Effects of a Management – Labor Context and Team Play on Ultimatum Game Outcomes*

Hal Arkes
Ohio State University

John H. Kagel
Ohio State University

Dimitry Mezhvinsky
Ohio State University

8/6/2015

Individual players an ultimatum game labeled as “managers” made significantly lower offers than in a generic-label control treatment. Acceptance rates for “labor” were the same or higher as well. Two-person teams showed more mixed results, with outcomes varying with the amount of money at stake. The use of business context and the presence of a teammate each result in smaller offers and higher acceptance rates. However, these effects did not compound. Instead, the team effect largely replaced the context effect. But with meaningful context there were no consistent differences between teams and individuals.

Key words: ultimatum game, context versus generic, teams versus individuals
JEL classification: C78, C91, C92, D03

* This research has been partially supported by National Science Foundation grant SES-1226460. Opinions, findings, conclusions or recommendations offered here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Economists and psychologists have both noted that the labeling of an economic game significantly influences the participants' outcomes. When Rege and Telle (2004) told participants that contributions in a public good game would go into a "community box," the magnitude of the contributions was significantly higher than when the contributions would be placed in a "box." Analogous findings were obtained in many other public goods games (e.g., Dufwenberg, Gachter, & Hennig-Schmidt, 2011; Cookson, 2000; Eriksson & Strimberg, 2014; Pillutla & Chen, 1999). Similar results have also been demonstrated in prisoners' dilemma games (e.g., Batson & Moran, 1999; Eiser & Bhavnani, 1974; Ellingsen, Johannesson, Mollerstrom & Munkhammar, 2012; Kay & Ross, 2003; Liberman, Samuels, & Ross, 2004; McDaniel & Sistrunk, 1991; Zhong, Loewenstein, & Murnighan, 2007). For example, Ellingsen et al. (2012) found that, when a prisoners' dilemma game was described as the "Community Game," cooperation was significantly greater than when the same dilemma was described as the "Stock Market Game."

In this paper we ascertain if mere labeling can influence behavior in an economic game as simple and direct as the ultimatum game (Guth, Schmittberger, & Schwarze, 1982). In the ultimatum game a Proposer is endowed with an amount of money, often $10. The Proposer then determines a split of the money between herself and a Responder. If the Responder accepts the proposed split, the money is divided according to the proposal. If the Responder rejects the proposed split, neither party receives any money.

To the best of our knowledge, there has been only one study in which the description of an ultimatum game has been manipulated. Hoffman, McCabe, Schachat, and Smith (1994) found that, when the game was described as a seller setting a price for a good which a potential buyer can accept or reject, Proposers' offers were approximately 10 percent lower than when the usual Proposer/Responder terminology was used. Responders' rejection rates did not differ between the two versions of the labeling. Note that Hoffman et al. (1994) varied the entire milieu: one was a purchase situation, and the other was not.

The experiment varied the labels assigned to participants: Either the standard (generic) "Proposer" and "Responder" labels or the labels of "Management" and "Labor," within a meaningful context for the relationship between the two players. We conducted the generic and context treatments with individual decision makers and repeated both treatments using two-

person teams as the decision-making unit. Each session featured only one labeling system and one type of decision maker.

Prior Research: Priming through Labels

How might the label of an economic game influence the behavior of the participants? A vast number of "priming" studies have demonstrated that making a concept salient in even a surreptitious manner can influence subsequent behavior. For example, Bargh, Chen, and Burrows (1996) presented participants with three types of five-word sequences, each of which had to be unscrambled in order to form a coherence sentence. One type contained word sequences whose constituent words formed a sentence related to rudeness. A second type contained word sequences whose constituent words formed a sentence related to politeness, and a third type contained neutral sentences. Following the completion of this task, the participant was supposed to go to the experimenter outside the room in order to receive a second task. However the experimenter was talking to another person and ignored the participant who was seeking this second task. The dependent variable was the amount of time the participant waited before interrupting the experimenter's conversation. Those who had been primed with rude words interrupted the experimenter much sooner than persons who were in either of the other two groups.

More closely related to the research in the current paper, Shih, Pittinsky, and Ambady (1999) showed that Asian-American women primed with their Asian identity exhibited superior performance on a mathematics test, and those primed with their female identity exhibited inferior performance on the math test, both performance levels being compared to women in a control group. The authors suggested that the stereotypical roles of Asians and women in the domain of mathematics were responsible for the differential performance levels.

Eriksson and Strimling (2014, p. 360) succinctly summarize the influence of labels in this manner: "Various authors have used different words to describe what is essentially the same explanation of the effectiveness of label framing. In sum the common view is that various mental models of situations are developed in the course of life experiences. These mental models may include norms and other constraints on one's own behavior as well as expectations on others' behavior. An individual's behavior in a given situation will depend on which mental model is selected to serve as the mental frame of the situation." In their research on the influence of labels on economic games other than the ultimatum game, Dufwenberg, et al. (2011), Herr (1986), Kay

and Ross, (2003), and Pillutla and Chen (1999) all suggest that frames influence what behavior norms are operative at a given time and therefore what behavior is appropriate. The questions posed here are (i) will the labeling of the participants in a situation as straightforward as the ultimatum game have a significant effect for individuals, and (ii) if so, will it carry over to teams, which the literature shows are typically more thoughtful and "rational" in the sense of behaving in closer conformity to standard selfish economic man?[1]

Treatment 1: The Effect of Labels on Individual Decision Makers

Experimental Design and Procedures: The core ultimatum game was the same across all treatments. At the beginning of an experimental session, participants were assigned the role of either a proposer or a responder, with equal numbers of each. The assigned role remained the same for each participant for the duration of the experiment. For each round of a session, proposers and responders were randomly and anonymously paired together. Subjects were encouraged to treat each round as a one-shot game by the payment rule, which compensated subjects on the basis of the outcome in a single, randomly determined, round.

Each round, the proposer was presented with a sum of money to split between herself and the responder, in the amount of either $10, $14, or $18, with each sum appearing exactly three times over the course of nine rounds. We varied the size of sum to be split to determine whether the amount at stake affected the impact of the terminology used. The sum to be divided was randomly assigned across the nine rounds in each session to minimize order effects.

Once a proposer selected a proposed split of the available sum, this split was presented to the corresponding responder. The responder observed the dollar values proposed and chose either to "Accept" or "Reject" the split. If the responder chose to "Accept," the split the proposer selected was implemented for that round. If the responder chose to "Reject," the proposer and responder both received nothing for that round.

