subjects who tried to work together across all knowledge levels, and adjacent knowledge levels by payoff condition. We compare primary (1^o), secondary (2^o), tertiary (3^o), and common knowledge (CK). The comparisons across all knowledge levels have three degrees of freedom, and the comparisons across adjacent knowledge levels have one degree of freedom.

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 3

Proportion of Participants Reporting Different Levels of Knowledge in Each Condition in Experiment 3

Condition	Reported level of knowledge				
	Primary	Secondary	Tertiary	Common	Unclassifiable[a]
Primary	.931	.008	.008	.008	.046
Secondary	.020	.899	.013	.007	.060
Tertiary	0	.230[b]	.637	.044	.089
Common	.007	0	.015	.837	.141

Note. Participants' perceived knowledge level by (actual) knowledge condition. Participants' perceived knowledge level was assessed using comprehension questions. Accurate judgments are those on the diagonal, and are given again in the last column

[a]Unclassifiable errors correspond to patterns of errors that were logically inconsistent (e.g., reporting that they had tertiary knowledge, but not primary knowledge), incomplete, or in which participants reported the correct level of knowledge but chose "can't tell" rather than "yes" for some level of knowledge that they did have.

[b]According to a sign test, tertiary knowledge was mistaken for secondary knowledge more frequently than for common knowledge, $p < .001$.

		Butcher's options	
		Work Together (Hot Dogs)	Work Alone (Chicken Wings)
Baker's options	Work Together (Hot Dogs)	$1.10, $1.10	$0, $1.00
	Work Alone (Dinner Rolls)	$1.00, $0	$1.00, $1.00

Figure 1. An interaction between a butcher and a baker in Experiment 1. The baker chooses a row and the butcher chooses a column. The four cells show the payoffs (baker's payoff, butcher's payoff) for each combination of choices. These payoffs generate a coordination game, specifically a *stag hunt game* (Skyrms, 2004), in which one equilibrium is better for both players than another equilibrium. The other three payoff conditions substitute $2.00, $5.00, or $10.00 for the $1.00 payoff, and $2.10, $5.10, or $10.10 for the $1.10 payoff.

