Effort and Performance: What Distinguishes Interacting and Noninteracting Groups from Individuals?

Tibor Besedeš,* Cary Deck,† Sarah Quintanar,‡ Sudipta Sarangi,§ and Mikhail Shor‖

We study how group membership affects behavior both when group members can and cannot interact with each other. Our goal is to isolate the contrasting forces that spring from group membership: a free-riding incentive leading to reduced effort and a sense of social responsibility that increases effort. In an environment with varying task difficulty and individual decision making as the benchmark, we show that the free-riding effect is stronger. Group members significantly reduce their effort in situations where they share the outcome but are unable to communicate. When group members share outcomes and can interact, they outperform groups without communication and individuals. We show that these groups do as well as the best constituent member would have done on his or her own.

JEL Classification: C92, D71, Z13

1. Introduction

Economists have recently started paying more attention to group decision making as many economic decisions from the family dinner table to the corporate boardroom reflect the opinions of groups (see recent surveys by Charness and Sutter 2012 and Kugler, Kausel, and Kocher 2012). Within a group, individuals offer input into the decision and collectively share in the resulting outcome. Despite the conventional wisdom that two heads are better than one, researchers are more equivocal about the ability of groups to make better decisions. For

* Georgia Institute of Technology, School of Economics, Georgia Institute of Technology, Atlanta, GA 30332-0615, USA; E-mail besedes@gatech.edu.

† University of Arkansas and Economic Science Institute, Chapman University, 425 WCOB, Department of Economics, University of Arkansas, Fayetteville, AR 72701, USA; E-mail cdeck@walton.uark.edu.

‡ University of Arkansas–Little Rock, Department of Economics and Finance, University of Arkansas–Little Rock, 2801 South University Ave., Little Rock, AR 72204, USA; E-mail smquintanar@ualr.edu; corresponding author.

§ Louisiana State University; DIW Berlin; and NSF, Department of Economics, 2300 BEC, Louisiana State University, Baton Rouge, LA 70803, USA; National Science Foundation, Arlington, VA 22230, USA; DIW Berlin, 10117 Berlin, Germany; E-mail sarangi@lsu.edu.

‖ University of Connecticut, Department of Economics, University of Connecticut, 341 Mansfield Road, Storrs, CT 06269-1063, USA; E-mail mike.shor@uconn.edu.

example, team-managed mutual funds do no better, and sometimes worse, than funds managed by individuals (Prather and Middleton 2002; Chen et al. 2004). Conversely, others have documented "assembly bonus effects" where groups outperform even their most capable members (Laughlin, Bonner, and Miner 2002). Even without interaction, group membership and the interdependence of members' payoffs can, in themselves, alter individual decision making (Charness, Rigotti, and Rustichini 2007; Sutter 2009). Thus, group performance depends on both the effort individuals bring *to* the group and the interaction *within* the group that enables its members to make a collective decision. The goal of this article is to begin to disentangle these two effects.

Individuals often exert a different level of effort when making decisions as part of a group versus for themselves. Because personal responsibility for decisions is diluted in a group setting, members may free ride (in the microeconomic sense) or engage in "social loafing" by reducing personal effort when part of a group (Latané, Williams, and Harkins 1979; Karau and Williams 1993).[1] Conversely, as one's decisions impact the payoffs of other group members, altruism, social pressures, shared responsibility, social identity, and group salience may lead to increased effort (Wagner 1995; Charness, Rigotti and Rustichini 2007; Sutter 2009). This positive aspect of group membership may be referred to as its social responsibility effect. Thus, group membership may induce two opposing forces on an individual's provision of effort.

Whatever effort each group member brings, the group translates individual problem-solving approaches into a single collective action. Some groups are able to identify the member with the greatest task-specific expertise (Henry 1993). In other instances, groups create knowledge, resulting in a strategy superior to what any member could obtain alone. For example, Charness, Karni, and Levin (2007, 2010) find that interaction improves the likelihood of correctly answering questions concerning stochastic dominance and conjunctive events. On the other hand, Tindale et al. (1996) find that groups often favor intuitive but incorrect answers in tasks requiring an understanding of probability. We employ an experimental design that allows us to observe the effect of group membership independent of the effects of interaction for a specific type of multistate choice problem. Past research includes experiments in which groups make a joint decision[2] and experiments where subjects are members of groups but do not make a joint decision.[3] The former is concerned with the effect of groups on collective behavior, while the latter is concerned with the effect of group membership on individual actions. Similar to Sutter (2009), our experiment includes both types of groups.

Our subjects participate either as individuals or in one of two group treatments in a series of choice tasks. The tasks are context free but may be thought of as selecting an insurance plan (option) from among several that cover (pay a fixed prize in) some eventualities (states of the world) but not others. The probability of each eventuality is provided to subjects. Therefore, options can be easily and objectively ranked based on each option's probability of payment and irrespective of a respondent's risk attitude. Selecting the optimal option thus requires only

---

[1] We define a situation where an individual provides lower effort expecting others to provide more effort toward a better common outcome as one characterized by free riding. The social dilemmas literature on the other hand considers a situation to exhibit free riding only if an individual generates negative spillovers in the pursuit of their own self-interest without having to pay for these externalities. Of course, our notion of free riding in the social dilemmas literature might be explained as the lack of other-regarding preferences.

[2] Such as Blinder and Morgan (2005, 2008), Cooper and Kagel (2005), Kocher and Sutter (2007), and Sutter (2009).

[3] Such as Charness, Rigotti, and Rustichini (2007), Chen and Li (2009), Hargreaves-Heap and Zizzo (2009), Sutter (2009), Charness, Karni, and Levin (2010), Ioannou, Qi, and Rustichini (2011), and Morita and Servátka (2011).

(cognitive) effort in the form of calculating and comparing the probabilities of receiving payment associated with each option. In the *individual* treatment, subjects make decisions and earn payments on their own. In the *interacting groups* treatment, subjects complete the task in groups of three, engaging in free-form face-to-face discussion. Each group makes a joint decision, and members earn identical resulting payments. In the *noninteracting groups* treatment, subjects are placed in groups of three but make individual decisions without any communication with other group members. Decisions of a single group member, selected at random, determine each group member's identical payment. Groups with interaction have both payoff commonality and joint decision making, while noninteracting groups have payoff commonality but individual decision making.[4]

Observe that comparing individual and interacting groups is quite natural to find out if free riding matters in the interacting groups. However, it is challenging to isolate the effects of free riding from the benefits of collaboration in this setting. By introducing noninteracting groups and comparing them to individuals, we can see the effect of free riding by comparing the change in behavior as task difficulty varies. While the social responsibility remains the same for easy and hard tasks, it is costlier to solve a hard task providing a strong incentive for free riding. Finally, note that the comparison of noninteracting groups and interacting groups provides evidence of the benefit of collaboration while allowing for free riding to vary.

While the individual and interacting group treatments have many obvious parallels outside of the lab, the noninteracting group treatment does not. This ability to create counterfactual situations is a major advantage of laboratory experiments. Even when those counterfactuals do not mimic a real-world example, they enable the dissection of hypotheses that may not be possible otherwise. Our somewhat artificial noninteracting groups allow us to address two questions. First, how does group membership, in itself, influence individual effort as measured by performance in the absence of interaction? Second, how does interaction within groups affect effort and the optimality of decisions controlling for the commonality of payoffs?

We report three main results. First, members of noninteracting groups engage in free riding, resulting in a loss of $1.20 on a $20 task payoff. They perform slightly worse than subjects in the individual treatment across all tasks, making an optimal decision in 67% of tasks as compared to 72% for individuals. However, as task complexity increases, raising the cost of (cognitive) effort, the performance disparity between noninteracting groups and individuals widens, with noninteracting groups making an optimal decision in less than half of the most complex tasks (47%), while individuals do so in two-thirds of these tasks (65%). Free riding is primarily observed among men with no significant free riding observed for women.[5] We conjecture that the uncertainty about whether a member's effort will impact payoffs increases free-riding tendencies. Each member's effort pays off with a constant 1/3 probability, thus

---

[4] There are potentially several other differences between the two types of groups as implemented in the lab. For example, each interacting group moves to a separate room whereas noninteracting group members make decisions in the same room as people in other noninteracting groups and those making the decisions individually. Interacting groups may reduce social distance as group members could discuss nontask-related items. It is certainly possible to operationalize the two types of groups in different ways, something that could be investigated in future work.

[5] Conclusions on the role of sex in free riding based on public goods experiments have been mixed. Nowell and Tinker (1994) report more free riding by women, while Brown-Kruse and Hummels (1993) report the opposite and Cadsby and Maynes (1998) find no difference.

preserving the benefit of investing more effort, while the cost of providing effort increases with task complexity. This free-riding effect appears to outweigh the social concerns created by payoff commonality.

Our second result is that interacting groups do as well as their best individuals, but not better, making optimal decisions in 87% of the tasks. The fact that they are able to interact makes them effective aggregators of information rather than knowledge creators for this type of a problem. Given that they are all fully responsible for making decision as opposed to having the responsibility one-third of the time, we find that there is no free riding as task difficulty varies. Not surprisingly, this suggests that payoff commonality and the social concerns it involves are important as long as they do not require very costly effort.

Our third result relates to the saliency of group membership. Charness, Rigotti, and Rustichini (2007) show that groups affect strategic decision making when group membership is made sufficiently salient through payoff dependence or observation of play by group members. Sutter (2009) extends those results to a nonstrategic setting and finds that individuals who are part of groups but cannot communicate yield similar decisions to those achieved by interacting groups. These studies show that the performance of noninteracting groups depends on the level of group saliency, which is induced by common payoff. Our results qualify this conclusion in the sense that payoff commonality on its own may not be sufficient to induce group saliency, or alternatively, that the nature of payoff commonality is important.