Subjects were assigned to either one of two treatments. The first treatment employed generic instructions with the ultimatum game characterized using generic labels ("Proposer" and "Responder") along with the typical set of instructions employed in an ultimatum game. The second treatment employed contextualized instructions with proposers labeled as "Management"

[1] See Charness and Sutter (2012) and Kugler et al. (2012) for surveys of the literature.

and responders referred to as "Labor." Only one treatment, either generic or context, was employed in each session.

In contrast to the generic treatment, the context treatment instructed players to think of themselves as either the management of a company or as representing the workers of the company. Additionally, players were asked to think of the sum to be split as part of a "profit sharing" plan. Specifically, subjects in the context treatment were primed as follows: "In this company there is profit sharing, so that the amount of money to be split up depends on how much profit the company made that year. The amount of money available in each round will vary, sometimes being $10, sometimes $14, and sometimes $18. You can think of $10 as the amount the company has available in a bad year, $14 is what they have available in an average year, and $18 is what they have available in a good year. On each round you will make an offer to Labor that specifies how much of the money you will receive and how much of the money they will receive." For both treatments, it was made clear to all participants that in each round they were paired against another randomly chosen participant in the room.

With the exception of the different terminology and the hypothetical context presented in the context treatment, treatments were identical. The terminology and context employed were intended to prime participants to think of a familiar setting featuring an unequal power dynamic between managers and workers in a company. We wanted to ascertain if this priming would affect participants' attitudes about what constitutes an acceptable split.

All sessions were conducted at The Ohio State University Economics Experimental Laboratory using ORSEE software (Greiner, 2004). Participants were seated at individual workstations, each with a computer that participants used to enter their choices. The experiment was run using z-Tree software (Fischbacher, 2007). Instructions were presented to participants in written form and read aloud prior to the start of each session.[2] There were 3 generic-label (proposer/responder) individual subject sessions (GI) with between 16 and 20 subjects in each session, and 3 context-label (management/labor) individual subject sessions (CI) with between 16 and 18 subjects in each session. The participants were primarily undergraduate students attending the University. Average payoffs including a $5 show-up payment were $12.42 in the GI treatment and $11.35 in the CI treatment. The duration of each session was approximately 20 minutes, including the time necessary to read the instructions.

[2] See http://econ.ohio-state.edu/kagel/Ultimatum_Game_Instructions.pdf for the instructions used.

Experimental Results: Figure 1 reports shares offered for both the generic and context treatments, distinguishing between accepted and rejected offers for the different pie sizes, pooling together the three rounds in which each pie size was offered. The first thing to notice is that for all three pie sizes the frequency of equal splits is lower in the context treatment, with these offers displaced in favor of lower offers. The average share proposed is significantly lower with context for all three pie sizes, with proposers' share increasing modestly with increasing amounts of money at stake for both generic and context treatments (see Table 1).

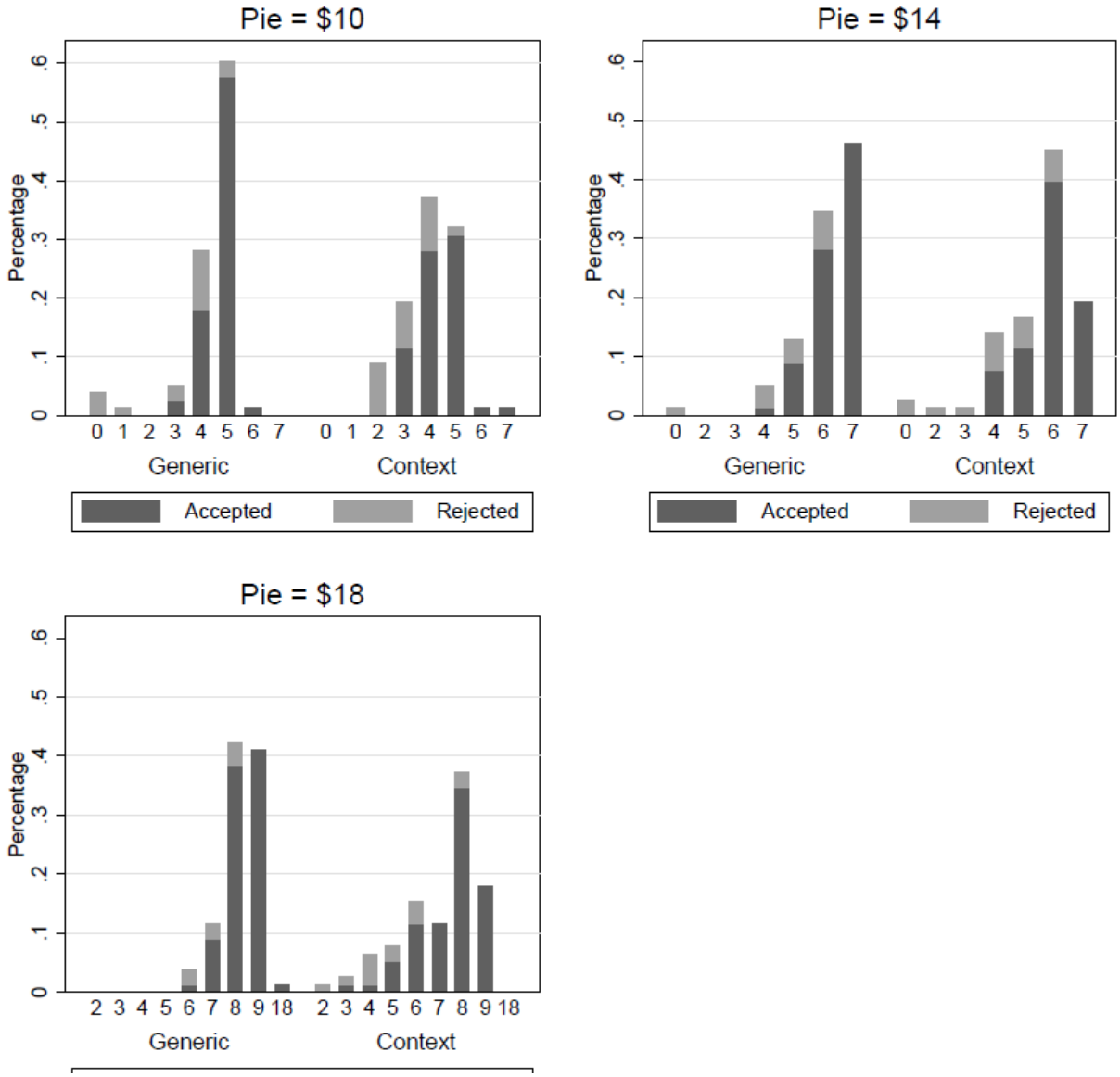

Figure 1: Individuals' acceptance and rejections at each pie size.