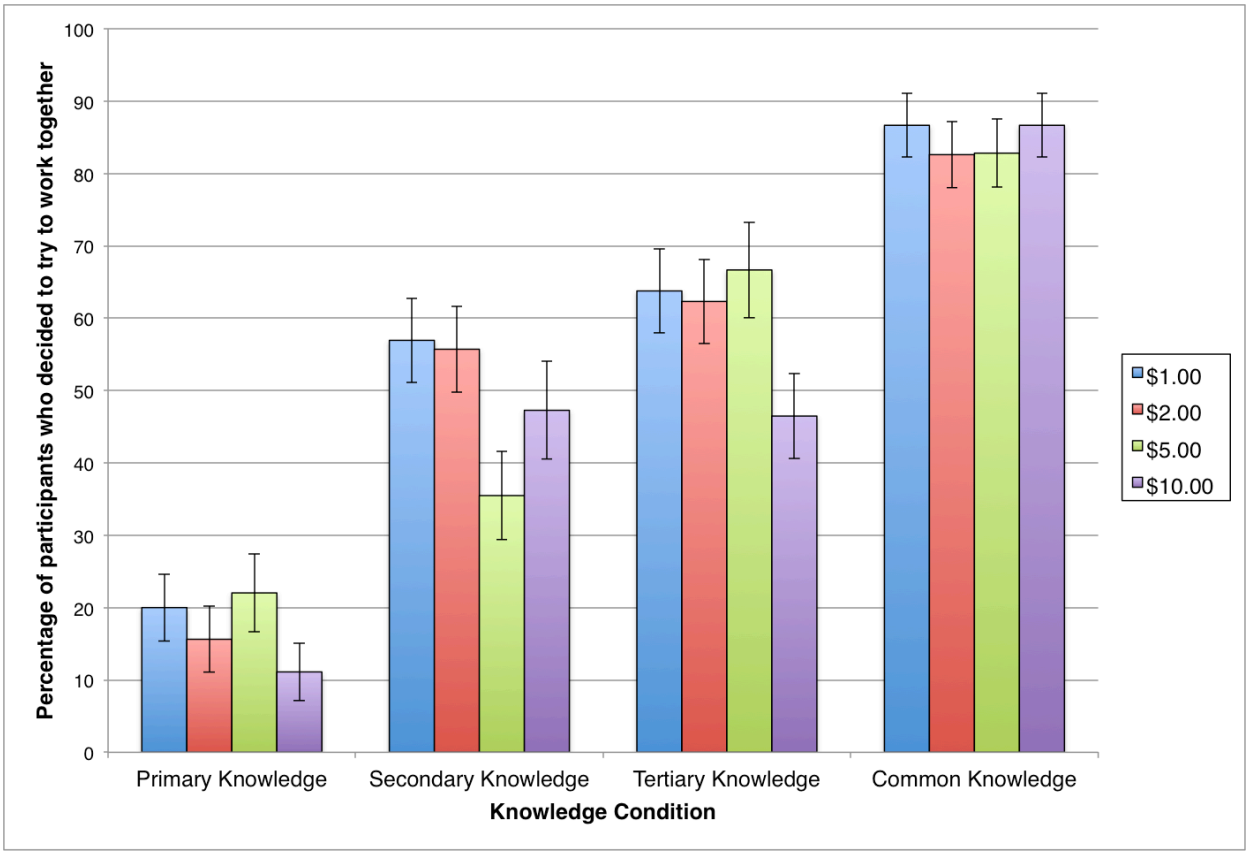


Figure 2. Percentage of participants who tried to work together in Experiment 1, organized by knowledge condition and payoff condition. Error bars represent standard error.

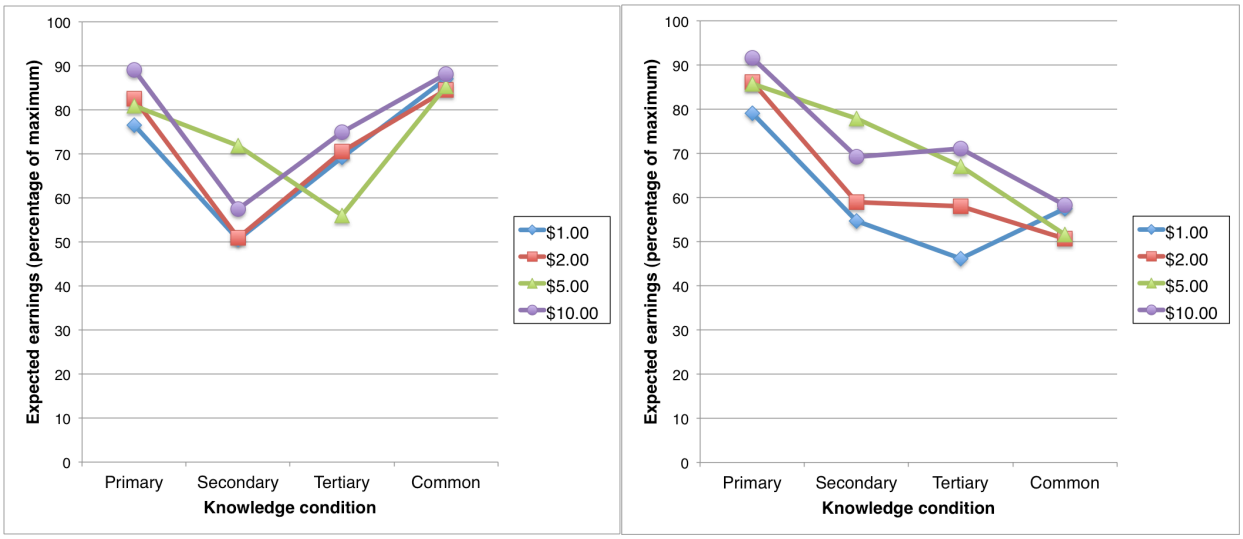


Figure 3. Average expected earnings as a percentage of maximum possible earnings for Experiment 1 (a) and Experiment 2 (b) by knowledge condition. We calculated *expected earnings* as the average amount a participant would earn across all possible pairings with the other participants.

