BRIEF REPORT

People believe each other to be selfish hedonic maximizers

Stefania De Vito · Jean-François Bonnefon

Abstract Current computational models of theory of mind typically assume that humans believe each other to selfishly maximize utility, for a conception of utility that makes it indistinguishable from personal gains. We argue that this conception is at odds with established facts about human altruism, as well as the altruism that humans expect from each other. We report two experiments showing that people expect other agents to selfishly maximize their pleasure, even when these other agents behave altruistically. Accordingly, defining utility as pleasure permits us to reconcile the assumption that humans expect each other to selfishly maximize utility with the fact that humans expect each other to behave altruistically.

Keywords Theory of mind · Altruism · Rationality · Utility · Pleasure

Whether it is called *mind reading*, *folk psychology*, or *theory of mind* (ToM), humans have a uniquely developed ability to reason about the beliefs and desires of others, in order to predict or explain their behavior. Understanding this ability is a key step in understanding our evolutionary trajectory (Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007; Whiten & Erdal, 2012), and computational models of ToM are accordingly found in biology (Yoshida, Dolan, & Friston, 2008), economics (Robalino & Robson, 2012), and psychology (Baker, Saxe, & Tenenbaum, 2009). Although these three fields formalize ToM in substantially different ways, all models appear to share a fundamental assumption, known as *mutual rationality attribution*. According to that assumption, people believe each other to selfishly maximize their expected utility.

In psychological studies and models of ToM, *utility* is usually indistinguishable from *personal gain*. Indeed, these studies and models commonly consider a single agent, whose utility is maximized by obtaining a personal benefit for the least possible cost or effort. For example, the experiments reported by Baker et al. (2009) feature a single individual navigating a maze in order to reach a goal object: Utility is maximized by reaching a goal object at a minimal energy cost. Comparably, the experiments reported in Scott and Baillargeon (2013) feature an agent reaching for one of two similar toys, one being more accessible than the other: Utility is maximized by reaching for the most accessible toy. Similarly, the examples introducing the rationality assumption in Gergely and Csibra's (2003) review article all feature a single agent, whose utility is typically maximized by walking the shortest path between a starting and a target location.

That utility should be indistinguishable from personal gain is arguably not an issue in situations that feature a single agent. Issues arise, though, as soon as an agent's action can impact another agent's outcome. In these situations, identifying utility with personal gains would be at odds with one of the best-established facts about how humans make their decisions: Humans do not only care about their personal gains, but also about the gains of others. For example, humans are much more likely than chimpanzees to display various forms of altruism, such as taking costly actions that confer economic benefits on other individuals, or preferring globally beneficial actions to actions that only benefit themselves (Fehr & Fischbacher, 2003; Horner, Carter, Suchak, & de Waal, 2011; Melis et al., 2011).

It would seem reasonable to expect that altruism, and other-regarding preferences in general, would be captured by ToM. To our knowledge, though, the literature on ToM has not addressed the important question of whether people expect

S. De Vito · J.-F. Bonnefon
Centre National de la Recherche Scientifique, Toulouse, France

J.-F. Bonnefon (✉)
CLLE, Maison de la recherche, Université de Toulouse le Mirail,
5 allées A. Machado, 31058 Toulouse Cedex 9, France
e-mail: bonnefon@univ-tlse2.fr

others to selfishly maximize personal gains or to maximize another utility function. Other literatures, though, can provide useful (if not always consistent) context for such an investigation.

First, research in behavioral economics has repeatedly shown that many people are willing to trust unknown partners in social dilemmas such as investment games or prisoners' dilemmas (for reviews, see Cooper & Kagel, 2014; Johnson & Mislin, 2011). If people believed others to maximize personal gains, they would be unlikely to trust strangers whose gains were maximized by abusing trust. Accordingly, behavioral economics suggests that people do not expect others to maximize personal gains.