Our noninteracting groups differ from those in Charness, Rigotti, and Rustichini (2007) and Sutter (2009) in the way individual decisions translate into group outcomes. In Charness, Rigotti, and Rustichini (2007) each individual receives a payment resulting from his or her actions as well as a third of the payment received by all other members of his or her group. Group members take turns ''playing'' the game, and in some treatments those members who are not actively playing are able to observe their active member. Members of noninteracting groups observe the active member's choices, though they are unable to participate. This presents an obvious social responsibility incentive for the participating member to perform well and is used as a way to vary group saliency. Though Sutter (2009) is likely the most similar to our design, the payoff structure for group members is very different. Each group member is solely responsible for one-third of all decisions with all group members' performances summed to arrive at the group's payment.

In our experiment each member of a noninteracting group makes every decision. One randomly chosen member's decision is solely responsible for the entire group outcome, while other group members' decisions are undisclosed. Our goal is to understand free riding by varying the difficulty of the task that group members face. Hence, in our design it is necessary for all group members to tackle all the tasks. Compared to dividing tasks among group members, our design changes the balance of incentives between free riding and social responsibility in ways that are difficult to identify otherwise. Experiments that make each group member indispensable and identifiable by design will eliminate or substantially reduce the incentive to free ride. Our probabilistic design preserves incentives for free riding alongside the possibility of identification and ex ante indispensability in noninteracting groups.

An important aspect of our experimental design is that we compare group and individual performance on an intellective, nonstrategic task where choices can be objectively ranked from best to worst, at least for standard models. A number of past studies on group decision making have used judgmental tasks involving a strategic setting in which decision optimality depends

on beliefs about other players[6] or a task in which decision optimality depends on idiosyncratic personal traits.[7] In contrast, choices in our experiment are invariant to personal traits and require only that subjects prefer more money to less. Our effort departs from studies that examine nonstrategic play such as Gillet, Schram, and Sonnemans (2009)[8] and Sutter (2009) in one important dimension: *Our design makes it possible to vary the difficulty of the task by changing the number of options and the number of states describing each option.* It is precisely this variability that allows us to examine the balance between the free riding and social responsibility forces by allowing us to increase the effort required to solve the task while preserving the benefit from solving it.

Our results suggests that, statistically speaking, the best member's likely outcome serves as an upper bound of what an interacting group can achieve through joint effort, which is far better than what the group would do in the absence of interaction. This implies that groups neither create knowledge nor, on average, suppress the most superior problem-solving approaches. Taken together, our results suggest that payoff commonality is insufficient on its own to make group membership salient or, alternatively, that free riding can be a stronger incentive than that offered by group saliency. Yet, when groups interact, they can effectively identify and adopt the problem-solving approach of their strongest members.

2. Experimental Design and Procedures

Our experiment consists of either individuals or groups completing a series of decision tasks in a task booklet. In every task there are a number of mutually exclusive states that occur with known probability. Subjects choose among a set of options where an option covers a given set of states. The tasks are identical to those used by Besedeš et al. (2012b) in their study of individual decision making among the elderly. Figure 1 illustrates a task with four options, denoted A, B, C, and D. Options differ in the states they cover and no two options cover identical states. States are denoted and presented as 100 colored beads to be drawn from an urn. In Figure 1, there are 8 lime, 36 pink, 45 white, and 11 green beads. After all subjects complete their tasks, the task to be used for payment is randomly determined. Then 100 colored beads corresponding to the states of the chosen task are placed into a container, and one is

---

[6] For example, in bargaining games, Cason and Mui (1997) find more altruism among groups than individuals, while Bornstein and Yaniv (1998) and Luhan, Kocher, and Sutter (2009) find the opposite. In trust games, Kugler et al. (2007) find that groups send less than individuals in the first stage, while Cox (2002) finds no significant differences. Groups are better at deducing optimal strategies in *p*-beauty contests (Kocher and Sutter 2005), centipede games (Bornstein, Kugler, and Ziegelmeyer 2004), and signaling games (Cooper and Kagel 2005) but are no better at eliminating dominated strategies (Cooper and Kagel 2009) and are more likely to overbid in common value auctions (Cox and Hayne 2006; Sutter, Kocher, and Strauss 2009). Ioannou, Qi, and Rustichini (2011) argue that the role of group identity on individuals has been exaggerated.

[7] For example, differences between group and individual decision making may conflate decision-making processes with participants' other-regarding preferences (as in bargaining experiments, e.g., Cason and Mui 1997; Luhan, Kocher, and Sutter 2009), risk tolerance (as in lottery experiments, e.g., Baker, Laury, and Williams 2008; Masclet et al. 2009; Deck et al. 2010), or other personal traits.

[8] Gillet, Schram, and Sonnemnans (2011) conduct a common pool dilemma experiment finding that groups make qualitatively better decisions in a nonstrategic setting. Groups are more competitive than individuals in a strategic setting with their efficiency relative to individuals dependent on the nature of the joint decision-making process.

BEADS	#	OPTIONS **Circle the letter option of your choice.**			
		A	B	C	D
Lime	8	✓		✓	
Pink	36	✓			
White	45		✓		✓
Green	11	✓	✓	✓	

Figure 1. Sample Choice Task

drawn. Should a pink bead be drawn and the chosen option contains pink (only option A in Figure 1), a $20 payment is earned in addition to a $5 participation payment. If a green bead is drawn when green is not included in the chosen option (only option D in Figure 1), only the $5 participation payment is earned. If a lime or a white bead is drawn, they will result in payment only if the chosen option contains the drawn color.

As subjects entered the lab, they were randomly assigned to one of three concurrently conducted treatments: (i) individual, (ii) interacting group, or (iii) noninteracting group. Both interacting and noninteracting groups consist of three subjects each. The following procedure was followed to randomly assign subjects to one of the treatments. Upon arrival, each subject in the session drew a colored ball from an urn that determined the subject's treatment assignment. The urn contained balls of three different colors representing the three treatments. Subjects did not know at the time of drawing a ball the color associated with each treatment. To create groups, the first three subjects to draw the color representing a specific group treatment were assigned to one group. Those assigned to the individual and noninteracting group treatments were directed to one large room in the lab where they were seated at separate cubicles. Subjects in the noninteracting group treatment were first introduced to their group and were seated next to their group members, but they were not allowed to speak to one another during the experiment. Each member was told to complete his or her own task booklet individually.

Each interacting group was taken to a private room. Each member read the instructions individually, allowing each to form his or her own opinion on the best procedure to solve the tasks. After all members finished reading the instructions, an experimenter gave the group one pen and one task booklet. At this point, group members were allowed to talk and interact, and they were required to complete a single task booklet as a group. On completing the task, all interacting group members were moved to adjoining private cubicles in the large room with the other subjects who had participated as individuals or members of noninteracting groups. At this point all subjects in the session completed a survey booklet.

After all survey booklets were all completed, one member of each noninteracting group was randomly chosen to have his or her decision determine the payment for the group. All

Table 1. Experimental Design

States			PDF1			PDF2			12 Options											
			States			States			8 Options											
									4 Options											
4	8	12	4	8	12	4	8	12	A	B	C	D	E	F	G	H	I	J	K	L
Lime	Lime	Lime		2	2		7	7												
	Purple	Purple	8		3	28		5												
	Purple	Orange		6	2		21	7												
		Lt Blue			1			9												
Pink	Pink	Pink		22	18		11	6												
		Yellow	36		4	24		5												
	Blue	Blue		14	14		13	13												
White	White	White		11	11		8	8												
	Brown	Brown	45	34	19	26	18	7												
		Red			15			11												
Green	Green	Green		8	8		13	13												
	Navy	Navy	11	3	3	22	9	9												

noninteracting group members earned the same amount of money based on this randomly chosen member's decision. The booklet chosen for payment was revealed to all members, so that each member in a group knew who made the decision that determined their payment. Booklets of the other two group members were kept private. For both interacting and noninteracting groups, once payoffs were determined, group members went together to receive their payment.[9]

The first task subjects observed is a small three-option three-state task designed as a familiarization tool and used as an introduction to the experiment. Subsequently, each subject is presented with 18 tasks constituting a 3×3×2 within-subject design (Table 1). The first dimension denotes the number of options, the second the number of states, and the third the probability distribution over states. Tasks have four, eight, or 12 options each described by four, eight, or 12 states (colors of beads). Two different probability distributions of colored beads are used. In PDF 1, some colored beads are more likely than others, while in PDF 2, each colored bead is roughly equally likely to be drawn. Figure 1 presents the four-option, four-state, PDF2 task. Subjects can calculate the expected payoff of an option by summing the probabilities (number of beads) of states covered by that option.

The unique optimal choice is always the option containing the largest number of beads, since that option has the highest likelihood of yielding a $20 payment. Nevertheless, past

---

[9] While making decisions in the private room, interacting groups were video recorded. As the analysis of the recordings did not yield qualitatively different results or any significant insights, it can be found in Appendix A.

experiments indicate that most subjects do not select optimally; indeed, many subjects use a heuristic that involves selecting the option that covers the most states, rather than the sum of the states' probabilities (Besedeš et al. 2012a). This task is well suited to addressing our research questions. First, as previously noted, it is an intellective task that allows for objective comparisons of individual and group performance.[10] Second, even when a group member recognizes the optimal decision rule, he or she nevertheless must win over adherents to the suboptimal but intuitively appealing rule to select the option that covers the most states. Past research has found that a simple, intuitive, though incorrect approach often triumphs in groups over truth (Tindale et al. 1996). Last, we can manipulate task complexity by changing the number of states and options in a task. This allows us to examine free riding as a function of the effort required.