Table 1
Average Dollar Amount Offered: Individuals
[shares in brackets]

Pie Size	Mean (Standard deviation of the mean)		
	Generic	Context	Generic - Context
$10	$4.38 ($0.91) [0.44]	$4.01 ($0.85) [0.40]	$0.37*
$14	$6.15 ($0.78) [0.44]	$5.50 ($1.07) [0.39]	$0.65**
$18	$8.35 ($0.98) [0.46]	$7.06 ($1.62) [0.39]	$1.29***

*p <0.10, **p < 0.05, ***p < (0.01) Differences based on Wilcoxon rank-sum test using mean proposer averages as the unit of observation.

Figure 2 shows the frequency with which each offer was accepted, distinguishing between the generic and context treatments. The figure omits offer sizes that are observed fewer than nine times for each pie size in the data.[3] In all cases, shares are accepted at the same rate, or higher, with context. Table 2 reports expected payoff maximizing offers along with modal offers. At pie size $10, the expected payoff maximizing offer is $5 for both treatments.[4] At larger pie sizes, expected payoff maximizing offers are lower by $1 in the context treatment, with corresponding higher payoffs for proposers. Modal offers are commonly equal to the payoff maximizing offer. When they are not, they are within $1 and typically quite close in frequency.[5]

[3] Following Roth et al. (1991) we planned to use frequencies of 10 or more which hold in all cases except for pie size $18, where context had frequencies of 14, 29, 9, 12 for offers of $9, $8, $7 and $6, respectively. Frequency of 9 was included for the sake of consecutive dollar offers.

[4] It's mystifying why anyone would turn down an equal split, but it happened 4% of the time here. We view this as inherent noise, or trembles, in the experiment.

[5] For the $10 Context case, there are 29 proposals at $4, with 25 at $5. For the $18 Generic case there are 32 offers at $9 and 33 at $8. The one outlier here is the $18 Context, with 29 offers of $8 and 9 offers at $7.

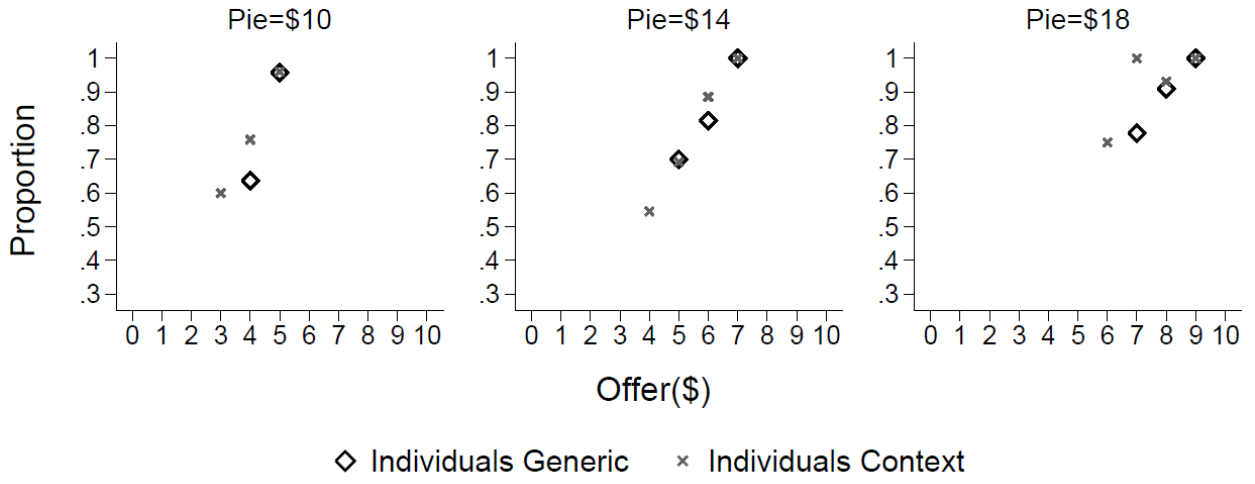

Figure 2: Acceptance rates for individuals: Generic versus Context

Table 2
Expected Payoff Maximizing Offers: Individuals
(Proposers Payoff in parentheses)

	Expected Value Maximizing Offer		Modal Offer	
Pie Size	Generic	Context	Generic	Context
$10	$5 ($4.80)	$5 ($4.80)	$5	$4
$14	$7 ($7.00)	$6 ($7.12)	$7	$6
$18	$8 ($9.10)	$7 ($11.00)	$8	$8

Conclusion 1: For individuals, the priming effect of the management-labor context resulted in lower average offers at all pie sizes, along with consistently fewer, or the same, rejection rates at each pie size. The net effect is that the payoff maximizing offer was $1 lower at pie sizes $14 and $18, and the same at $10. That is, the priming effect of meaningful context triggered the expected response of favoring proposers over responders in this simplest of games. A secondary result is that the expected value maximizing offer, as a percentage of the amount of money to be split, decreased at larger dollar values with context, and at the $18 dollar value for the generic

treatment. The fact that modal offers are at, or within $1, of payoff maximizing proposals, shows that proposers had a pretty good idea of what constitutes an acceptable offer.

Discussion: The results for individuals indicate that the priming worked in the way one might expect.[6] These findings are in keeping with results reported in Ball, Eckel, Grossman, and Zame (2001) who showed that assigning "stars" (little star stickers along with a public announcement to that effect) to buyers and sellers in a double oral auction influenced equilibrium outcomes: When the potential sellers were the star awardees, the price at which the items were sold was significantly higher than when the buyers were the star awardees. This held regardless of whether the stars were awarded randomly or to those who scored well on a quiz. The relative status of the two groups impacted transaction prices, with the "stars" inducing differential status, along with corresponding norms for who should achieve better outcomes in the double auction. In the ultimatum game there is an inherent asymmetry in the power relationship between proposers and responders. The management/labor frame accentuated this asymmetric relationship, in line with the perceived asymmetric power relationship between management and labor outside the lab.