Second, social psychology research has suggested that even though people might not act in a self-interested manner, they still expect others to act self-interestedly (e.g., Kruger & Gilovich, 1999; Miller, 1999). This literature on the norm of self-interest suggests that people are cynics who expect others to do whatever increases their benefits. Accordingly, research in social psychology (contra research in behavioral economics) suggests that people do expect others to maximize personal gains.

Third, research on reasoning has suggested yet another nuance to the issue of whether people expect others to be self-interested. This literature has shown that when people are asked to predict the actions of unknown fictive characters, they tend to assume a limited form of altruism in these characters. For example, people assume that fictive characters will do something that benefits others, as long as it does not cost them anything (Bonnefon, 2009; Bonnefon & Sloman, 2013; Thompson, Evans, & Handley, 2005). Accordingly, research on reasoning suggests that people expect others to maximize the gains of others, as long as their personal gains are not affected.

In sum, no conclusive body of data speaks to the question of whether people expect each other to maximize personal gains or maximize another utility function. In most psychological models of ToM, utility is indistinguishable from personal gains, and some research in social psychology indeed supports that equivalence. In contrast, research on reasoning and decision making suggests that people expect each other to care about more than their personal gains—while leaving open the question of the utility function that people believe each other to maximize.

In this article, we suggest that humans believe each other to be utility maximizers, but for a meaning of utility that allows for seemingly altruistic behavior. This meaning is actually the original meaning of "utility"—that is, pleasure (Bentham, 1789/1948). In other words, we claim that humans believe each other to be hedonic maximizers, who make choices that selfishly maximize their net affect (henceforth, we will use the terms "pleasure" and "net affect" interchangeably).

Hedonic maximization is compatible with apparent altruism. Indeed, a hedonic maximizer may take a costly action that confers a benefit to a recipient, if it feels better to take than not to take the action. Accordingly, we believe that it is possible for models of ToM to both capture other-regarding preferences such as altruism and incorporate the assumption of mutual rationality attribution. The key to this reconciliation is to define utility in terms of pleasure, rather than personal gain. Note that we do not assume total independence or mutual exclusiveness between pleasure and personal gains. What we assume is that sometimes the action that maximizes pleasure is not the same as the action that maximizes personal gains—and we believe that people expect others to take the action that maximizes their pleasure, rather than the action that maximizes their personal gains.

In the rest of this article, we report two experiment exploring the predictions derived from this claim. In the first experiment, we presented participants with minimal scenarios in which an actor could take an action, with various possible outcomes for the actor and a recipient, and measured both the perceived expected pleasures of the actor and recipient and the perceived likelihood that the actor would take action. In the second experiment, we presented participants with a different form of minimal scenario, in which an actor could take an action that could result in different degrees of pleasure for the actor or for a recipient. Participants' goal was to predict whether the actor would take the action. Information was missing, though, on the expected pleasure of either the actor or the recipient. We expected that participants would seek to know the actor's expected pleasure, but that they would not seek to know about the recipient's expected pleasure. Furthermore, we explored a possible moderator of this effect—namely, the relation between actor and recipient (friends or strangers).

In both experiments, the sample sizes were large enough for an analysis of variance (ANOVA) to detect a moderate effect with a 95 % chance. The numbers of participants were set in advance at 200 for Experiment 1 and 100 for Experiment 2, after which we excluded the few participants who did not return a full set of responses. No dependent variable was measured other than the ones mentioned in each Method section, and all experimental conditions are reported.

Experiment 1

Method

The participants were 198 adults (72 men, 126 women; mean age = 25 years) recruited through the Mechanical Turk online platform and directed to a Qualtrics online survey. Validation studies have repeatedly shown that Mechanical Turk samples were at least as representative of the U.S. population as are

traditional subject pools, and studies of judgment and decision making have returned the same results with Mechanical Turk samples as with traditional subject pools (Berinsky, Huber, & Lenz, 2012; Goodman, Cryder, & Cheema, 2013; Horton, Rand, & Zeckhauser, 2011; Paolacci, Chandler, & Ipeirotis, 2010).