Tasks are given to subjects in the form of a response booklet that lists the 19 tasks on separate pages. Subjects record their responses in the booklet with a provided pen. To control for order effects, three different versions of the response booklet are used to vary the order of the tasks. Subjects were not allowed to go backwards in their task booklets, a rule enforced by experimenters.[11] After completing the response booklet, each subject independently and privately completed a survey provided in a separate booklet.[12] The survey included questions about subject demographics and a series of simple math questions to control for the acumen required to compare options.[13]

The experiment was conducted in the Behavioral Business Research Laboratory at the University of Arkansas in the spring of 2010. Subjects were recruited from undergraduate businesses classes. A total of 150 individuals participated in sessions over the course of three days. These included 30 subjects each in the individual and noninteracting group treatments, and 90 subjects (30 groups) in the interacting groups treatment. The subject pool was 33% female, 80% white (non-Hispanic), and averaged 20.2 years of age. Subjects took on average almost 16 minutes to complete the task booklet portion of the experiment. The time spent on the booklet varied across the three treatments with noninteracting group members spending the least amount of time at just under 13 minutes. Subjects in the individual treatment spent just over 15 minutes, while subjects in the interacting group treatment spent the most time on the experiment, just under 20 minutes.[14] Similar to many of our results that follow, differences between the time taken by interacting groups and the other two treatments are statistically significant (Mann-Whitney $p < 0.006$), while that between the individual and noninteracting group treatments is not (Mann-Whitney $p = 0.269$).

---

[10] This ranking holds for models such as expected utility theory that depend on the set of final payoff amounts and the probability associated with each. It is possible to construct exotic context-specific models where this ranking might not hold. In Figure 1, one could evaluate option C under PDF1 not as giving a 19% chance of generating a payment, but as giving mutually exclusive 8% and 11% chances. Depending on how one weights these probabilities, option C may be more or less preferred to option D's 45% chance of payment.

[11] We employed a two-pronged enforcement: experimenters observed the subjects throughout the experiment, and decisions were marked with a special marker that made it impossible to secretly change a decision.

[12] Both the experimental task booklet and the survey instrument are available on request.

[13] The survey also included the three-question cognitive reflection test (CRT, Frederick 2005). However, it appears that the treatment, which was randomly assigned, impacts response to the CRT because there were significant differences across treatments in CRT scores collected posttask even though there were no differences in other demographic characteristics between treatment groups. This suggests that whether one participated in group or individual decision making may affect cognitive reflection on subsequent tasks.

[14] Interacting groups have the lowest standard deviation (6 minutes), followed by noninteracting groups (over 7 minutes) and individuals (over 15 minutes).
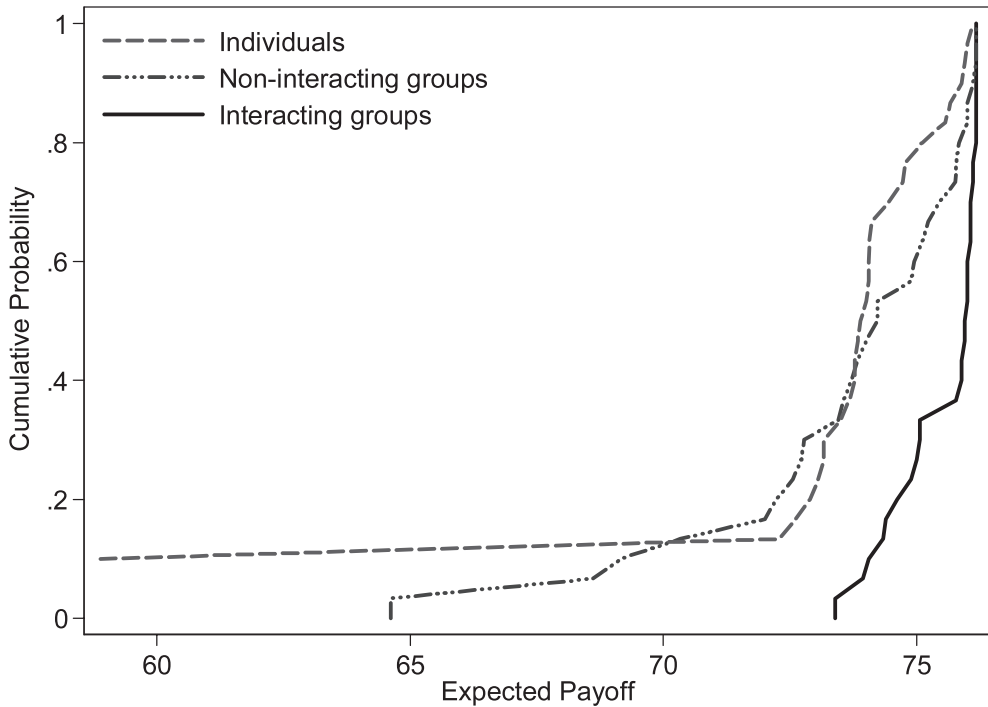
Figure 2. Cumulative Distribution of Expected Payoffs by Treatment

3. Results

Overall Performance

We begin with a comparison of overall performance across treatments. Each task has a unique optimal option associated with the highest probability of payment. Groups with interaction make the optimal decision in 87%, followed by individuals in 72% and noninteracting groups in 67% of all tasks. The difference between interacting groups and the other two treatments is highly significant (Mann-Whitney $p < 0.004$).[15] In fact, interacting groups select the optimal option more frequently than subjects in the two other treatments on *each* of the 18 experimental tasks. Differences between the noninteracting groups and individuals are not significant ($p > 0.100$).[16]

Figure 2 presents the distribution of expected payoffs by treatment. An optimal choice on each task would result in an expected payoff of 76.2% across all tasks. Nearly one-quarter of all interacting groups achieve this outcome, selecting the optimal option in each task. Again, we find that interacting groups significantly outperform subjects in both the individual and noninteracting group treatments, with the interacting group distribution of payoffs stochastically dominating the other two treatments (Kolmogorov-Smirnov $p < 0.004$). For

---

[15] Three subjects in the individual treatment and one subject in the noninteracting treatment failed to provide a choice for one of their 19 tasks. Our statistical results are not sensitive to dropping these four tasks or to coding them as the minimum, average, or maximum obtainable payoffs on that task.

[16] The comparisons are qualitatively unchanged if we use the average expected payoff as the measure to compare performance, rather than the frequency of selecting the optimal option.

example, while 80% of all interacting groups achieve an expected probability of payment above 75%, less than half of subjects in the individual treatment and less than one-third of subjects in the noninteracting group treatment do so.

Individual Effort and Free Riding

Next, we compare performance in the individual treatment to that in the noninteracting group treatment. In both treatments, subjects complete the tasks independently and without any assistance from others. However, the noninteracting group introduces two countervailing incentives. First, group membership and payoff dependence may encourage higher effort through, for example, a sense of responsibility for the welfare of others in the event one's decisions are binding upon all group members. Second, effort is subject to a free-riding incentive because a member can benefit from other group members' efforts, and one's own actions have a two-thirds chance of being inconsequential. Which of these incentives dominates determines whether noninteracting group members perform better or worse than individuals.

To examine if free riding is exhibited, we take advantage of our experimental design by comparing performance on tasks of varying difficulty. A task with 12 options and 12 states, for example, requires more cognitive effort to identify the unique optimal option than a task with only four options and four states. As task difficulty increases, the demands on effort increase. The benefit from exerting that effort remains constant as each task is equally likely to be selected for payment. We would expect performance to decline with task complexity across all treatments. However, if members of noninteracting groups are free riding on the effort of others, we would expect a greater discrepancy between noninteracting groups and individuals on hard tasks than on easy ones.

Although we cannot directly observe an individual's cognitive effort, we make the assumption that the effort subjects invest in solving a problem is reflected in their performance. If subjects are not putting forth effort, then we will observe performance that is no better than random choice. In other words, in order to solve the task, subjects must be putting forth some effort to come to a decision. Noninteracting groups are correct on nearly one out of every two hard tasks, whereas the rate would be one in 12 if subjects guessed randomly. Another way we can link performance to effort is by considering participants' survey responses to the question, "Would you please explain how you made your decisions in the experiment, noting any factors that contributed to your behavior (please feel free to use the back of this sheet if you need more room)." There are obvious patterns when segmenting responses by the number of optimal choices made during the experiment. Those participants, in both the individual and group treatments, with all optimal choices used words and phrases such as "each person added up columns" and "we would check each other's arithmetic." These comments anecdotally illustrate both an understanding of how to solve the problem as well as time and consideration (effort) put into solving. Similarly, individuals and group members who selected few if any optimal options provided more vague responses like "chose the most obvious," "went kind of fast through the end because I was distracted," and "I guessed the choice in which I felt was the most correct."

We examine the frequency of selecting the optimal option in relatively harder and easier tasks in Table 2. We define harder tasks as those with 12 options and 12 states while all other tasks are defined as "easier."[17] Overall, subjects are much more likely to select the optimal

---

[17] Similar results follow from a less-restrictive definition of a harder task as one with at least eight options and at least eight states.

Table 2. Frequency of Optimal Choice in Tasks of Varying Difficulty

	Noninteracting Group (%)	Individual (%)
Easier tasks	70	73
Harder tasks	47	65

option on easier tasks than harder ones (Wilcoxon $p < 0.001$). Individuals select the optimal option in 73% of easier tasks, while noninteracting group members do so in 70% (Mann-Whitney $p = 0.593$). However, individuals select the optimal option in 65% of harder tasks, while noninteracting group members do so in only 47% (Mann-Whitney $p = 0.032$). Expected payoffs follow the same pattern, with individuals earning the same payoffs as noninteracting subjects on easier tasks (Mann-Whitney $p = 0.525$), but significantly higher payoffs on harder tasks (Mann-Whitney $p = 0.019$).

We conclude that subjects free ride in noninteracting groups. Of course, free riding can occur to varying degrees, from slightly lowering the effort to providing no effort and choosing randomly. Evidence suggests that the extent of free riding is limited. For example, members of noninteracting groups select the optimal option in nearly half of harder tasks, which is below the rate in the individual treatment but also well above the one in 12 chance implied by random choice.

To determine if subject-specific differences across treatments can account for this result, we estimate the determinants of optimal choice in a panel probit model (Table 3). With 30 subjects in the individual and noninteracting group treatments, and with each subject making 18 decisions of interest, there are a total of 1076 observed decisions, accounting for four nonresponses. In the first column, we include controls from our postexperiment survey for a subject's sex, race, and the number of correctly answered basic math questions (Math Score), with each variable reflecting a subject's individual characteristics and not that of any other group members for noninteracting groups. Additionally, to identify whether there are significant differences between treatments in easier and harder tasks, we incorporate treatment-specific dummies for task types. The reference category is the individual treatment in easy tasks.