Treatment 2: The Effect of Labels on Decision Makers in Teams

Treatment 2 was a replication of Treatment 1 with one important exception: the proposers and responders were each two-person teams acting as proposers and responders. The motivation for this treatment was three fold: (i) to determine if labels would have the same effect on teams as individuals, (ii) to compare teams with individuals, and (iii) to use the transcripts from the team chats to obtain better insight into the motivational forces and concerns of players in this game. In conducting the experiment, we were cognizant of the fact that past research in both economics and psychology made no clear prediction one way or the other on whether the labels would have the same effect for teams. But with respect to teams versus individuals, the typical result is that teams are "more rational" than individuals in the sense of coming closer to the

[6] In this respect it would be interesting to investigate this same priming effect in a country with much stronger labor unions.

subgame perfect equilibrium prediction of a minimum ($1) offer which is accepted regardless of the sum in question (see Bornstein & Yaniv, 1998).[7]

In a comprehensive review of group performance and decision making in psychology experiments, Kerr and Tindale (2004, p. 625) concluded that "The ubiquitous finding across many decades of research . . . is that groups usually fall short of reasonable potential productivity baselines . . . they exhibit process losses." However it is not clear what a process loss comprises in the ultimatum game. If it means being more influenced by irrelevant labels, then the difference between the management/labor and the neutral contexts would be greater for teams versus individuals. If it means being oblivious to the norms primed by the labels, then one might expect a smaller, or non-existent, difference between the two contexts with teams.

Reviews of the team literature in economic and game theory experiments (Charness & Sutter, 2012; Kugler et al. 2012) also yield ambiguous predictions with respect to priming effects. Although teams are generally considered to be more rational, as defined above, the implications of this for priming through labels is not entirely clear. On the one hand, one could argue that this should effectively eliminate the impact of priming as teams "see through" the labels, ignoring the meaning responses they were intended to induce. On the other hand, one could argue that the team discussions would amplify any labeling effects as "management" and "labor" discuss their options. Further, we are unaware of any game theory experiments investigating this question.

At the same time, based on previous research, we would expect clear differences between team and individual behavior, particularly for the generic case. Bornstein and Yaniv (1998) found that, in an ultimatum game with three-person teams, the teams offered less than the individuals. That is, teams were more rational in the sense that they behaved closer to the game-theoretic solution (although still quite far away from offering the smallest possible positive share). Bornstein and Yaniv's experiment was conducted using the standard generic format with agents labeled as proposers and responders.

Again, with meaningful context there are no prior studies to reference. But consider the following: Assume that really low shares will be more or less universally rejected (which they are). Because meaningful context with the individuals resulted in significantly lower shares

[7] Zero shares also correspond to the subgame perfect equilibrium but leave responders indifferent between accepting and rejecting.

offered, this would limit the room for teams to push offers even lower and hope for them to be accepted. That is, practically speaking there is a floor on how low offers can be while having a respectable chance of being accepted. This, in turn, leaves less room for a team effect along the lines reported in Bornstein and Yaniv (1998) for our team context treatment.

Experimental Design and Procedures: The generic and context treatments used exactly the same terminology, and priming procedures as for individuals. Two person teams were randomly formed, for both proposers and responders, with teammates and roles staying the same for the duration of the session. A player's identity was never revealed to her teammate, and players were instructed not to identify themselves.

In each bargaining round, teams were randomly and anonymously paired with each other. Subjects were seated at separate computer terminals, communicating with each other through a free form digital chat box. Proposers were given three minutes to determine how to split the money with the responders. Once a proposer team agreed on a split, the offer was sent to the responder team who had three minutes to agree whether to accept or reject the offer.[8] Further, responders were able to chat with their teammate while proposers were making their decisions, and proposers were able to chat with their teammates while responders were making their decisions. If a team could not agree on a decision once the three minutes had expired they received an additional 30 seconds to reach agreement without the chat box. The instructions laid out rules for how choices would be determined if they could not reach agreement after that. Ultimately, we did not see any cases of disagreement among teammates, though there were several cases in which one teammate failed to enter a choice relying on the fact that if only one member of the team made a choice, this would be recorded as the team's choice.

The team sessions proceeded the same way as the individual treatments, with each teammate receiving the full amount her team received in the round selected for compensation. That is, teammates did not split their team's earnings.

Four sessions each of each team treatment were conducted, for a total of 48 and 56 teams for the context and generic session, respectively, so as to have essentially the same power as the individual sessions. The average payoff for each subject, including a $5 show-up payment, was $12.00 in the generic treatment and $9.17 with context.

---

[8] All teams received an additional minute to chat during the first round in order to allow time for participants to get acquainted with the chat box.

Experimental Results: Results are first reported along the same lines as the individual subject sessions comparing the generic and context treatments directly. After that comparisons are reported for teams versus individuals. We conclude with a brief section analyzing the team chats to see what can be learned about players' thought processes and motivations.

Generic versus context conditions within teams: Figure 3 reports shares offered for both the team sessions, distinguishing between accepted and rejected offers for the different pie sizes, pooling over the three rounds each pie size was offered. Again, there are large numbers of equal splits at all pie sizes, with average shares offered not keeping up with the increase in the amount of money to be split (see Table 3). Comparing between treatments, average shares offered are greater for all pie sizes with context, the opposite of what was reported for individuals. But the differences are not as large and fail to achieve statistical significance even at the 10% level for the $14 pie.

Figure 3: Teams' acceptance and rejection at each pie size.

Table 3
Average Dollar Amount Offered: Teams
[shares in brackets]

Pie Size	Mean (Standard deviation of the mean)		
	Generic	Context	Generic - Context
$10	$4.01 ($0.68) [0.40]	$4.35 ($0.54) [0.43]	-$0.34**
$14	$5.73 ($0.78) [0.41]	$5.92 ($0.93) [0.42]	-$0.19
$18	$7.10 ($1.11) [0.39]	$7.63 ($1.20) [0.42]	-$0.53*

*p <0.10, **p < 0.05, ***p < (0.01) Differences based on Wilcoxon rank-sum test using mean proposer averages as the unit of observation.

Figure 4 shows the frequency with which teams accepted offers, distinguishing between the generic and context cases, again restricted to offer sizes that are observed nine or more times in the data. The pattern here is different from that of individuals. With pie sizes $10 and $14, the differences between generic and context acceptance rates are negligible. However, for the $18 pie, acceptance rates for generic teams are substantially higher. This is the opposite of the pattern for individuals. Table 4 reports expected payoff maximizing offers along with modal offers. Modal and payoff maximizing offers for the $10 pie are both $4. At $14, although mean proposals are a bit lower for the generic treatment (see Table 3), the payoff maximizing offer's expected value is $0.52 higher with context. However, at $18, this is reversed, with the payoff maximizing offer at $6 for generic versus $8 for context, largely as a result of the higher acceptance rates at $6 for generic teams. As was the case for individuals, modal offers are commonly equal to payoff maximizing offers, the one exception being context with pie size $14.