Our participants were presented with one of four scenarios, in which an actor had an opportunity to take an action. The outcome of that action could be positive or negative for the actor, and positive or negative for the recipient (see Table 1).

After reading the scenario, participants rated the likelihood that the actor would experience four positive emotions (joy, pride, gratitude, and relief) and four negative emotions (sadness, anger, disappointment, and guilt) if the action was taken, and the likelihood that the actor would experience these same emotions if the action was not taken. Participants gave their responses by adjusting the position of a slider on a line anchored at *extremely unlikely* on the left and *extremely likely* on the right. The position of the slider was coded as a decimal number between 0 and 1. All measures appeared in randomized order.

The expected net affect of the actor was then computed, for each participant, as a function of these 16 emotional ratings. Specifically, we first computed the weighted sum of the likelihoods of all emotions, with a weight of +1 for positive emotions resulting from action and negative emotions resulting from inaction, and a weight of –1 for negative emotions resulting from action and positive emotions resulting from inaction. We then scaled this weighted sum to a number between 0 and 1, which served as our index of the expected net affect of the actor. The principle underlying this coding system was to capture the *net* affect resulting from acting, as compared to not acting (hence, the weights of +1 and –1 for positive emotions resulting from acting and positive emotions resulting from not acting, respectively; negative emotions were reverse-coded), and to turn this weight into an *expected* net affect by multiplying each emotion weight by the likelihood of experiencing this emotion.

This process was repeated for the recipient, in order to compute the expected net affect of the recipient. Finally, participants rated the likelihood that the actor would take the action, using the same slider as for the other questions. In sum, we manipulated the actor's and recipient's outcomes, computed the actor's and recipient's expected affects, and measured the perceived likelihood of the actor taking action.

Because we described the outcomes in general terms ("better off" and "worse off"), participants were likely to interpret these outcomes as a mixture of personal gains and affect. Ideally, we would like to show that participants expected the actors to selfishly maximize their expected pleasure, but not to selfishly maximize the mixture of personal gains and pleasure implied by a descriptor such as "better off." In operational

terms, we needed to conduct two analyses. In the first analysis, we sought to show that both the actor's and the recipient's outcomes influenced the likelihood of action (i.e., that participants did not expect actors to selfishly maximize their outcomes). In the second analysis, we sought to show that only the actor's expected affect, not the recipient's, would affect the likelihood of action (i.e., that participants expected actors to selfishly maximize their pleasure).

Results

Figure 1 displays the average expected net affects of the actor and the recipient, in each situation, as well as the perceived likelihood that the actor would take the action. People took into account both the actor's and the recipient's outcomes when making a prediction about the actor's decision. Indeed, an ANOVA showed that the likelihood of action was predicted by the actor's outcome, $F(1, 194) = 175.8$, $p < .001$, $\eta^2 = .47$, but also by the outcome for the recipient, $F(1, 194) = 17.3$, $p < .001$, $\eta^2 = .08$, and their interaction, $F(1, 194) = 7.6$, $p = .006$, $\eta^2 = .04$.[1] This interaction effect seems to mostly capture the difference between mutualism and selfishness: Actors were seen as being highly likely (83 % chance) to take an action that benefited both themselves and a recipient, but less likely (61 % chance) to take an action that benefited themselves while hurting a recipient.

So far, these data suggest that participants took into account both the actor's and the recipient's outcomes. However, according to our main claim, participants should expect the actors to be selfish hedonic maximizers. That is, they should predict the likelihood of actions on the basis of the actor's expected pleasure, and the actor's only. To test this prediction, we regressed the perceived likelihood of action on the expected pleasure of the actor and that of the recipient.

Figure 2 displays the partial regression plots reflecting the relation between the perceived likelihood of action and its two predictors. The perceived likelihood of action was uniquely predicted by the expected pleasure of the actor ($\beta = .72$, $p < .001$), whereas the expected pleasure of the recipient was not a significant predictor ($\beta = -.06$, $p = .28$). Our main prediction is thus grounded in the data. Participants expected others to be selfish hedonic maximizers, who optimized their own pleasure without consideration for the pleasure of others.