Confirming our aggregate results, there is no significant difference between individuals and noninteracting groups on easier tasks, while noninteracting groups do significantly worse on harder tasks. Thus, we find evidence of free riding with noninteracting group members being less likely to invest effort in harder problems than individuals.

In the second column we add a set of dummy variables indicating the position of a specific task in the sequence of all 18 tasks seen by subjects. These round-order dummies control for any possible order effects. The inclusion of round-order dummies does not affect our results qualitatively, slightly increasing the magnitude of the two statistically significant coefficients. In the final column we include a measure of the amount of time each subject spent on the entire experiment, measured in minutes.[18] The inclusion of time does not affect the estimates qualitatively. The time variable itself is estimated with a positive and statistically significant coefficient indicating that subjects who spend more time on the entire experiment tend to do better. However, we caution that time is likely highly endogenous, preventing a causal

---

[18] Due to the difficulty of observing time spent on a single task in in-person experiments, we do not have task-specific time measurements.

Table 3. Optimal Choice in Individual and Noninteracting Treatments

	(1)	(2)	(3)
Harder task	−0.191	0.271	0.316
	(0.154)	(0.214)	(0.226)
Noninteracting group	−0.587***	−0.620***	−0.689***
× Harder task	(0.222)	(0.262)	(0.261)
Noninteracting group	−0.160	−0.202	−0.235
× Easier task	(0.143)	(0.160)	(0.133)
Male	0.075	0.078	0.084
	(0.123)	(0.138)	(0.119)
White	0.518***	0.587***	0.355**
	(0.155)	(0.174)	(0.139)
Math score	0.021	0.023	−0.009
	(0.061)	(0.069)	(0.061)
Total time			0.057***
			(0.015)
Constant	0.168	−0.333	−0.885**
	(0.260)	(0.287)	(0.386)
Observations	1076	1076	1076
Log pseudo-likelihood	−634	−563	−524
Round order dummies	No	Yes	Yes

Estimated coefficients with robust standard errors clustered by subject in parentheses.
  * $p < 0.1$.
 ** $p < 0.05$.
*** $p < 0.01$.

interpretation. It may be that more time leads to better decision making or that better decision makers tend to spend more time deciding.

   We next examine whether the tendency to free ride is higher for some subjects than others. Figure 3 illustrates performance in individual and noninteracting group treatments across three subject characteristics: males versus females, low versus high math scores, and nonwhite versus white subjects. A subject is classified as a high math score if he or she answered at least four of the five math questions in the postexperiment survey correctly.[19] Each panel indicates how a subset of subjects performed on tasks of varying complexity in the two treatments.

   We note that each subset exhibits some decline in performance on harder tasks when a member of a noninteracting group and that, at first glance, the slope of this decline seems similar for almost all subgroups. However, a significant difference is observed between women and men. Women select the optimal option slightly more in the noninteracting group treatment than in the individual treatment on easier tasks and slightly less on harder tasks. However, these differences are not significant (Mann-Whitney $p > 0.429$). Men, who also exhibit no significant difference in performance across treatments on easier tasks ($p = 0.216$), do exhibit significant differences on harder tasks ($p = 0.039$). Thus, men appear to free ride by decreasing effort on harder tasks upon joining a group. Women, conversely, do not.

   We examine these potential determinants of free riding more formally in Table 4. We append to our previous probit model separate dummy variables for noninteracting group performance on easier and harder tasks for male, white, and high math subjects, in addition to

---

[19] The average number of correct math answers is very similar across the three treatments, with 3.33 correct answers for both individual and noninteracting group treatments and 3.41 for the interacting group treatment.
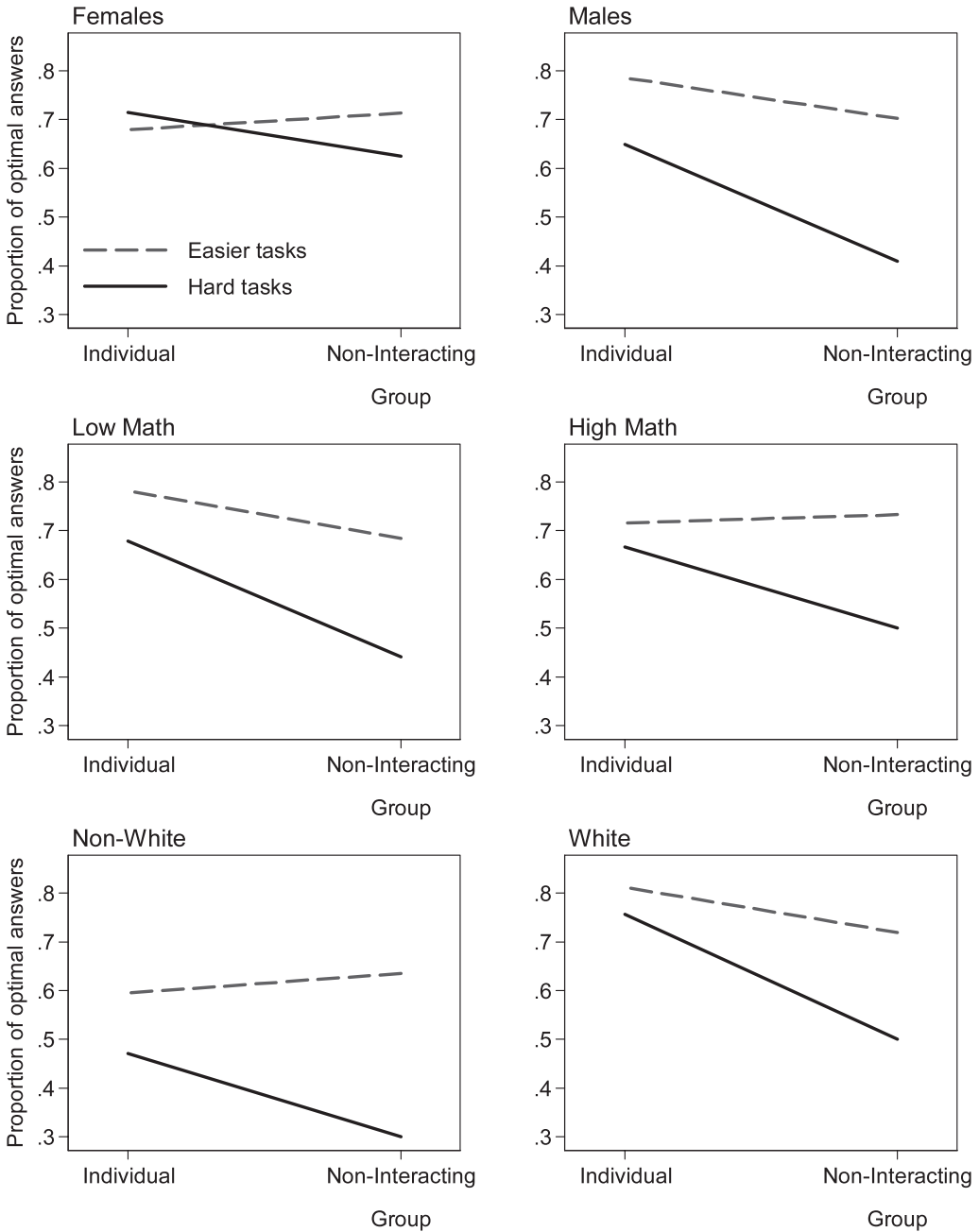
Figure 3. Performance on Hard and Easier Tasks in Noninteracting Group versus Individual Treatments

a general dummy variable for harder tasks.[20] The dummy for harder tasks is not statistically significant, while males do worse on harder tasks in noninteracting groups. This indicates that free riding, and the resulting reduced performance of noninteracting groups in harder tasks, is primarily the result of effort reduction on the part of males ($p = 0.007$). In columns 2 and 3 we

---

[20] We examined the effect of other demographic variables (age, own and parents' education) and measures of risk attitudes. These variables do not contribute significantly either individually or collectively and do not change the sign or significance of the variables of interest.

Table 4. Free Riding in Noninteracting Groups by Subject Characteristics

	(1)	(2)	(3)
Harder task	−0.230	0.221	0.285
	(0.153)	(0.215)	(0.223)
Noninteracting group	−0.847***	−0.980***	−1.090***
× Harder task × Male	(0.313)	(0.361)	(0.331)
Noninteracting group	−0.161	−0.128	0.147
× Harder task × White	(0.298)	(0.360)	(0.354)
Noninteracting group	0.348	0.399	0.152
× Harder task × High math	(0.270)	(0.324)	(0.319)
Noninteracting group	−0.236	−0.277	−0.340
× Easier task × Male	(0.276)	(0.311)	(0.215)
Noninteracting group	−0.255	−0.304	−0.003
× Easier task × White	(0.223)	(0.251)	(0.198)
Noninteracting group	0.316	0.373*	0.182
× Easier task × High math	(0.199)	(0.226)	(0.178)
Male	0.191	0.211	0.286**
	(0.189)	(0.211)	(0.146)
White	0.626***	0.717***	0.355**
	(0.193)	(0.215)	(0.168)
Math score	−0.050	−0.061	−0.048
	(0.072)	(0.082)	(0.071)
Total time			0.056***
			(0.015)
Constant	0.272	−0.209	−0.846**
	(0.306)	(0.343)	(0.398)
Observations	1076	1076	1076
Log pseudo-likelihood	−628	−555	−519
Round order dummies	No	Yes	Yes

Estimated coefficients with robust standard errors clustered by subject in parentheses.
  * $p < 0.1$.
 ** $p < 0.05$.
*** $p < 0.01$.

add the round-order dummies and the total time spent in the experiment. With round-order dummies included, high math subjects seem to do marginally better on hard tasks in noninteracting groups, but that result disappears with the inclusion of time and is not different from high math subjects' improved performance on easier tasks. The caveat about interpreting the time variable still remains.