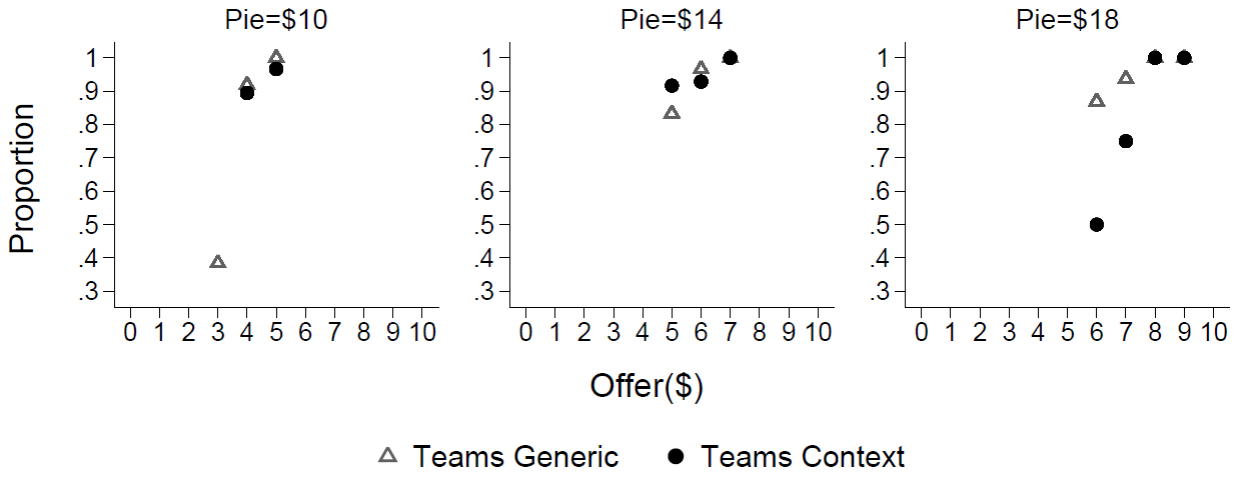

Figure 4: Acceptance rates teams: generic versus context.

Table 4
Expected Payoff Maximizing Offers: Teams
(Proposers Payoff in parentheses)

	Expected Value Maximizing Offer		Modal Offer	
Pie Size	Generic	Context	Generic	Context
$10	$4 ($5.52)	$4 (5.34)	$4	$4
$14	$6 ($7.76)	$5 (8.28)	$6	$6
$18	$6 ($10.44)	$8 ($10.00)	$6	$8

Conclusion 2: The context labels do not have the same effect for teams as they had for individuals. Compared to the results for individuals, among teams there are substantially smaller differences between the generic and context conditions in the average dollar value of offers for each pie size, with these differences favoring the generic, as opposed to the context treatment. Among teams the differences in rejection rates between generic and context conditions at each dollar value are essentially the same up to pie size $18, where acceptance rates are substantially

higher for the generic treatment. In contrast, for individuals acceptance rates for the generic treatment were greater for all pie sizes.

We will look to the team dialogues, reported at the end of this section, to better understand the lower acceptance rate at $18 for teams in the context-label condition.