---

[1] The expected pleasure of the actor and that of the recipient showed a similar sensitivity to both payoffs. Regarding the expected pleasure of the actor, the ANOVA detected an effect of the actor's payoff, $F(1, 194) = 167.7$, $p < .001$, $\eta^2 = .46$; an effect of the recipient's payoff, $F(1, 194) = 63.8$, $p < .001$, $\eta^2 = .25$; but no interaction effect, $F(1, 194) = 2.3$, $p = .13$, $\eta^2 = .01$. Regarding the expected pleasure of the recipient, the ANOVA detected an effect of the actor's payoff, $F(1, 194) = 14.0$, $p < .001$, $\eta^2 = .06$; an effect of the recipient's payoff, $F(1, 194) = 395.6$, $p < .001$, $\eta^2 = .67$; and an interaction effect, $F(1, 194) = 9.2$, $p = .003$, $\eta^2 = .05$.

Table 1 Text of the four scenarios used in Experiment 1

	Recipient Worse Off	Recipient Better Off
Actor worse off	**Spite.** Mike and Bob are in a situation where, if Mike takes action, both Mike and Bob will be worse off.	**Altruism.** Mike and Bob are in a situation where, if Mike takes action, Mike will be worse off and Bob will be better off.
Actor better off	**Selfishness.** Mike and Bob are in a situation where, if Mike takes action, Mike will be better off and Bob will be worse off.	**Mutualism.** Mike and Bob are in a situation where, if Mike takes action, both Mike and Bob will be better off.

The labels "Spite," "Selfishness," "Altruism," and "Mutualism" were not shown to participants.

Experiment 2

In the second experiment, we attempted to conceptually replicate the main result of Experiment 1 using a different approach. We wanted to show that participants would deem it useful to know about the expected pleasure of the actor, but not about the expected pleasure of the recipient, when trying to predict whether the actor would take action. Furthermore, we explored a possible moderator of this effect, by telling participants that the actor and the recipient were either friends or strangers.

Method

The participants were 91 adults (mean age = 33 years) recruited through the Mechanical Turk online platform and directed to a Qualtrics online survey. They were presented with scenarios in which an actor had an opportunity to take a certain action that would affect both the actor and the recipient.

Half of the participants were told that the actor and the recipient were close friends (the *friends* condition), and the other half were told that the actor and the recipient did not know each other well (the *strangers* condition). In both groups, participants were asked about the empathy between the actor and the recipient, by indicating how likely it was that Tom and Bob felt good and felt bad about the same things. Participants responded by adjusting the position of a slider on a line anchored at *extremely unlikely* on the left and *extremely likely* on the right. The position of the slider was coded as a decimal number between 0 and 1.

Depending on the scenario, participants were informed about how the actor or the recipient would feel as a consequence of the action. These feelings could be positive or negative (see Table 2).[2] After reading each scenario,
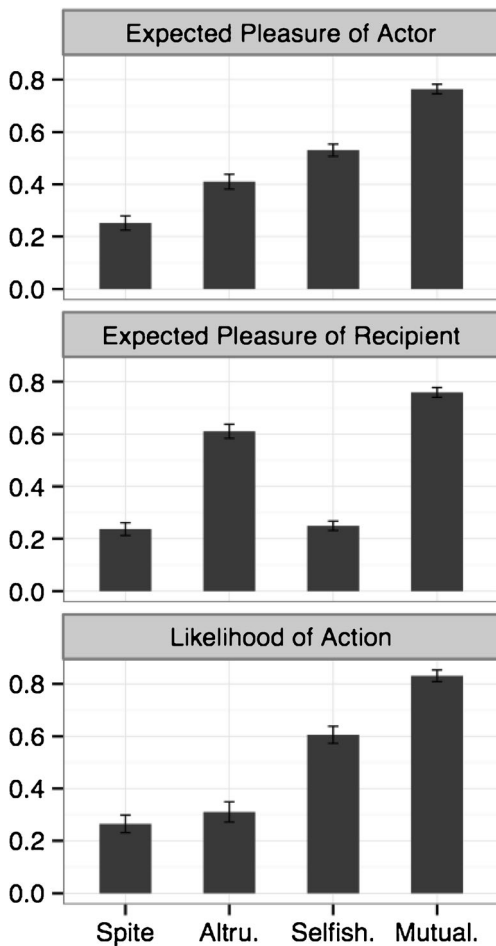