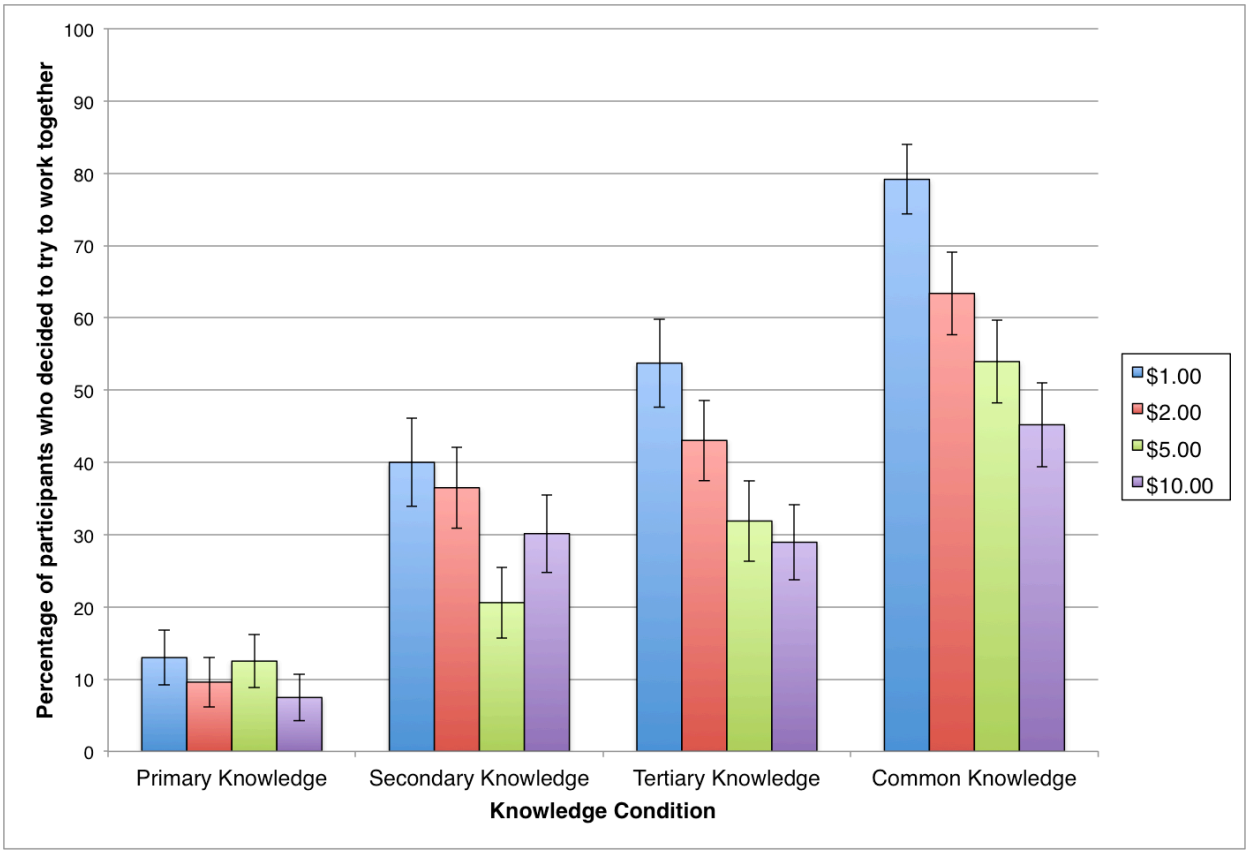


Figure 4. Percentage of participants who tried to work together in Experiment 2, organized by knowledge condition and payoff condition. Error bars represent standard error.