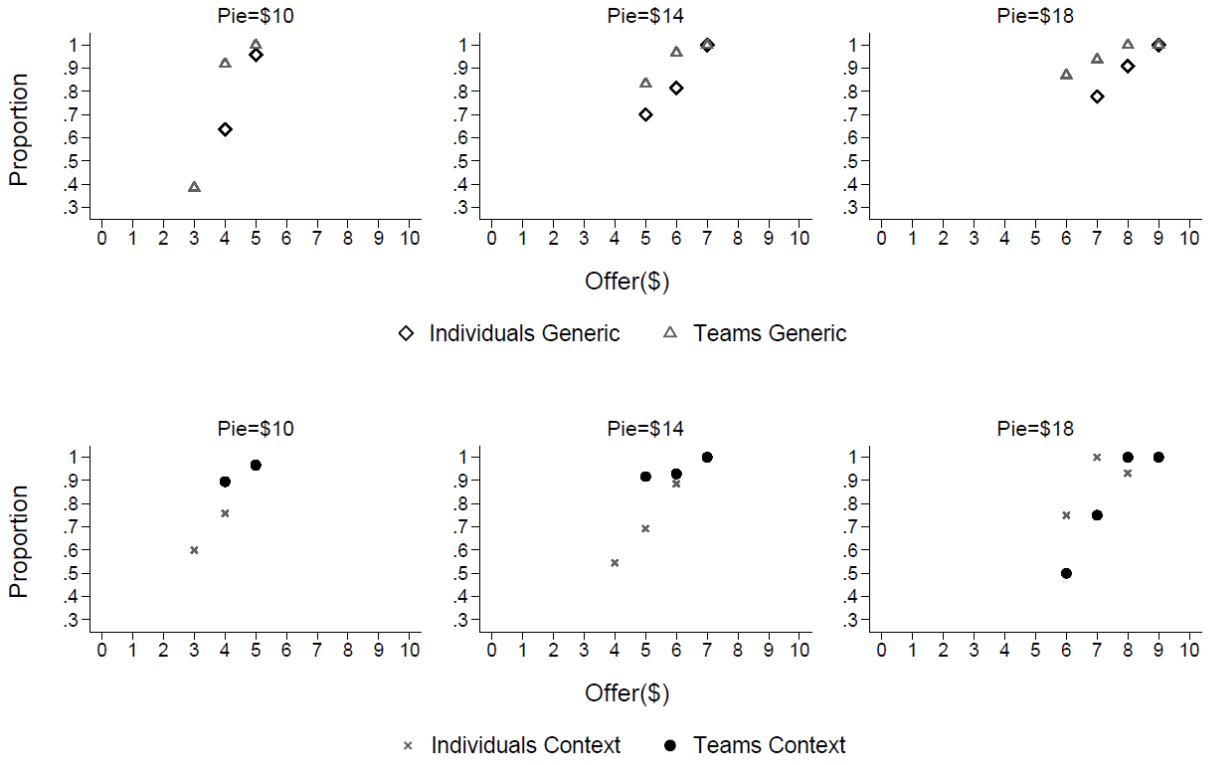

Figure 5: Acceptance rates: teams versus individuals

Table 5

Comparing Teams with Individuals

	Average Dollar Amounts Offered (standard error of the mean in parentheses)					
	Generic			Context		
Pie Size	Individuals	Teams	Indiv-Team	Individuals	Teams	Indiv-Team
$10	$4.38 ($0.91)	$4.01 ($0.68)	$0.37**	$4.01 ($0.85)	$4.35 ($0.54)	-$0.34
$14	$6.15 ($0.78)	$5.73 ($0.78)	$0.42**	$5.50 ($1.07)	$5.92 ($0.93)	-$0.42
$18	$8.35 ($0.98)	$7.10 ($1.11)	$1.25***	$7.06 ($1.62)	$7.63 ($1.20)	-$0.57
	Payoff Maximizing Offers (Proposer's Share in Parentheses)					
	Generic			Context		
Pie Size	Individuals	Teams	Indiv-Team	Individuals	Teams	Indiv-Team
$10	$5.00 ($4.80)	$4.00 ($5.52)	$1.00 (-$0.72)	$5.00 ($4.80)	$4.00 ($5.34)	$1.00 (-$0.54)
$14	$7.00 ($7.00)	$6.00 ($7.76)	$1.00 (-$0.76)	$6.00 ($7.12)	$5.00 ($8.28)	$1.00 ($-1.16)
$18	$8.00 ($9.10)	$6.00 ($10.44)	$2.00 (-$1.34)	$7.00 ($11.00)	$8.00 ($10.00)	-$1.00 ($1.00)

*p <0.10, **p < 0.05, ***p < (0.01) Differences based on Wilcoxon rank-sum test using mean proposer averages as the unit of observation.

Team versus individual behavior: Table 5 compares team and individual behavior within the two treatments. For the generic treatment teams consistently offer significantly lower shares, with a lower payoff maximizing offer as well. The latter causes consistently higher proposer payoffs as rejection rates are the same or higher for individuals for each dollar value (see Figure 5, top panel).

With context, the average dollar share offered by teams is higher at each pie size but none of the differences are statistically significant at conventional levels. Further, with the exception of the $18 pie, the teams' payoff maximizing offer is lower than for individuals, resulting in higher proposer payoffs, as rejection rates are the same or lower at each dollar value (sometimes substantially lower; see Figure 5 bottom panel). The context $18 pie is a real outlier here as

lower offers ($6 and $7) are rejected at much higher rates for teams than for individuals. These higher rejection rates help explain why the payoff maximizing offer is higher at $18 for teams, resulting in lower team payoffs. But this still leaves the substantially higher rejection rates for $18 teams unexplained, and out of line, with the differences in rejection rates at lower pie sizes.

Conclusion 3: Comparing team with individual subject play, generic labeling leads teams to offer significantly less on average, and have the same, or higher, acceptance rates at each offer level. This is consistent with the notion that teams are more "rational" than individuals, as this is closer to (but still far away from) the subgame perfect equilibrium outcome. With context, individuals offer less on average than teams, but these differences are not significant. This is consistent with the idea that there is a lower bound to what can be offered in the ultimatum game with any reasonable chance of being accepted, so that there is essentially no room to go much lower. Teams are also more likely than individuals to accept lower offers at pie sizes $10 and $14, again consistent with the idea that teams are more "rational" than individuals. However, there is an important outlier here at $18, with teams having much higher rejection rates than individuals. This in turn accounts for the higher payoff maximizing offers for individuals with the $18 pie.

Team chat data: Analysis of team chats for the context treatment is reported in Table 6. Four categories were established: (1) Cases where neither of the terms "management" or "labor" are used; (2) Cases where the terms "management" and "labor" were simple placeholders for "proposer" and "responder" as they lacked any meaning beyond what would be associated with the generic terms; (3) Cases where the terms "management" and "labor" evoked meaning related to the sense in which the terms are used outside the lab, meaning which could have potentially affected offers or acceptance of those offers; and (4) Cases where teams clearly "saw through" the labels in the sense of explicitly noting that they had no meaning beyond what generic terms for proposer and responder would be. These counts are recorded irrespective of whether the team was a proposer or a responder. Examples of each of these latter three cases are reported in Table 7. Pooling across pie sizes, coders agreed on the classifications well over 90% of the time. [9]

[9] What disagreements there were primarily involved the frequency with which category 2 was coded.

Table 6

Frequency Teams Engaged in Different Types of Chats

(used one or more times in a session)

Chat Type	Use of Terms Management and Labor	Frequency[1]
1	No mention of terms	20.8%
2	Terms used strictly as placeholders for proposer and responder	41.7%
3	Terms evoked meaning responses as used outside the lab which could have affected offers or responses.	27.1%
4	Teams clearly "saw through' the labels	8.3%

Table 7

Chats used as placeholders (2):

If we ever do get offered $1, I bet it's going to be the one where management gets $18 to play with… If they ever do that, well just have to mug them on the way out. Make them distribute the wealth." "or we just reject that one and take a loss of a dollar" "Losing a dollar to spite somebody out of $17, I like your attitude."

"ok let's do 11 7" "I feel like they have to accept almost anything within reason or nobody gets any money..." "that is right. So we can split more for management and less for labor, maybe?" "i would think so, let's make sure they accept this one and then we can possibly change the next round's amounts. "

Chats evoking meaning responses (3):

"offering them 50% is a bit much, considering the difference in jobs... what do you think?"

"If we were actually in management we would be able to know the cost implications of our decisions, as well as the business demands. i would say up a dollar to 12. I would assume that the plant needs $5 to operate at a good standard, so giving them an extra dollar is the equivalent to giving them a 20% raise"

"We need to form a union. This is an outrage!!!!"