Fig. 1 Participants were presented with minimal scenarios in which the actor could take a spiteful, altruistic, selfish, or mutually beneficial action. In each case, participants provided 32 ratings of expected emotions, from which we computed the expected pleasure of the actor (*top*) and the expected pleasure of the recipient (*middle*). Participants also rated the likelihood that the actor would take the action (*bottom*)

[2] To make sure that these descriptors ("feel good" and "feel bad") exclusively implied pleasure, and not the same mixture of pleasure and personal gains as in Experiment 1, we conducted a manipulation check on the Mechanical Turk, drawing on a sample of 38 participants (27 men, 11 women; mean age = 38 years). These participants were asked about the meanings of "feel good" and "feel bad." The question prompt about "feel good" read: "In the sentence 'Mike and Bob are in a situation where if Mike takes action, Bob will feel good', the expression 'feel good' means that. . . ." The responses were "Bob will experience a personal gain" (10-point scale from *absolutely not* to *absolutely*) and "Bob will experience a positive emotion" (10-point scale from *absolutely not* to *absolutely*). Participants judged that to "feel good" meant to experience a positive emotion ($M = 9.4$, $SE = 0.2$) much more than a personal gain ($M = 3.8$, $SE = 0.5$), $t(37) = 11.0$, $p < .001$. They also judged that to "feel bad" meant to experience a negative emotion ($M = 9.2$, $SE = 0.3$) much more than a personal loss ($M = 2.8$, $SE = 0.5$), $t(37) = 11.4$, $p < .001$.
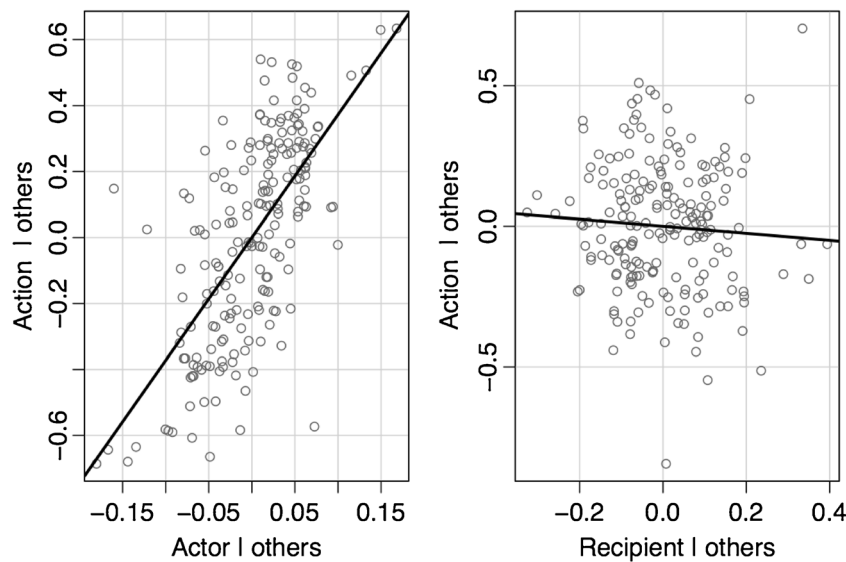
Fig. 2 Partial regression plots displaying the results of the multiple regression of the perceived likelihood of action on the expected pleasures of the actor (*left*) and the recipient (*right*)

participants indicated how useful it would be to know how the other agent would feel as a result of the action, in order to predict whether the actor would take the action.

