Experimental methods: When and why contextual instructions are important

Aleksandr Alekseev [a], Gary Charness [b,*], Uri Gneezy [c,d]

[a] *Department of Economics, Georgia State University, Atlanta, United States*
[b] *Department of Economics, University of California, Santa Barbara, United States*
[c] *Rady School of Management, University of California, San Diego, United States*
[d] *CREED, University of Amsterdam, The Netherlands*

ARTICLE INFO

ABSTRACT

An important methodological issue in experimental research is the extent to which one should use context-rich or abstract language in the instructions for an experiment. The traditional use of abstract context in experimental economics is commonly viewed as a way to achieve experimental control. However, there are some advantages to using context-framed instructions, such as "employer and worker" instead of "player 1 and player 2." Meaningful context can enhance understanding of an environment and reduce confusion among participants, particularly when a task requires sophisticated reasoning, and hence may yield responses of better quality. In emotionally-charged research questions, such as pollution or bribes, contextual instructions may affect behavior in the experiment, but this effect may be appropriate as it relates to the research question. Our review of the evidence from the literature indicates that in the great majority of cases meaningful language is either useful or produces no change in behavior. Nevertheless, a few important considerations are worth keeping in mind when using rich context. Finally, we see the choice of context as being an expansion of the experimenter's toolkit and a factor to consider in experimental design.

1. Introduction

In one of the seminal papers that paved the way for how we run experiments in economics, Smith (1976) discusses induced values, and mentions a "precautionary corollary". He writes: ". . .the experimenter may be tempted to add 'realism' by giving the abstract experimental commodity a name such as 'wheat', or otherwise attempt to use instructions to simulate the alleged circumstances of a particular market. This runs the danger of so enriching induced values that control over valuation is lost...Consequently, it may be preferable *not* to embellish the instructions with well-intentional attempts at 'realism'. Let the explicit reward structure be the singular source of valuation, insofar as this is possible."

Smith's approach of keeping the instructions abstract and context free became the gold standard in experimental economics. Due to this standard paradigm that context should be avoided and that the structure of the game is paramount,

---

* Corresponding author.
E-mail addresses: aalexeev2@student.gsu.edu (A. Alekseev), charness@econ.ucsb.edu, gary.charness@ucsb.edu (G. Charness), ugneezy@ucsd.edu (U. Gneezy).

research into the effects of context was largely neglected in experimental economics for many years. Yet an important point is that participants in experiments may very well supply their own context for the choice at hand. To the extent that this exogenous context varies across people, supplying only an abstract context may lead to less control than is desirable.[1]

Notable early exceptions to the context-free rule include papers on tax evasion (Alm et al., 1992), public goods (Elliott et al., 1998), and the ratchet effect (Cooper et al., 1999). Below we will discuss these studies and more recent systematic research on the effect of context. Table 1 provides a summary of studies that systematically change context.[2]

It is important to note that the literature we discuss below supports the argument that when instructions are embedded in a meaningful, as opposed to abstract, context, behavior may change. While it is clear that context might affect behavior, the question we address in this paper is whether this is (always) an undesirable effect. In particular, we will survey findings showing that in some cases adding context increases the understanding of the instructions by participants. Calling participants in a negotiation game "Buyer" and "Seller" instead of "Player A" and "Player B" might change the results—but is this changed behavior less (or more) representative of the behavior that we are interested in studying?

One can also think of the use of meaningful context as an additional arrow in the experimenter's quiver. If a researcher were interested whether workplace re-organization has adverse effects on cooperation, it would be useful to have a control treatment where cooperation is prevalent (we shall see that cooperative framing is feasible). She could then consider other treatments where the workplace environment is varied systematically to determine which changes make cooperation more difficult. If one wishes to study which factors ameliorate bubbles in asset markets, having a control treatment where bubbles are present is advisable. In the first case, it turns out that cooperative framing is useful; in the latter case, bubbles turn out to be more likely to manifest in an abstract setting. The researcher can chose the setting that is most useful for her research question.

2. Learning and understanding

The earliest (and arguably most-studied) application of context is enhancing understanding of experimental tasks and speeding up the learning process. It is typically more difficult for most people to operate with abstract terms rather than meaningful ones, especially when a task requires sophisticated reasoning. Decisions of confused people are usually of little interest to researchers, and the use of meaningful terms seems to provide a simple solution to the problem. There is substantial evidence that context does reduce confusion about instructions and leads to better choices.[3]

One of the early examples of this is the study by Wason and Shapiro (1971), who used an individual-choice task that asked participants to evaluate a claim. The task, which represents an example of a logical reasoning problem, has been extensively studied since then and is now known as "Wason's Selection Task." A participant in this task is presented with four cards that lie on a table so that only one side is visible. The cards are labeled with letters and numbers, such as "D, K, 3, 7." Each card is known to have one of these letters on one side and one of these numbers on the other side. The participant is then asked to evaluate the following claim "Every card which has a D on one side has a 3 on the other side."

The evaluation of the claim is done through the selection of two cards that will be flipped to reveal their other side. The two cards that allow validation of the claim are "D" and "7." If the "D" card does not have a "3" on the flip side or the "7" card has a "D" on the flip side, that would falsify the claim. Otherwise, the claim would be validated. Typically only less than 10% of the participants select the right cards (Griggs and Cox, 1982). The most common mistake is to choose the "3" card instead of a "7."