Chats where team "see through" context (4):

"dont let the words, management and labor fool u. they exist to make sense to us cultural wise, think of it as A and B, as equals in which case 50-50"

A somewhat surprisingly (to us at least), a number of teams (20.8%) never used the labels "management" and "labor." Adding these to the frequency with which the terms were strictly

used as labels accounts for 62.5% of the teams. For 27.1% of the teams (13 out of 48) the labels were used to refer to their explicit meaning from outside the laboratory, which could impact their behavior within the experiment. Further, as in the psychology experiments reported on below, the mere labeling and context invoked could well affect behavior, resulting in lower offers. Four teams "saw through" the labels. By this we mean that the team members noted the relation of the labels to the roles of management and labor in the "real world" but also noted that these real world implications should not influence their behavior in the experiment.[10]

We also coded for "economic man" among responders in both the generic and context treatments; these comprised discussions to the effect that more money is better than nothing in responding to offers, and their partner agreed. There is not much difference on this dimension between the two team treatments: 78.3% for context and 82.1% for generic.[11] At the same time, a number of these same teams spoke about rejecting really low offers, with a number of them doing so. Further, a number of teams rejected a particular dollar offer at a larger pie size that they had accepted with a smaller pie: 7 and 11 times for context and generic, respectively.[12] At the same time, there was only one case where a team rejected an offer at a lower pie size, but accepted it with a larger pie.[13] This is consistent with the other regarding preference literature where both absolute and relative payoffs matter (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999).

To summarize: A large number of teams (62.5%) either never used the management-labor labels in their discussions or used the terms strictly as placeholders for the proposer-responder terms in the generic treatment. A significant minority (27.1%) reacted to the context using the meaning of the labels as they are used outside the lab. This may have led them to reject very unequal shares when there was $18 available for profit sharing in "a good year." Few teams explicitly "saw through" the use of meaningful context, explicitly noting that the management-labor terms were designed to "make sense to us cultural wise," but should essentially be ignored.

Discussion

---

[10] Two of these teams had previously used the labels in relationship to their use outside the lab.

[11] These percentages are for cases where one or both coders coded for this. Coders were in agreement over 80% of the time. An example of this is as follows: "accept?" "yep" 'proposer have their advantage):" "yeah they do … i'd definitely do that if i was in their shoes though so i cant even get mad about it"

[12] The corresponding numbers for individuals were 10 and 9 for context and generic treatments.

[13] There were however reversals *within* a given pie size where a dollar offer was initially rejected, but in a later round accepted. This occurred in 3 and 5 instances for teams under context and generic labels, respectively, and 4 times in each treatment for individuals.

The data reported indicate that in large measure teams are more "rational" than individuals in that they come closer to the subgame perfect equilibrium prediction – offering the same or less, and being as willing or more so to accept these lower offers. As a result, proposer teams had higher expected payoffs than individuals at their payoff maximizing offers, with this result more pronounced for the generic case. However, the one clear case where these results do not hold is with context at pie size $18 where teams are more likely to reject low dollar offers (offers of $6 and $7), with proposers earning lower expected payoffs at the payoff maximizing offer. [14] One possibility here is that at pie size $18 context had the most impact as the following quote suggests from a team rejecting one of these low offers: "it was a good year for the company, we as labor, probably made that happen and they wont give us 50%." Recall that the context treatment explicitly noted that the $18 pie was a result of an above average year in an organization with profit sharing, so that this comment makes perfectly good sense in terms of its meaning and implication outside the lab.[15]

Then the question is why this same effect was not observed for the $18 pie with individuals. The psychology literature on team versus individual behavior notes that there is "… substantial evidence that when group members think about themselves as a group (and thus share a social identity) they begin to behave in ways that protect the group from harm or enhance its overall welfare. Many of the implications of this bias are positive …, but there are situations where it prevents the group form making good decisions." (Tindale, Talbot, & Martinez, 2013). One leading example of this is the referred to as the "discontinuity effect" whereby in a repeated prisoner's dilemma game with an uncertain end point individuals tend to cooperate more than teams to begin with. Because it is difficult for teams to re-establish cooperation once it has been violated, team's initial lower cooperation level relative to that of individuals dooms the teams to a lower level of overall cooperation (Schopler & Insko, 1992; Wildschut et al., 2003).[16]

---

[14] Pooling the data a chi-square test for differences in rejection rates for these $6 and $7 offers is significant at the 5% level (two-tailed test) using Fisher's exact test. Treating each session as the unit of observation and conducting a Mann-Whitney test for differences in means, we can reject the null at the 10% level (two tailed). Offers below $6 for the $18 pie (which are rare) were rejected at about the same rate: 3/5 context, 4/8 generic.

[15] In addition there were several calls for going on "strike" for teams in the context treatment.

[16] Kagel and McGee (2005) report on a series of finitely repeated prisoner dilemma games, which indicate that this "discontinuity effect" can be overcome by starting over again with a new opponent. In that experiment, the discontinuity effect holds for the first super-game, but not in subsequent super-games with new opponents. Team dialogues indicate that the breakdown in the first super-game continues even when teams recognize that they would be better off to play cooperatively, recognizing that they have to wait for the next super-game to get started with cooperation to see if their opponent "wants to play nice."

General Discussion

The first principal result of the present paper is that the business context resulted in the solitary proposers making significantly smaller offers which were accepted at the same or higher rates than in the generic context. The net result is that at the payoff maximizing offer, expected proposer payoffs were the same for the generic and business contexts at pie size $10, but larger for "management" at pie size $14 and substantially larger at pie size $18. Related results have been reported in the psychology literature, but unlike the experiment reported here, they did not use financial incentives. Kay, Wheeler, Bargh, and Ross (2004, Experiment 3) presented pictures of business-related objects, such as a briefcase and a boardroom table, to some participants before they assumed the role of proposers in an ultimatum game. Other participants were exposed to pictures of neutral objects. All participants had to rank in terms of height the objects to which they were exposed after which they participated in an ultimatum game. Those who had been exposed to the business-related objects offered less to the responders than did persons who had been exposed to the neutral objects. Kay et al. (2004) hypothesized that objects can cue behavioral norms, and business or economic objects cue the norms of competitiveness and self-interest. Adherence to such norms would result in smaller offers made to responders. We hypothesize that the management/labor labels cued the same behavioral norms as did the economically-related objects in the Kay et al. (2004) research.[17]

Molinsky, Grant, and Margolis (2012, p. 28) summarized these and related results from the psychology literature hypothesizing that instantiating an economic schema reduces compassion for others: ". . . exposure to economics increases the salience of knowledge structures prioritizing rationality, efficiency, and self-interest—and that when these knowledge structures are activated, people act with less compassion."[18] Taken together, these results from the psychology literature, even absent payoffs contingent on behavior, suggest that the management/labor labels we used would diminish the proposers' consideration of the welfare of the responders. Hence the proposers in our economic label group would make smaller offers.