As always, participants gave their responses by adjusting the position of a slider on a line, this time anchored at *not at all useful* on the left and *extremely useful* on the right. The position of the slider was coded as a decimal number between 0 and 1.

Results

We conducted a $2 \times 2 \times 2$ ANOVA in which the predictors were the relation between the agents (friends or strangers), the valence of the known hedonic outcome (good or bad), and the agent about whom information was missing (actor or recipient). The dependent variable was the usefulness of obtaining the missing information, in order to predict whether the actor would take action.

The analysis detected two main effects and an interaction. First, and in line with the results of Experiment 1, participants

indicated that information about the actor's hedonic outcome was more important (.69) than information about the recipient's hedonic outcome (.41), $F(1, 89) = 70.9$, $p < .001$, $\eta^2 = .40$. This difference was moderated by the relation between the two agents, as detected by an interaction effect, $F(1, 89) = 18.5$, $p < .001$, $\eta^2 = .10$. The analysis also detected a main effect of the relation between the two agents, $F(1, 89) = 17.2$, $p < .001$, $\eta^2 = .16$, whose interpretation seems unnecessary in the light of the interaction effect. The analysis did not reveal any other effect, all Fs < 1.5, all ps > .23

As is shown in Fig. 3, the relation between the two agents does not make a difference when information is missing about the hedonic outcome of the actor (this information is always considered useful). The relation between the two agents makes a large difference when information is missing about the hedonic outcome of the recipient. This information is considered useless when the agents are strangers, and useful when the agents are friends.

Empathy ratings allow for a better understanding of this effect. Expectedly, the empathy ratings were higher in the friends than in the strangers condition, $t(90) = 4.1$, $p < .001$.

Table 2 Text of the four scenarios used in Experiment 2

	Informed About Actor	Informed About Recipient
Positive feelings	Tom and Bob are in a situation where, if Tom takes action, Tom will feel good.	Tom and Bob are in a situation where, if Tom takes action, Bob will feel good.
Negative feelings	Tom and Bob are in a situation where, if Tom takes action, Tom will feel bad.	Tom and Bob are in a situation where, if Tom takes action, Bob will feel bad.
Your goal is to predict whether Tom will take this action. To make this prediction, how useful is it to know . . .	whether Bob would feel good or bad as a result?	whether Tom would feel good or bad as a result?

In addition, participants in the friends condition were told that Tom and Bob were good friends, and participants in the strangers condition were told that Tom and Bob did not know each other much.
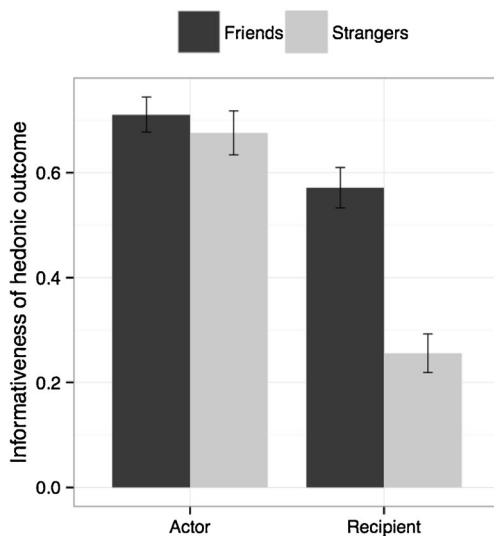
Fig. 3 Informativeness of the hedonic outcomes of actor and recipient, as a function of the relation between the two agents. For predicting whether the actor will take action, it is always useful to know how the actor will feel in case of action, but knowing how the recipient will feel is useful only if the two agents are friends

As is shown in Fig. 4, they were also correlated with the perceived usefulness of knowing about the recipient's hedonic outcome. The correlation was .42 ($p < .001$) when the outcome was negative for the actor, and .23 ($p = .03$) when the outcome was positive for the actor.