In the absence of interaction, joining a group reduces overall performance. In terms of the dichotomy between free riding (which is expected to reduce effort in groups) and social responsibility (which is expected to increase effort), the free-riding effect is a stronger force for males, while women do neither better nor worse in noninteracting groups than as individuals. Additionally, we find no evidence that the tendency to free ride is predicted by race or mathematical ability.

We end this section by examining the dispersion of choices and its implications for payoffs. We do so by comparing the difference between the highest and lowest probability of payment based on actual choices of members of noninteracting groups in each task. The average spread for noninteracting groups is 8.06 percentage points in easy tasks, meaning that on average the best choice selected by a group member has 8 percentage point higher

probability of payment than the worst choice. For hard tasks the average spread is 10.90 percentage points. With a payment of \$20 at stake, this implies an expected loss of between \$1.61 on easy tasks and \$2.20 on hard tasks if the worst member in the noninteracting group determines the payoff instead of the best member.

We can compare these amounts to what the average spread would be if noninteracting group members behaved as if they were making individual decisions without the incentive to free ride and without any social responsibility. We can obtain such groups by creating all 4060 possible hypothetical three-member groups using subjects in the individual treatment. In these hypothetical groups, the average spreads between the worst and best choice are only 5.03 percentage points in easy tasks and 5.22 percentage points in hard tasks, suggesting approximate potential losses of only \$1 given the \$20 stakes. Thus, free riding in noninteracting groups in hard tasks may result in more than twice the loss relative to individuals working alone. We next turn to examining the effect of groups that are free to interact and communicate.
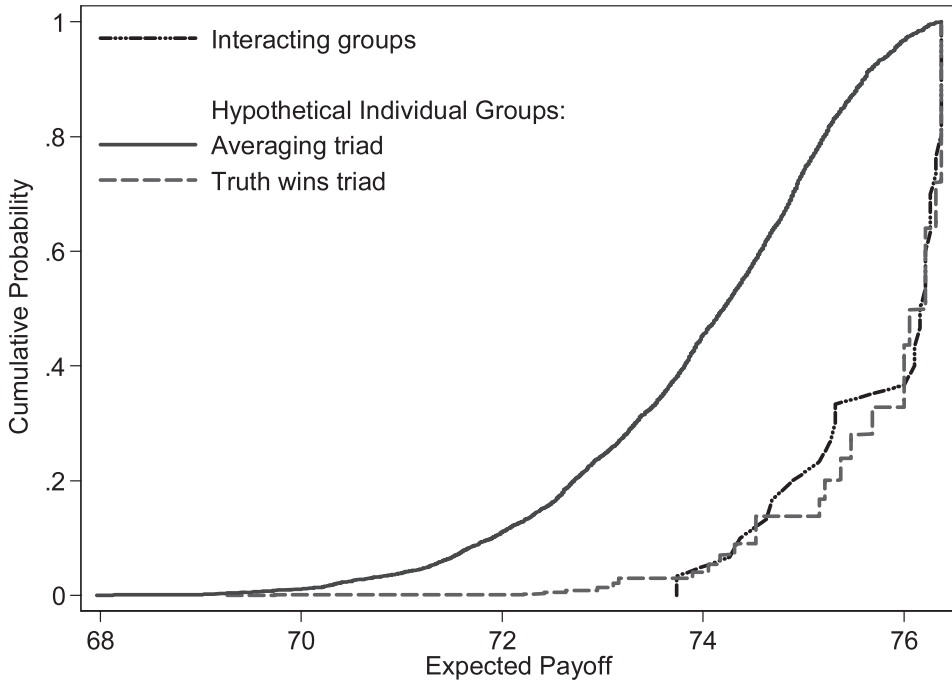
Interaction

Before we can examine how group interaction affects the balance between free riding and social responsibility, we need to examine how groups use each member's knowledge. When individuals interact and collaborate on a common decision, the degree of success depends on both the group's aggregation of its members' knowledge and on the group's ability to create knowledge beyond what any one member possesses. Aggregation can take several forms. If a group member is chosen to solve the problem for reasons uncorrelated with ability (e.g., charisma), then groups would do as well as individuals, on average. A proportionality or majority procedure can be expected to reinforce predominant attitudes of its members. In the best case, the approach of the most capable member is adopted, a so-called "truth wins" standard (Steiner 1972; Davis 1973; Cooper and Kagel 2005). If groups create knowledge and do not merely aggregate it, then groups exhibit "assembly bonus effects" by which performance exceeds even what the most capable member could have achieved on his or her own (Laughlin, Bonner, and Miner 2002). However, some results suggest that "assembly bonus effects" and even "truth wins" are rare, because groups rarely perform as well as their best member (Tindale and Larson 1992; MacCoun 1998; Kerr and Tindale 2004; Forysth 2009).

To examine which of the aggregation benchmarks best describes our data, we compare the outcomes of interacting groups with the aggregated judgments of the same number of subjects in the individual treatment. We formulate all 4060 possible combinations of three subjects from the individual treatment. We call these three-member hypothetical groups "triads." For each triad, we calculate both the highest payoff of the three individuals (a "truth wins triad") and the average payoff of the three individuals (an "averaging triad"). These hypothetical payoffs of triads are compared to the actual payoffs of interacting groups.

In Figure 4 we present the cumulative distribution of payoffs for interacting groups and for both averaging triads and truth wins triads. Interacting groups do not appear to select one member randomly to make the decision for the group as the performance of an interacting group is far better than averaging triads in expectation (Mann-Whitney $p < 0.001$) and

Figure 4. Performance of Interacting Groups and Hypothetical Groups of Individuals

stochastically dominates averaging triads (Kolmogorov-Smirnov $p < 0.001$).[21] However, the performance of interacting groups is indistinguishable from that of the truth wins triads. Both in expectation and in distribution, we cannot reject that interacting groups do as well ($p > 0.594$ for both Mann-Whitney and Kolmogorov-Smirnov tests).

Simply put, the performance of interacting groups is statistically similar to the performance of the best-performing group member. However, this does not appear to be a literal description of group dynamics. If groups simply adopted the optimal decision rule whenever one of their group members understood it, we would see variance in performance *across* groups (based on whether or not such a member exists), but not across decisions *within* groups. Instead, while seven interacting groups never select a suboptimal option, seven other groups select a suboptimal option at least four times. In addition, suboptimal choices should cease once group members understand the optimal decision rule, which we can reasonably expect to occur relatively early in the experiment. Yet we find that of the 23 groups that make at least one suboptimal decision, only two groups make their last such choice in the first two rounds. The remaining 21 groups make the last suboptimal decision in the latter half of the experiment (in rounds 9 through 18) with 10 groups making their last suboptimal choice in the last two rounds.[22] Thus, many groups fail to adopt the optimal decision-making rule

---

[21] Both the parametric and nonparametric test results also hold with $p < 0.001$ if variance is adjusted using Abrevaya's (2008) recombinant estimator.

[22] Because some groups see a hard task in the seventeenth round, such a large concentration of the last mistake at the end of the experiment may be due to fatigue. However, only one group made its last mistake on a hard task in round 17. We present a more extensive argument against a large role for fatigue at the end of this section.

consistently.[23] Nevertheless, we conclude that while groups, statistically, are great aggregators of existing knowledge, they do not outperform the truth wins triad and thus do not (on average) create knowledge.

In the previous subsection, we identified free riding among male subjects in noninteracting groups. These subjects performed significantly worse on harder tasks as members of noninteracting groups than as individuals. In Table 5 we compare performance in individual and interacting treatments using similar variables as in Table 4, again estimating a panel probit model.[24] The individual and interacting group treatments are not directly comparable (we do not observe individual choices in the interacting groups treatment), and so we adopt two empirical strategies. First, we analyze the likelihood of selecting the optimal strategy at the subject level with each interacting group member inheriting the outcome of the group. With 30 subjects in the individual treatment and 90 in interacting group treatment, we have a total of 2157 observations (each subject makes 18 decisions, with three in the individual treatment not making a decision on a task). We conduct a weighted probit, with each interacting group member receiving one-third weight. Second, we perform the same analysis at the group level. Here we compare the actual decisions of interacting groups with those of our 4060 "truth wins" triads to reflect hypothetical group performance from the individual treatment. Demographic variables (male, white, math score) reflect the average of each of these variables for the three members of an interacting group or truth wins triad. For example, a group with two male members and one female member is assigned a value of two-thirds for the variable "male." We apply a weight of 30/4060 to each triad to equalize the relative importance of each treatment.[25] This approach leaves us with 213,710 "observations" on decisions made by individuals in a total of 4090 groups (with each group composed of three individuals making 18 decisions per group).

The subject-level analysis assumes that each interacting group subject did as well as the entire group, while the group level analysis assumes that each subject in the individual treatment did as well as the best of each three-member triads in which he or she is a member. Both approaches yield the same overall result: While performance depends on subject and task characteristics, there is no evidence of free riding. Across the three specifications, while race and sex affect overall performance at the group level, neither leads to any significant differences between easier and harder tasks. Men—the group that exhibited free riding in noninteracting groups—do not appear to affect performance in interacting groups.

The only factor consistently contributing to differences between easier and harder tasks is math acumen. Our results indicate that interacting groups with higher math individuals do significantly better on easy tasks, but also somewhat worse on hard tasks.[26] Charbonnier et al.

---

[23] Anecdotally, some groups had discussions and arguments between using the optimal decision rule (selecting the option with the largest frequency of states) and one that seemed more intuitive to group members (selecting the option with the largest quantity of states). At least two groups settled on compromises, limiting consideration to the two or three options with the highest number of states before taking account of probabilities, or using probabilities to handle "ties" among options with the same number of states, This illustrates that truth need not always win but also offers an explanation for better performance in groups even when optimality is not obtained.

[24] On one hand, some individuals may have even greater incentive to free ride in interacting groups than noninteracting ones, especially if they perceive their effort as dispensable (Jones 1984; Karau and Williams 1993). On the other hand, the greater saliency of group membership brought about by joint decision making and interaction may reduce psychological incentives to reduce effort (Wagner 1995; Charness, Rigotti, and Rustichini 2007; Sutter 2009).