We also looked at how two-person teams behaved in the ultimatum game for the generic and context treatments. Comparing generic teams with generic individuals, teams offered less and were more likely to have their offers accepted than individuals, thus replicating Bornstein

[18] They report two experiments consistent with this hypothesis as well.

and Yaniv's (1998) earlier results. With context the results were more mixed. Compared to management individuals, management teams' expected payoffs at their payoff maximizing offers were larger for the $10 and $14 pies. But at $18 the outcome was reversed, largely as a result of teams rejecting low dollar payoffs at substantially higher rates than individuals. Two conjoined factors would seem to lie behind this last result: (i) the psychology literature on team versus individual behavior notes that considering themselves as a group frequently leads teams to favorable group outcomes, but this consideration can also lead to worse outcomes under certain circumstances, and (ii) meaningful context can prime different ways of framing an issue compared to a generic framework. Specifically, at pie size $18 the priming effect for "labor" rejecting unequal offers would tend to be maximized, at the they were asked to think of the money to be divided as a result of a profit sharing plan with the $18 pie a result of a "good year." These higher rejection rates at pie size $18 are less "rational" than for individuals in that outcomes are further away from the subgame perfect equilibrium prediction. But they may well have contributed to "management" teams offering higher shares than their individual counterparts at $18, resulting in high average payoffs for "labor" in the team treatment.[19]

Our results generally support the notion that the environment in which an economic behavior occurs can have a significant influence on a participant's choices. One might presume that the straightforward incentives in an economic situation such as the ultimatum game would be salient enough to overwhelm the sway of a mere label or the presence of a teammate. However there is a substantial literature showing that the priming of a particular schema or "mental model" can elicit norms and constraints that can significantly affect behavior (Eriksson & Strimling, 2014). By instantiating a management-labor context or by exploiting the implications of group dynamics the impact of incentives can be substantially altered.

---

[19] Vohs, Mead, and Goode (2006, Experiment 6) demonstrated a related result: Compared to those exposed to either neutral cues or no cues, persons who were cued with monetary words versus pictures subsequently donated less to charity.

[19] One might also presume that managers, thinking from the point of view of labor, would offer more equitable shares at $18.

References

Ball, S., Eckel, C., Grossman, P. J., Zame, W. (2001). Status in markets. *Quarterly Journal of Economics, 116*(1), 161-188. doi: 10.1162/003355301556374

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*(2), 230-244.

Batson, D. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology, 29*, 909-924.

Bolton, G. E. & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review, 90 (1)*, 166-193.

Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: Are groups more 'rational' players? *Experimental Economics, 1*(1), 101-108.

Charness, G., & Sutter, M. (2012). Groups make better self-interested decisions. *Journal of Economic Perspectives.* 26 (3): 157-176. doi: 10.1257/jep.26.3.157

Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics, 3*, 55-79.

Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior, 73*, 459-478. doi: 10.1257/jep.26.3.157

Eiser, J. R., & Bhavnani, K-K. (1974). The effect of situational meaning on the behavior of subjects in the prisoner's dilemma game. *European Journal of Social Psychology, 4*(1), 93-97.

Ellingsen, T., Johannesson, M., Mollerstrom, J., & Munkhammar, S. (2012). *Games and Economic Behavior, 76*, 117-130. doi.org/10.1016/j.geb.2012.05.007

Eriksson, K., & Strimling, P. (2014). Spontaneous associations and label framing have similar effects in the public goods game. *Judgment and Decision Making, 9*(5), 360-372.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics, 114 (3)*, 817-868.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments.

Experimental Economics, *10*(2), 171-178. doi: 10.1007/s10683-006-9159-4

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. *Forschung und Wissenschaftliches Rechnen*. Vol. 63 (pp. 79-93). Gottingen, Germany: Gesellschaft fur wissenschaftliche Datenverarbeitung.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367-388.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology, 51*, 1106-1115.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. L. (1994). Preferences, property rights and anonymity in bargaining games. *Games and Economic Behavior, 7*, 346-380.

Kagel, J. H. & McGee, P. (2015). Team versus individual play in finitely repeated prisoner dilemma games. *AEJ: Micro* (in press)

Kay, A. C., & Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construals on behavioral intentions in the prisoner's dilemma. *Journal of Experimental Social Psychology, 39*, 634-643. doi:10.1016/S0022-1031(03)00057-X

Kay, A. C., Wheeler, S. C., Bargh, J. A., & Ross, L. (2004). Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes, 95*, 83-96. doi:10.1016/ j.obhdp.2004.06.003

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*, 623-655. doi: 10.1146/annurev.psych.55.090902.142009

Kugler, T., Kausel E.E., & Kocher M.G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(4): 471-482.

Liberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves.

Personality and Social Psychology Bulletin, 30(9), 1175-1185. doi: 10.1177/0146167204264004

McDaniel, W. C., & Sistrunk, F. (1991). Management dilemmas and decisions: Impact of framing and anticipated responses. *Journal of Conflict Resolution, 35*, 21-42.

Molinsky, A. L., Grant, A. M., & Margolis, J. D. (2012). The bedside manner of homo economicus: How and why priming an economic schema reduces compassion. *Organizational Behavior and Human Decision Processes, 119*, 27-37. doi:10.1016/j.obhdp.2012.05.001

Pillutla, M. M., & Chen, X. (1999). Social norms and cooperation in social dilemmas: The fffects of context and feedback. *Organizational Behavior and Human Decision Processes, 78*, 81-103.

Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics, 88*, 1625-1644. doi:10.1016/S0047-2727(03)00021-5

Roth, A. E, Prashnika, V. Okuno-Fujiwara, M. & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, and Tokyo: an experimental study. *American Economic Review*, *81*, 1068-95.

Schopler, J., & Insko, C. A. (1992). The discontinuity effect in interpersonal and group relations: Generality and mediation. In W. Stroebe & H. Miles (Eds.), *European review of social psychology* (Vol. 3) (pp. 121-151). New York: Wiley.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83.

Tindale, R. S., Talbot, M., & Martinez, R. (2013). Decision making. In J. Levine (Ed.), *Group processes* (pp. 165-192). Hove, United Kingdom: Psychology Press.

Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science,* (17 Nov., *314*), 1154-1156.

Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A., & Schopler, J. (2003). Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect. *Psychological Bulletin, 129*, 698-722. doi: 10.1037/0033-2909.129.5.698

Zhong, C-B., Loewenstein, J., & Murnighan, J. K. (2007). Speaking the same language: The cooperative effects of labeling in the prisoner's dilemma. *Journal of Conflict Resolution, 51*(3), 431-456. doi: 10.1177/0022002707300834