These results suggest that the reason why it was useful to know how the recipient would feel in the friends condition was because empathy was high in this condition between the actor and recipient. In other words, knowing about the hedonic
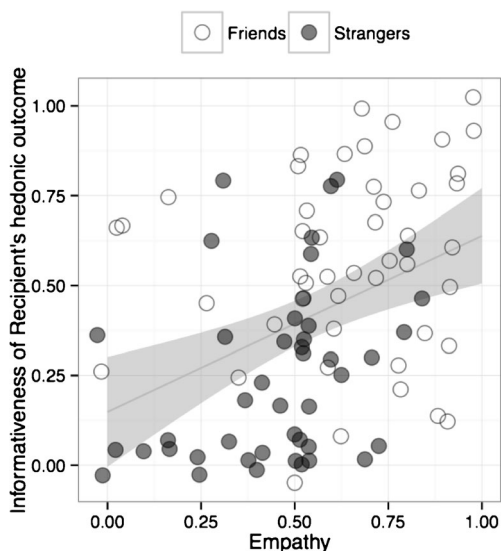

Fig. 4 The greater the empathy between actor and recipient, the more useful it is to know how the recipient will feel, in order to predict whether the actor will take action. Note that the empathy ratings tend to cluster to the right in the friends condition

outcome of the recipient was deemed useful in the friends condition because it provided indirect information about how the actor would feel.

In sum, Experiment 2 provided additional evidence for the claim that people believe each other to be selfish hedonic maximizers. Our key measure was whether participants would think it useful to know how the recipient would feel, in order to predict whether the actor would take action. In the default situation in which the actor and recipient did not know each other, participants did not think it was useful to know how the recipient would feel. In the situation in which the actor and recipient were friends, participants believed that it was useful to know how the recipient would feel, but only to the extent that it gave them additional information about the feelings of the actor.

General discussion

Current psychological models of theory of mind typically assume that humans believe each other to maximize utility. This begs the question of which utility function people believe each other to maximize. In existing models, utility is indistinguishable from personal gains, mostly because these models only feature single-agent situations (e.g., Baker et al., 2009; Gergely & Csibra, 2003; Scott & Baillargeon, 2013). We have argued that as soon as another agent enters the picture, defining utility as personal gains is at odds with well-established facts about human altruism. Our results confirmed that people did not believe others to maximize their personal gains.

Most importantly, though, our results showed that people did expect each other to maximize utility, in the hedonic sense. Thus, we are in a position to rephrase the assumption of mutual rationality attribution featured in models of ToM. Our new, hedonic version of mutual rationality attribution can be phrased thus: People believe each other to selfishly maximize their expected pleasure. This hedonic version is compatible with existing models, because pleasure and personal gains are (presumably) strongly correlated in the single-agent situations that existing models address. The hedonic version of mutual rationality attribution becomes especially helpful when addressing multiagent situations such as the one featured in our experiments, rather than the single-agent situations usually featured in the psychological literature.

We acknowledge that our scenarios are only a first step toward full-blown strategic contexts. Our scenarios actually resemble dictator games, in that the recipient is totally passive. Although we believe that these simple two-agent situations are already interesting, and easily encountered in everyday life, further research will be needed to generalize our conclusion to complex strategic contexts. An intriguing possibility is that

the participants in our experiments might not have considered the mental states of the recipient because the actor himself was not required to engage ToM when making his decision. Strategic contexts in which actors must carefully consider the emotions of recipients might prompt participants themselves to consider these emotions. Moving toward more complex contexts might also provide opportunities to orthogonally manipulate pleasure and personal gains, and to assess their respective impacts in situations of mutual exclusiveness.

Beyond their specific contributions to modeling, our conclusions are of potential interest to the whole range of fields investigating ToM within and outside psychology. First, new results about ToM in healthy adults will inform accounts of ToM development (Apperly, Samson, & Humphreys, 2009). Specifically, research in developmental and differential psychology will be required in order to learn when and how individuals start seeing each other as selfish hedonic maximizers, and whether this tendency can be viewed as a heuristic shortcut.