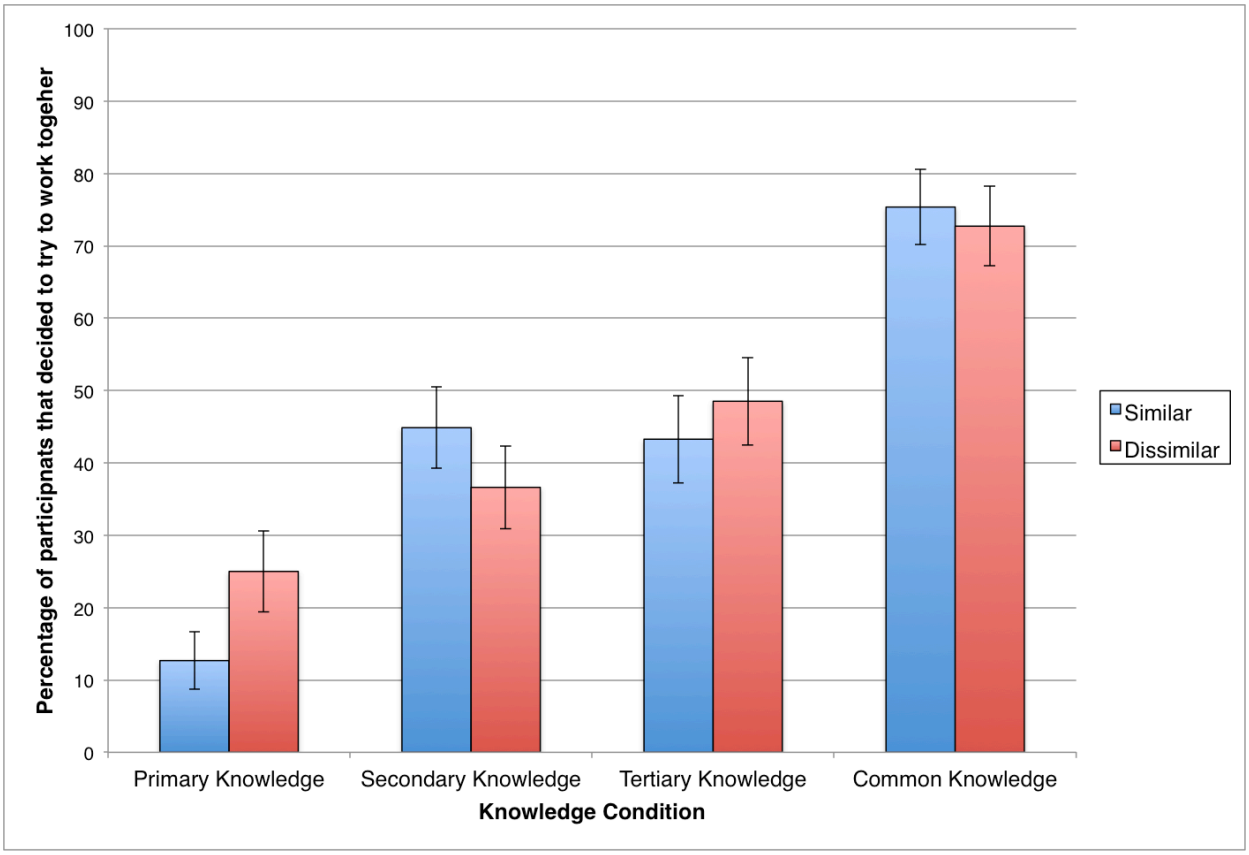


Figure 5. Percentage of participants who tried to work together in Experiment 3, organized by knowledge condition and similarity/dissimilarity condition. Error bars represent standard error.