[25] There are no qualitative differences in our results if we do not use any weighting to equalize the contributions of each unit of observation.

[26] The impact on hard tasks is significant when considering the analysis at the subject level, but not the group level.

Table 5. Optimal Choice in Individual and Interacting Group Treatments

	Subject Level			Group Level		
	(1)	(2)	(3)	(4)	(5)	(6)
Harder task	−0.244*	0.143	0.160	−0.183***	0.202***	0.250***
	(0.140)	(0.190)	(0.196)	(0.014)	(0.037)	(0.042)
Interacting group	0.101	0.066	−0.034	0.132	0.019	0.057
× Harder task	(0.270)	(0.296)	(0.262)	(0.540)	(0.607)	(0.627)
× Male						
Interacting group	0.168	0.238	0.421	0.239	0.416	0.997*
× Harder task	(0.279)	(0.316)	(0.313)	(0.437)	(0.490)	(0.565)
× White						
Interacting group	−0.287	−0.385*	−0.446**	−0.476	−0.620*	−0.582
× Harder task	(0.187)	(0.207)	(0.214)	(0.324)	(0.339)	(0.359)
× High math score						
Interacting group	0.314	0.366	0.260	0.298	0.373	0.445
× Easier task	(0.217)	(0.238)	(0.205)	(0.328)	(0.367)	(0.408)
× Male						
Interacting group	−0.047	−0.084	0.103	0.014	−0.043	0.527
× Easier task	(0.214)	(0.235)	(0.243)	(0.298)	(0.334)	(0.415)
× White						
Interacting group	0.505***	0.563***	0.531***	0.536***	0.593***	0.680***
× Easier task	(0.131)	(0.146)	(0.145)	(0.177)	(0.195)	(0.252)
× High math score						
Male	0.060	0.061	0.182	0.159***	0.181***	0.300***
	(0.193)	(0.212)	(0.150)	(0.014)	(0.016)	(0.018)
White	0.479**	0.553***	0.284**	0.653***	0.753***	0.373***
	(0.197)	(0.215)	(0.192)	(0.017)	(0.019)	(0.014)
Math score	−0.051	−0.058	−0.011	0.005	0.005	0.052***
	(0.077)	(0.085)	(0.058)	(0.009)	(0.010)	(0.010)
Total time			0.050***			0.062***
			(0.012)			(0.002)
Interacting group			−0.015			−0.044***
× Total time			(0.013)			(0.017)
Constant	0.508	0.026	−0.763***	0.105**	−0.394***	−1.253***
	(0.310)	(0.359)	(0.292)	(0.044)	(0.079)	(0.085)
Observations	3231	3231	3231	858,750	858,750	858,750
Log pseudo-						
likelihood	−1508	−1357	−1297	−431,207	−381,806	−354,904
Round order						
dummies	No	Yes	Yes	No	Yes	Yes

Estimated coefficients with robust standard errors clustered by subject in parentheses.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

(1998) conjecture that those who perceive themselves as more capable than other group members may exert less effort as they perceive less individual glory from their effort in a group setting. To the extent that higher math individuals decrease effort in groups on harder tasks, they also appear to compensate with higher effort on easier tasks. Overall, subject characteristics cannot explain why interacting groups perform better in the aggregate. Our interpretation of this result is that it is the communication and interaction within the group that allows interacting groups to perform better, rather than a particular characteristic of its

members. On easier tasks, our results suggest that interacting groups may do better because of high math members, while their apparent worse performance in harder tasks does not hurt the entire group. Interacting groups who use more time to complete the task booklet seem to do worse, perhaps an indication of extended discussion or debate for those groups, which negatively impacts their performance.

Without round-order dummies (columns 1 and 4) there seems to be an overall worse performance on harder tasks, while the inclusion of round-order dummies and time indicate a slightly better performance on harder tasks at the group level. Of more importance is what their inclusion states about an alternative explanation for the lack of free riding. As our experiment is composed of 18 rounds (taking an average of 17 and one-half minutes for subjects to complete), it is possible that interacting groups perform better because they are better suited to dealing with fatigue. Subjects in the individual and noninteracting group treatments make all decisions on their own and are dealing with the effects of fatigue on their own. Subjects in the interacting groups may be less fatigued toward the end of the experiment because of their shared effort and interaction.[27] If fatigue plays a differential role, one would expect that the inclusion of round-order dummies and total time would result in some variables with the interacting group interaction becoming positive and statistically significant. That is not the case. In addition, an inspection of the coefficients on round-order dummies themselves[28] indicates that there is no concentration of large negative coefficients toward the end of the experiment. Relative to the first task (easy task in all three booklets), performance in the last six tasks is never worse at the subject level and is worse at the group level in tasks 13 and 17. Only the seventeenth task is a hard task in some booklets. To properly evaluate performance on the hard task in round 17, the coefficients on the hard task dummy and round 17 must be added together. Since they are of opposite signs they offset each other. In fact, the sum of the two coefficients is never statistically different from zero under any specification. Thus, we can conclude that fatigue is not a likely explanation for the superior performance of interacting groups.

4. Conclusion

The effect of group decision making depends on both the effect of group membership and the effect of interaction within groups. Group membership, in itself, introduces an additional sense of responsibility, especially if others are sharing in the fruits of one's labor. However, groups can also diffuse responsibility, providing incentives to reduce one's effort. Our design allows us to differentiate between the responsibility or accountability incentives and the incentive to free ride because they are reflected by the net outcome and different behaviors on harder tasks relative to easier ones.

Charness, Rigotti, and Rustichini (2007) and Sutter (2009) find that payoff commonality, in itself, leads to better decision making. In their designs, each group member is solely

---

[27] We are now clearly assuming that subjects in interacting groups are not fatigued by arguing and deliberating on which option to choose in each task!

[28] We chose not to present these to save on space. They are available on request.

responsible for a fraction of the group's decisions and payoffs. They demonstrate that group salience in addition to individual accountability and responsibility encourages better decision making. Our results indicate that the diffusion of responsibility that often accompanies groups is a negative, offsetting, and stronger force. By requiring every member to make a decision on every task while not communicating our design makes it possible for diffusion of responsibility to occur and reflect negatively on decision making. Charness, Rigotti, and Rustichini (2007) and Sutter (2009) find little evidence of the negative effects of the diffusion of responsibility. They do so with a design where rather than each member making a decision while not interacting, each member is solely responsible for making one-third of the group's decisions. This design maximizes the pressure the member potentially feels by being the group's representative, while eliminating the diffusion of responsibility. We have each subject responsible for all of the group's decisions with one-third probability of it determining the group's payoff. Our design allows members to still feel the responsibility pressure, while allowing for the diffusion of responsibility as well. While these are identical in expectation, they produce entirely different results.

Our evidence suggests that when each group member has to make a decision that may determine the group's outcome without communicating with other members, there is a significant negative effect on decision making. We interpret this negative net effect as the result of free-riding motives outweighing the social responsibility motives. Allowing for diffusion of responsibility while preventing communication increases the free-riding tendencies that manifest themselves negatively on decision making. The identified free-riding tendencies could be a consequence of payoff commonality alone not being sufficient to induce group saliency in the absence of group communication. In other words, the form of payoff commonality is important. Payoff commonality for a number of individuals may induce a group-like outcome provided there is no possibility for the diffusion of responsibility, as is the case in Charness, Rigotti, and Rustichini (2007) and Sutter (2009), but not in our experiment.

In our experiment, we observe that men engage in free riding while women do not, though this is fully tempered by group interaction. When groups make a joint decision in a setting where they can interact, we no longer find the free-riding effect. Thus, we agree with Charness, Rigotti, and Rustichini (2007) and Sutter (2009) that sufficient group salience improves performance within groups. We disagree that payoff commonality, alone, is sufficient to achieve such salience. However, allowing for communication offsets the free-riding incentives that exist in the absence of communication. We find that groups that are allowed to interact freely outperform both individuals and noninteracting groups by a wide margin, selecting the optimal option with a much higher frequency. In particular, interacting groups do as well as the best individual member would have done on his or her own. Thus, interacting groups appear simultaneously to minimize free riding and to be very good aggregators of existing knowledge. We conclude that better performance does not necessarily follow from group saliency, in itself, but from the interaction among group members.

Appendix A: Video Analysis

The video recordings of interacting groups give us a source of additional data to examine which aspects of group formation and interaction are most important in making optimal decisions. We use the recordings to determine the

Table A1. Video Analysis of Interacting Groups

| | Leader Definition | | |
	Talked First	Talked Most	Wrote Responses
Harder task	−0.665***	−0.665***	−0.665***
	(0.090)	(0.090)	(0.090)
Male	0.061	0.073	0.077
	(0.130)	(0.132)	(0.134)
White	0.172	0.177	0.160
	(0.139)	(0.143)	(0.143)
Math count	0.073	0.076	0.042
	(0.089)	(0.100)	(0.094)
Confused	−0.275	−0.263	−0.272
	(0.196)	(0.207)	(0.200)
Slacker	0.118	0.045	0.148
	(0.184)	(0.207)	(0.204)
Total time	−0.002	−0.004	−0.003
	(0.011)	(0.011)	(0.011)
Leader	0.120	−0.021	0.118
	(0.187)	(0.171)	(0.188)
Leader × High math	−0.235	−0.155	−0.061
	(0.221)	(0.219)	(0.207)
Constant	0.877**	0.938**	0.963**
	(0.391)	(0.398)	(0.414)
Observations	1620	1620	1620
Log pseudo-likelihood	−595	−595	−596

Robust standard errors in parentheses.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

leader of each group, whether there are slackers or confused members, and the amount of time it took groups to complete the experiment. We explore three definitions of a leader: the first to speak, the person who speaks the most, and the individual who wrote responses. A member is a slacker if he or she never participated in the decision-making process. Members are identified as being confused if they stated so at any point during the deliberations. In Table A1 we add this information to demographic characteristics and examine the performance of interacting groups at the individual level. The addition of these variables sheds no additional light on group deliberations. The group is significantly less likely to make an optimal decision on hard tasks, but this is the only characteristic that has a statistical impact on performance. The existence of a slacker or a confused member does not harm group performance. The total amount of time it took groups to complete the experiment plays no role in how well they do. There is no leader effect, because the leader dummy is never significant, regardless of how the leader is defined. Finally, we interacted the leader dummy with high math count (four or more correct). It is not significant regardless of how the leader is defined.