Second, social psychology and sociology have long struggled with the question of whether self-interested behavior can be considered a social norm (Miller, 1999; Spillman & Strand, 2013). Our results suggest reframing the question as whether the selfish pursuit of pleasure, rather than personal gain, can be considered a social norm. The pursuit of pleasure might well prove less historically and culturally contingent than the pursuit of personal gains. A social norm of pleasure maximization might, accordingly, prove to be a better description of social attitudes, and a better tool for eliciting altruistic behavior (Ratner & Miller, 2001).

Finally, in addition to their interest for computational ToM in biology and economics, our results can inform the long-standing philosophical controversy between psychological altruism and psychological egoism (Slote, 2011; Sober & Wilson, 2009; Steene, 2009). Psychological egoism is the view according to which seemingly altruistic actions are motivated by a selfish reward, typically a positive emotion. Psychological altruism is the view according to which at least some actions are not motivated by such a selfish reward. Our results do not compare in sophistication to the arguments that have been advanced for both views—but they do suggest that laypersons appear to agree with psychological egoism. Indeed, the participants in our study expected others to selfishly maximize pleasure, even when their actions had the appearance of altruism.

In sum, we have reported the first empirical resolution of a vexing question within and outside the field of ToM. Our results clarify that people expect each other to maximize their pleasure, without expecting others to maximize their personal gains. That is, people may very well expect others to sacrifice their resources for the welfare of friends or strangers—but apparently only to the extent that it feels good to do so.

References

Apperly, I. A., Samson, D., & Humphreys, G. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology, 45,* 190–201.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113,* 329–349.

Bentham, J. (1948). An introduction to the principle of morals and legislation. Oxford, UK: Blackwell. (Original work published in 1789)

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

Bonnefon, J. F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-theoretic leakage. *Psychological Review, 116,* 888–907.

Bonnefon, J. F., & Sloman, S. A. (2013). The causal structure of utility conditionals. *Cognitive Science, 37,* 193–209.

Cooper, D., & Kagel, J. (2014). Other regarding preferences: A survey of experimental results. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (Vol. 2). Princeton, NJ: Princeton University Press, (in press).

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425,* 785–791.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7,* 287–292. doi:10.1016/S1364-6613(03)00128-1

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

Herrmann, E., Call, J., Hernàndez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317,* 1360–1366.

Horner, V., Carter, J. D., Suchak, M., & de Waal, F. B. M. (2011). Spontaneous prosocial choice by chimpanzees. *Proceedings of the National Academy of Sciences, 108,* 13847–13851.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 4,* 399–342.

Johnson, N., & Mislin, A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology, 32,* 865–889.

Kruger, J., & Gilovich, T. (1999). "Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 76,* 743–753.

Melis, A. P., Warneken, F., Jensen, K., Schneider, A. C., Call, J., & Tomasello, M. (2011). Chimpanzees help conspecifics obtain food and non-food items. *Proceedings of the Royal Society B, 278,* 1405–1413.

Miller, D. T. (1999). The norm of self-interest. *American Psychologist, 54,* 1053–1060.

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

Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology, 81,* 5–16.

Robalino, N., & Robson, A. (2012). The economic approach to "theory of mind." *Philosophical Transactions of the Royal Society B, 367,* 2224–2233.

Scott, R. M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological Science, 24,* 466–474.

Slote, M. (2011). Egoism and emotion. *Philosophia, 41,* 313–335.

Sober, E., & Wilson, D. S. (2009). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Spillman, L., & Strand, M. (2013). Interest-oriented action. *Annual Review of Sociology, 39,* 85–104.

Steene, M. (2009). Why everyone acts altruistically all the time: What parodying psychological egoism can teach us. *Philosophia, 39,* 563–570.

Thompson, V. A., Evans, J., St, B. T., & Handley, S. J. (2005). Persuading and dissuading by conditional argument. *Journal of Memory and Language, 53,* 238–257.

Whiten, A., & Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society B, 367,* 2119–2129.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology, 4,* e1000254.