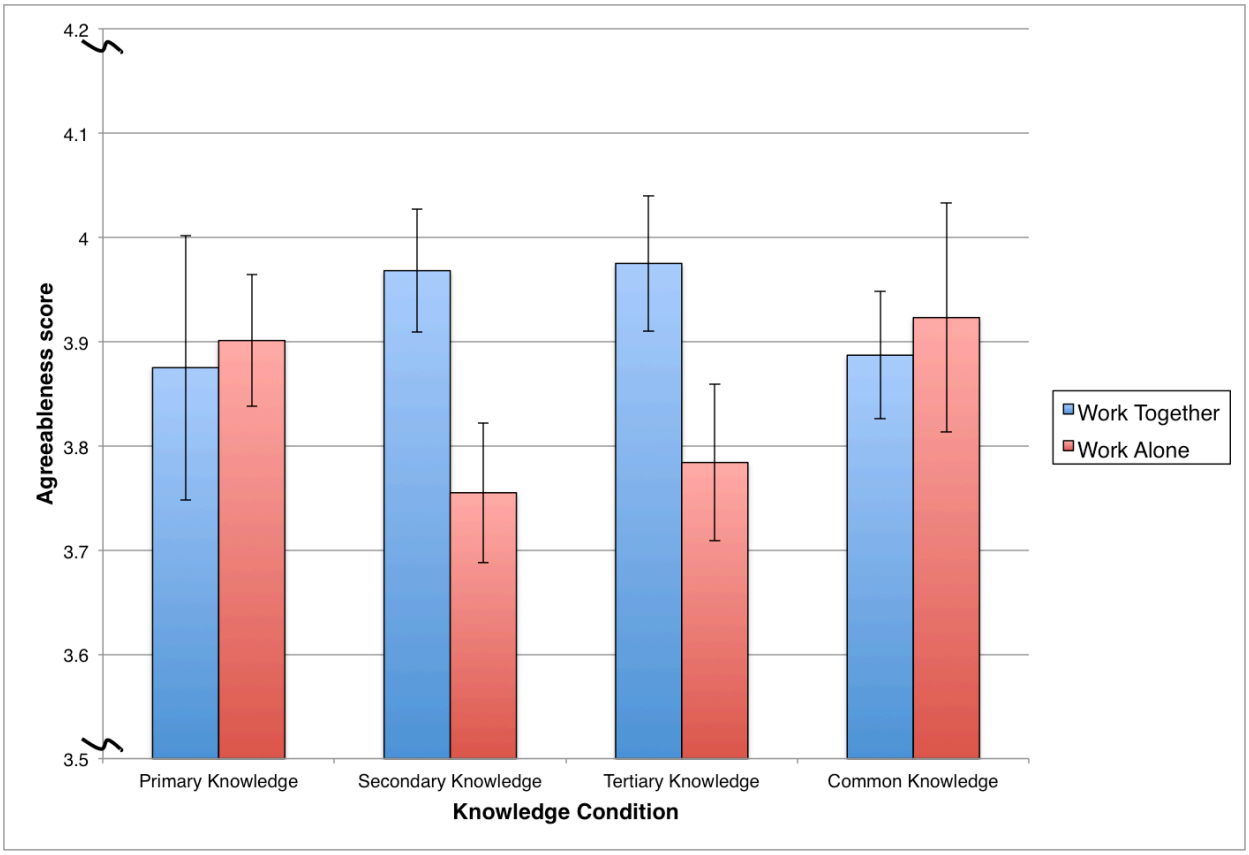


Figure 6. Average Agreeableness scale score for participants who tried to work together vs. participants who decided to work alone by knowledge condition in Experiment 3. The figure shows the abbreviated range of 3.5 – 4.2 (full range is 1 – 5). Error bars represent standard error.

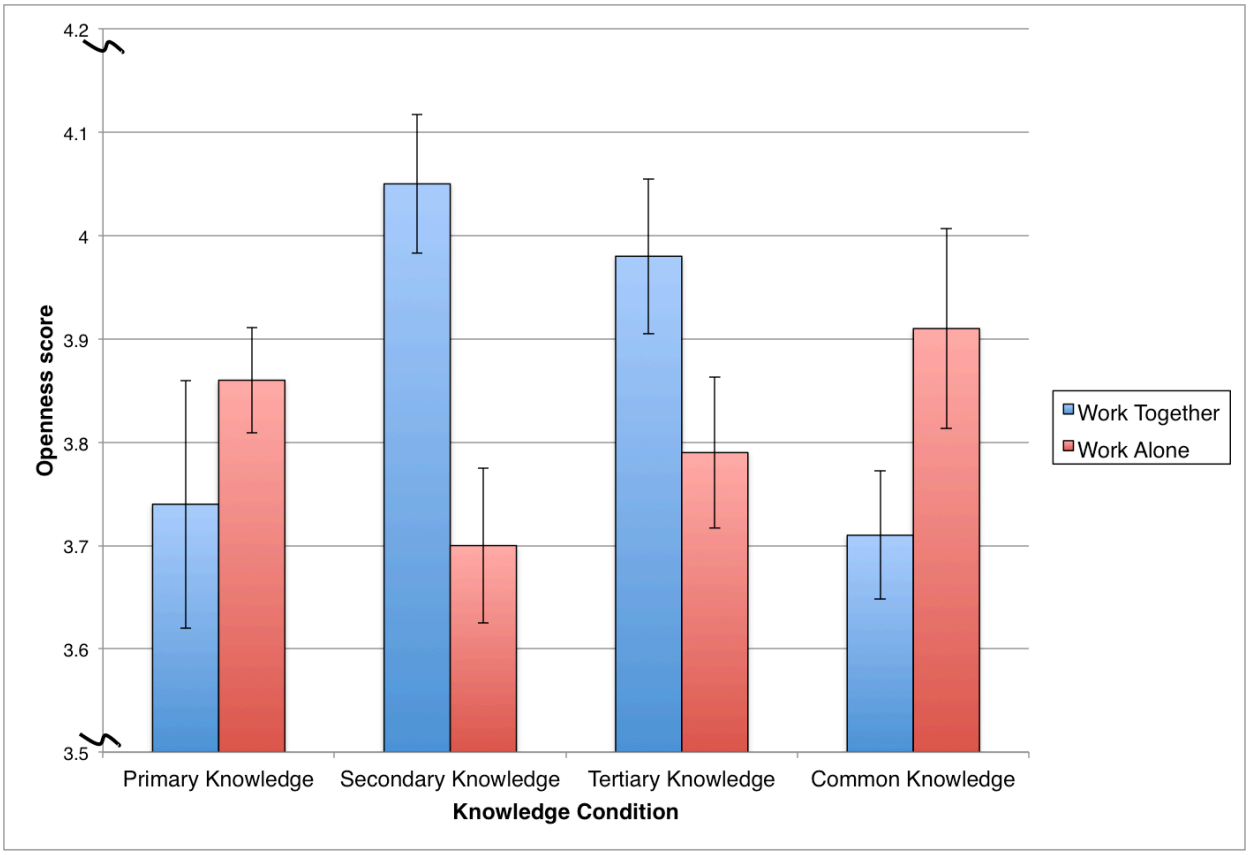


Figure 7. Average Openness scale score for participants who tried to work together vs. participants who decided to work alone by knowledge condition in Experiment 3. The figure shows the abbreviated range of 3.5 – 4.2 (full range is 1 – 5). Error bars represent standard error.