Appendix B: Experiment Instructions

Instructions for Individual Treatment

You will receive $5 for participating in this experiment and completing a brief survey. You can also earn an additional sum of money based on performance in the experiment. The experiment consists of 18 tasks. You will be given a booklet containing the 18 tasks, and each task is on a separate page in the booklet. It is important that you make the choices in the order in which they are presented in the task booklet. That is, you must complete the tasks in order, and once you complete a task you cannot go back to it. Please do not go back to any previous pages.

Each task requires the completion of a response form on which you will make a choice from a set of options appearing in a table such as the one below. In each task, you will select one of the options.

Table B1. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

There will be a container of colored beads and one bead will be randomly drawn from the container at the end of the experiment. A volunteer will conduct the drawing in front of you. The column "BEADS" will list the colors of beads in the container and the column "#" will list the number of beads of each color in the container. There will always be a total of 100 beads. Thus according to the above table the container will have 10 Red beads, 30 Orange beads, and 60 Yellow beads adding up to 100 beads. The chance that a particular color will be drawn is the number of beads of that color/100. In this example, there is a 30/100 = 30% chance that an orange bead will be drawn.

Under the "Options" heading will be a set of letters. The letters correspond to the different options that you may choose. In the example above, you could choose Option A, B, C, D, E, or F. Each option contains a series of marks corresponding to the colored beads. For example, Option C has a mark for the color red only while Option D has marks for both red and yellow. Alternatively note that Yellow beads are present in Options A, D, and E.

For each task you must choose only one option by circling the letter of your choice with the provided pen. Do not add any other marks on the page; just indicate your selected option by circling it. If you make a mistake or wish to change your response, please raise your hand and inform an experimenter. Circling multiple options or making additional marks without informing an experimenter may result in a loss of compensation.

After you have selected an option for each task, please close your booklet. You may then complete the brief survey.

Once everyone has finished, a volunteer will pick a number at random to determine which of the 18 tasks will be used to determine your payment. Note that even though you are making 18 decisions, only one randomly chosen task will affect your payment.

First the container will be filled with 100 colored beads according to the "#" column of the selected task. Then one bead will be randomly drawn from the container. If the option you chose for the selected task does not have a ✓ mark for the color of the bead drawn, you will leave with your $5 participation payment. However, if the option you chose does have a ✓ mark for the color of the bead drawn, you will receive $20. This will be in addition to the $5 participation payment, making your total earnings $25.

On the next page is an example. Suppose the following task was randomly selected and the person had chosen Option F by marking it as shown.

Table B2. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

If an orange bead is drawn from the container, then this person as well as anybody else who chose Option F would be paid the $5 participation payment plus $20 (for a total of $25). Also persons who chose Options B and E would receive the $20 (for a total of $25) since they contain a mark for orange. Anyone selecting Options A, C, or D would only receive the $5 participation payment.

After the drawing, a researcher will come to you to verify what you have earned. The researcher will give you a claim slip that you can use to collect your payment as you leave. When called, you will hand the claim slip to a researcher who will ask you to sign a receipt in exchange for your money. You will then drop your response booklet, survey, and pen in a large box. This process is designed to ensure that no one, including the researchers, can ever know the responses of any individual.

If you have any questions about the experiment, please ask now.

Otherwise, please wait quietly until you are taken to a room to complete the response booklet. Once there, you may open your response booklet and begin with Task 1. Keep in mind that you cannot go backwards through the booklet and should not skip around. Once you complete the booklet, close it and begin the survey. Please do not go back to the booklet once it has been closed.

Instructions for the Interacting Groups Treatment

You will receive $5 for participating in this experiment and completing a brief survey. You can also earn an additional sum of money based on performance in the experiment. The experiment consists of 18 tasks. You will be given a booklet containing the 18 tasks, and each task is on a separate page in the booklet.

You will be put into a group of three to complete the task booklet. Group members will be randomly chosen. You will all work together to make choices for the 18 tasks. It is important that you make the choices in the order in which they are presented in the experiment booklet. That is, you must complete the tasks in order, and once you complete a task you cannot go back to it. Please do not go back to any previous pages.

Each task requires the completion of a response form on which you will make a choice from a set of options appearing in a table such as the one below. In each task, you will select one of the options.

Table B3. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

There will be a container of colored beads and one bead will be randomly drawn from the container at the end of the experiment. A volunteer will conduct the drawing in front of you. The column "BEADS" will list the colors of beads in the container and the column "#" will list the number of beads of each color in the container. There will always be a total of 100 beads. Thus according to the above table the container will have 10 Red beads, 30 Orange beads, and 60 Yellow beads adding up to 100 beads. The chance that a particular color will be drawn is the number of beads of that color/100. In this example, there is a 30/100 = 30% chance that an orange bead will be drawn.

Under the "Options" heading will be a set of letters. The letters correspond to the different options that your group may choose. In the example above, you could choose Option A, B, C, D, E, or F. Each option contains a series of marks corresponding to the colored beads. For example, Option C has a mark for the color red only while Option D has marks for both red and yellow. Alternatively note that Yellow beads are present in Options A, D, and E.

For each task your group must choose only one option by circling the letter of your choice with the provided pen. Do not add any other marks on the page; just indicate your selected option by circling it. If you make a mistake or wish to change your response, please raise your hand and inform an experimenter. Circling multiple options or making additional marks without informing an experimenter may result in a loss of compensation.

After your group has selected an option for each task, please close your booklet. You may then complete the brief survey individually.

Once everyone has finished, a volunteer will pick a number at random to determine which of the 18 tasks will be used to determine your payment. Note that even though you are making 18 decisions, only one randomly chosen task will affect your payment.

First the container will be filled with 100 colored beads according to the "#" column of the selected task. Then one bead will be randomly drawn from the container. If the option your group chose for the selected task does not have a ✓ mark for the color of the bead drawn, every member of your group will leave with their $5 participation payment. However, if your group's chosen option does have a ✓mark for the color of the bead drawn, you will each receive an additional $20, making your total earnings $25.

On the next page is an example. Suppose the following task was randomly selected and the group had chosen Option F by marking it as shown.

Table B4. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

If an orange bead is drawn from the container, then everyone in this group as well as everyone in any other group who chose Option F would be paid the $5 participation payment plus $20 (for a total of $25 each). Also every member of groups that chose Options B and E would receive the $20 (for a total of $25) since they contain a mark for orange. Members of groups that selected Options A, C, or D would only receive the $5 participation payment.

After the drawing, a researcher will come to you to verify what you have earned. The researcher will give you a claim slip that you can use to collect your payment as you leave. When called, you will hand the claim slip to a researcher who will ask you to sign a receipt in exchange for your money. You will then drop your response booklet, survey, and blue ink pen in a large box. This process is designed to ensure that no one, including the researchers, can ever know the responses of any individual.

If you have any questions about the experiment, please ask now.

Otherwise, please wait quietly until you are taken to a room to complete the response booklet. Once there, you may open your group's response booklet and begin with Task 1. Keep in mind that you cannot go backwards through the booklet and should not skip around. Once you complete the booklet, close it and begin the survey. Please do not go back to the booklet once it has been closed.

Instructions for the Noninteracting Groups Treatment

You will receive $5 for participating in this experiment and completing a brief survey. You can also earn an additional sum of money based on performance in the experiment. The experiment consists of 18 tasks. You will be given a booklet containing the 18 tasks, and each task is on a separate page in the booklet.

You will be put into a group of three, but each of you will complete the task booklet individually. Group members will be randomly chosen. At the end of the experiment, the task booklet of one member of your group will be randomly selected and their decision will be used to determine the payoff for everyone in your group. It is important that you make the choices in the order in which they are presented in the experiment booklet. That is, you must complete the tasks in order, and once you complete a task you cannot go back to it. Please do not go back to any previous pages.

Each task requires the completion of a response form on which you will make a choice from a set of options appearing in a table such as the one below. In each task, you will select one of the options.

Table B5. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

There will be a container of colored beads and one bead will be randomly drawn from the container at the end of the experiment. A volunteer will conduct the drawing in front of you. The column ''BEADS'' will list the colors of beads in the container and the column ''#'' will list the number of beads of each color in the container. There will always be a total of 100 beads. Thus according to the above table the container will have 10 Red beads, 30 Orange beads, and 60 Yellow beads adding up to 100 beads. The chance that a particular color will be drawn is the number of beads of that color/100. In this example, there is a 30/100 = 30% chance that an orange bead will be drawn.

Under the ''Options'' heading will be a set of letters. The letters correspond to the different options that you may choose. In the example above, you could choose Option A, B, C, D, E, or F. Each option contains a series of marks corresponding to the colored beads. For example, Option C has a mark for the color red only while Option D has marks for both red and yellow. Alternatively note that Yellow beads are present in Options A, D, and E.

For each task you must choose only one option by circling the letter of your choice with the provided pen. Do not add any other marks on the page; just indicate your selected option by circling it. If you make a mistake or wish to change your response, please raise your hand and inform an experimenter. Circling multiple options or making additional marks without informing an experimenter may result in a loss of compensation.

After you have selected an option for each task, please close your booklet. You may then complete the brief survey individually.

Once everyone has finished, a volunteer will randomly determine which one group member's booklet will be used. In other words decisions made by one person in the group will determine the payoffs of all the others in the group. The volunteer will also pick a number at random to determine which of the 18 tasks will be used to determine your payment. Note that even though you are making 18 decisions, only one randomly chosen task will affect your payment.

First the container will be filled with 100 colored beads according to the ''#'' column of the selected task. Then one bead will be randomly drawn from the container. If the option your group chose for the selected task does not have a ✓ mark for the color of the bead drawn, every member of your group will leave with their $5 participation payment. However, if your group's chosen option does have a ✓ mark for the color of the bead drawn, you will all receive an additional $20, making your total earnings $25.

On the next page is an example. Suppose the following task was randomly selected and the group had chosen Option F by marking it as shown.

Table B6. Task

BEADS	#	OPTIONS Circle the letter option of your choice.					
		A	B	C	D	E	F
Red	10			✓	✓		✓
Orange	30		✓			✓	✓
Yellow	60	✓			✓	✓	

If an orange bead is drawn from the container, then everyone in this group as well as everyone in any other group who chose Option F would be paid the $5 participation payment plus $20 (for a total of $25 each). Also every member of groups that chose Options B and E would receive the $20 (for a total of $25) since they contain a mark for orange. Members of groups that selected Options A, C, or D would only receive the $5 participation payment.

After the drawing, a researcher will come to you to verify what you have earned. The researcher will give you a claim slip that you can use to collect your payment as you leave. When called, you will hand the claim slip to a researcher who will ask you to sign a receipt in exchange for your money. You will then drop your response booklet, survey, and blue ink pen in a large box. This process is designed to ensure that no one, including the researchers, can ever know the responses of any individual.

If you have any questions about the experiment, please ask now.

Otherwise, please wait quietly until you are taken to a room to complete the response booklet. Once there, you may open your group's response booklet and begin with Task 1. Keep in mind that you cannot go backwards through the booklet and should not skip around. Once you complete the booklet, close it and begin the survey. Please do not go back to the booklet once it has been closed.

References

Abrevaya, Jason. 2008. On recombinant estimation for experimental data. *Experimental Economics* 11(1):25–52.

Baker, Ronald J., II, Susan K. Laury, and Arlington W. Williams. 2008. Comparing small-group and individual behavior in lottery-choice experiments. *Southern Economic Journal* 75(2):367–82.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor. 2012a. Age effects and heuristics in decision making. *Review of Economics and Statistics* 94(2):58–95.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor. 2012b. Decision-making strategies and performance among seniors. *Journal of Economic Behavior and Organization* 81(2):524–33.

Blinder, Alan S., and John Morgan. 2005. Are two heads better than one? Monetary policy by committee. *Journal of Money, Credit, and Banking* 37(5):789–811.

Blinder, Alan S., and John Morgan. 2008. Leadership in groups: A monetary policy experiment. *International Journal of Central Banking* 4(4):117–50.

Bornstein, Gary, Tamar Kugler, and Anthony Ziegelmeyer. 2004. Individual and group decisions in the centipede game: Are groups more 'rational' players? *Journal of Experimental Social Psychology* 40(5):599–605.

Bornstein, Gary, and Ilan Yaniv. 1998. Individual and group behaviour in the ultimatum game: Are groups more 'rational' players? *Experimental Economics* 1(1):101–8.

Brown-Kruse, Jamie, and David Hummels. 1993. Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is? *Journal of Economic Behavior and Organization* 22(3):255–67.

Cadsby, C. Bram, and Elizabeth Maynes. 1998. Gender and free riding in a threshold public goods game: Experimental evidence. *Journal of Economic Behavior and Organization* 34(4):603–20.

Cason, Timothy N., and Vai-Lam Mui. 1997. A laboratory study of group polarization in the team dictator game. *Economic Journal* 107(444):1465–83.

Charbonnier, Emmanuelle, Pascal Huguet, Markus Brauer, and Jean-Marc Monteil. 1998. Social loafing and self-beliefs: People's collective effort depends on the extent to which they distinguish themselves as better than others. *Social Behavior and Personality* 26(4):329–40.

Charness, Gary, Edi Karni, and Dan Levin. 2007. Individual and group decision making under risk: An experimental study of Bayesian updating and violations of first-order stochastic dominance. *Journal of Risk and Uncertainty* 35(2):129–48.

Charness, Gary, Edi Karni, and Dan Levin. 2010. On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda. *Games and Economic Behavior* 68(2):551–56.

Charness, Gary, Luca Rigotti, and Aldo Rustichini. 2007. Individual behavior and group membership. *American Economic Review* 97(4):1340–52.

Charness, Gary, and Martin G. Sutter. 2012. Groups make better self-interested decisions. *Journal of Economic Perspectives* 26(3):157–76.

Chen, Joseph S., Harrison G. Hong, Ming Huang, and Jeffrey D. Kubik. 2004. Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review* 94(5):1276–1302.

Chen, Yan, and Sherry Xin Li. 2009. Group identity and social preferences. *American Economic Review* 99(1):431–57.

Cooper, David J., and John H. Kagel. 2005. Are two heads better than one? Team versus individual play in signaling games. *American Economic Review* 95(3):477–509.

Cooper, David J., and John H. Kagel. 2009. Equilibrium selection in signaling games with teams: Forward induction or faster adaptive learning? *Research in Economics* 63(4):216–24.

Cox, James C. 2002. Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In *Advances in Experimental Business Research*, edited by Rami Zwick and Amnon Rapoport. Dordrecht: Kluwer Academic Publishers, pp. 331–50.

Cox, James C., and Stephen C. Hayne. 2006. Barking up the right tree: Are small groups rational agents? *Experimental Economics* 9(3):209–22.

Davis, James H. 1973. Group decision and social interaction: A theory of social decision schemes. *Psychological Review* 80(2):97–125.

Deck, Cary, Jungmin Lee, Javier Reyes, and Chris Rosen. 2012. Risk taking behavior: An experimental analysis of individuals and pairs. *Southern Economic Journal* 79(2):277–99.

Forsyth, Donelson R. 2009. *Group Dynamics*. Belmont, CA: Cengage Learning.

Frederick, Shane. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* 19(4):25–42.

Gillet, Joris, Arthur Schram, and Joep Sonnemans. 2009. The tragedy of commons revisited: The importance of group decision-making. *Journal of Public Economics* 93(5–6):785–97.

Hargreaves-Heap, Shaun P., and Daniel John Zizzo. 2009. The value of groups. *American Economic Review* 99(1):295–323.

Henry, Rebecca A. 1993. Group judgment accuracy: Reliability and validity of post discussion confidence judgments. *Organizational Behavior and Human Decision Processes* 56(1):11–27.

Ioannou, Christos A., Shi Qi, and Aldo Rustichini. 2011. Group outcomes and reciprocity. Accessed May 10, 2012. Available http://core.kmi.open.ac.uk/download/pdf/23247.pdf.

Isenberg, Daniel J. 1986. Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology* 50(6):1141–51.

Jones, Gareth R. 1984. Task visibility, free riding, and shirking: Explaining the effect of structure and technology on employee behavior. *Academy of Management Review* 9(4):684–95.

Karau, Steven J., and Kipling D. Williams. 1993. Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology* 65(4):681–706.

Kerr, Norbert L., Robert J. MacCoun, and Geoffrey P. Kramer. 1996. Bias in judgment: Comparing individuals and groups. *Psychological Review* 103(4):687–719.

Kerr, Norbert L., and R. Scott Tindale. 2004. Group performance and decision making. *Annual Review of Psychology* 55(1):623–55.

Kocher, Martin G., and Matthias Sutter. 2005. The decision maker matters: Individual versus group behaviour in experimental beauty contest games. *Economic Journal* 115(500):200–23.

Kocher, Martin G., and Matthias Sutter. 2007. Individual versus group behavior and the role of the decision making procedure in gift-exchange experiments. *Empirica* 34(1):63–88.

Kugler, Tamar, Gary Bornstein, Martin G. Kocher, and Matthias Sutter. 2007. Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic Psychology* 28(6):646–57.

Kugler, Tamar, Edgar Kausel, and Martin G. Kocher. 2012. Are groups more rational than individuals? A review of interactive decision making in groups. *WIREs Cognitive Science* 3(4):471–82.

Latané, Bibb, Kipling Williams, and Stephen Harkins. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology* 37(6):822–32.

Laughlin, Patrick R., Bryan L. Bonner, and Andrew G. Miner. 2002. Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes* 88(2):605–20.

Luhan, Wolfgang J., Martin G. Kocher, and Matthias Sutter. 2009. Group polarization in the team dictator game reconsidered. *Experimental Economics* 12(1):26–41.

MacCoun, Robert J. 1998. Biases in the interpretation and use of research results. *Annual Review of Psychology* 49(1):259–87.

Masclet, David, Nathalie Colombier, Laurent Denant-Boemont, and Youenn Loheac. 2009. Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior and Organization* 70(3):470–84.

Nowell, Clifford, and Sarah Tinkler. 1994. The influence of gender on the provision of a public good. *Journal of Economic Behavior and Organization* 25(1):25–36.

Prather, Larry J., and Karen L. Middleton. 2002. Are N+1 heads better than one? The case of mutual fund managers. *Journal of Economic Behavior and Organization* 47(1):103–20.

Steiner, Ivan Dale. 1972. *Group processes and productivity*. New York: Academic Press.

Sutter, Matthias. 2009. Individual behavior and group membership: Comment. *American Economic Review* 99(5):2247–57.

Sutter, Matthias, Martin G. Kocher, and Sabine Strauss. 2009. Individuals and teams in auctions. *Oxford Economic Papers* 61(2):380–94.

Tindale, R. Scott, and James R. Larson, Jr. 1992. Assembly bonus effect or typical group performance: A comment on Michaelsen, Watson, and Black (1989). *Journal of Applied Psychology* 77(1):102–5.

Tindale, R. Scott, Christine M. Smith, Linda S. Thomas, Joseph Filkins, and Susan Sheffey. 1996. Shared representations and asymmetric social influence processes in small groups. In *Understanding Group Behavior: Consensual Action by Small Groups*, edited by Eric E. White and James H. Davis. Mahwah, NJ: Erlbaum, pp. 81–103.

Wagner, John A., III. 1995. Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management Journal* 38(1):152–72.