As an attempt to improve performance, Wason and Shapiro (1971) considered replacing the abstract labels with "thematic" ones. Instead of letters and numbers the cards now had cities "Manchester" and "Leeds," and modes of transportation "Car" and "Train" on them. The claim that a new group of participants was asked to validate was presented as "Every time I go to Manchester I travel by car." This change produced a dramatic positive effect on the likelihood of successes, which led the authors to claim that thematic (or meaningful) context may improve understanding of the task.

The problem of confusion with instructions can be particularly acute among certain participant pools. Participants who are easily confused make less sensible choices, making their data less valuable for researchers. On the other hand, the use of meaningful context may improve the quality of data from these subject pools, making them more accessible and generating better knowledge about the behavior of a wider population.

Chou et al. (2009) illustrate this point by looking at a subject pool of community-college students. The study finds that presenting a two-player guessing game in meaningful rather than abstract terms largely improves the performance of community college students. In one treatment, the game was presented in its standard form (abstract terms). Two players simultaneously choose a number between 0 and 100. The player whose number is closer to 2/3 of the average of the two

---

[1] Indeed, Harrison and List (2004) point out: "It is not the case that abstract, context-free experiments provide more general findings if the context itself is relevant to the performance of subjects. In fact . . . there is no control for the context that subjects might themselves impose on the abstract experimental task." Furthermore, Hennig-Schmidt et al. (2010) find that participants themselves construct context in order to make abstract terminology meaningful for them.

[2] We did not attempt to cover every existing paper that systematically varies context in economics studies. Our goal is to provide a summary that illustrates the main usages of context and its effect on participants' behavior over a wide range of scenarios.

[3] It is also true that there is heterogeneity with respect to how people respond to contextual changes. For example, Croson and Gneezy (2009) and Ellingsen et al. (2013) find that women react stronger to men to certain contextual changes.

Table 1
Summary of studies that systematically change context.

Study	Setting	Design	Results
A.Change of context changed behavior			
Wason and Shapiro (1971)	Wason's Selection Task	Abstract treatment used cards with letters and numbers (D, K, 3, 7). The rule was presented as "Every card which has a D on one side has a 3 on the other side." Thematic treatment used cards with cities and modes of transport. The rule was presented as "Every time I go to Manchester I travel by car."	Participants in the Thematic treatment were more likely to select the correct cards to evaluate the rule.
Samuelson and Allison (1994)	Public good/Common pool	Draws from the common pool were organized sequentially, with the first-movers being assigned a role of either "supervisor," "leader," or "guide."	Supervisors took significantly more from the common pool, while leaders and guides took an equal share.
Elliott et al. (1998)	Public good/Common pool	A public goods game with a choice between investing everything or nothing. Participants were primed using "news briefs" on either entrepreneur business strategies or cooperatives business strategies.	Cooperatively primed participants were more likely to contribute to the public fund than entrepreneur-primed participants.
Burnham et al. (2000).	Trust game	In one treatment the players were labeled as "partners," while in the other treatment as "opponents."	In the Partners treatment participants were more trusting and more trustworthy than in the Opponents treatment. Trustworthiness declined over time, and the effect of labeling vanished. The labeling effect on trust persisted.
Cooper and Kagel (2003)	Signaling game	Generic Context treatment used the language "A player," "B player," "A1 type," "A2 type," and did not explain situation modeled. Meaningful Context treatment used "Existing firm," "Other firm," "High cost" and "Low cost" types, decision to enter "this industry or some other industry," and described the situation of a competition between an existing firm and a potential entrant.	Strategic play by incumbents was more frequent in the Meaningful Context Treatment, but the effect waned over time. The play of entrants was unaffected by the context.
Kay and Ross (2003)	Prisoner's Dilemma	Participants were primed using a scrambled-sentence task with either words related to "cooperation" or "competition." Then they were asked to judge the appropriateness of different game labels such as "Community Game" or "Battle of Wits," related to either cooperative or competitive environment, and stated their possible action.	Cooperatively primed participants were more likely to associate the game with cooperative labels and state cooperative intentions.
Liberman et al. (2004)	Prisoner's Dilemma	In the first treatment the game was referred to as "Wall Street Game," and in the second treatment as "Community Game."	Participants cooperated more frequently in the "Community" than in the "Wall Street" game.
Barr and Serra (2009)	Bribery	Abstract treatment used the language of "player A," "player B," "player C," and "offers." Framed treatment used "private citizen," "public official," "other member of society," and "bribe," resp., it mentioned "corrupt services."	Under Framed treatment citizens were less likely to offer bribes, bribe amounts were higher conditional on offering (though not significantly), public officials were more likely to reject bribes (though not significantly).
Chou et al. (2009)	Guessing Game	The Baseline and Simplified treatments involved "choosing numbers." The Battle treatment used the language of a "war" and "battle on a hill," and "choosing how high to locate troops."	Participants were more likely to choose weakly dominant strategy in the Battle treatment than in the Baseline or Simplified treatments.
Cason and Raymond (2011)	Environmental	Neutral treatment used "coupons" being "traded," choosing a "number" that is "reported" to "inspector," and had no references to the situation being modeled. Environmental treatment used the language of "power plant managers" who could "buy permissions to pollute" or incur "abatement costs to avoid emissions" and had to "report pollution amounts" to "regulator," it described the situation being modeled.	In the Environmental treatment transaction volumes and compliance rates were lower, participants underreported pollution levels to the regulator.
Dufwenberg et al. (2011)	Public good/Common pool	Neutral treatment used the language of "experiment." Context treatment used the language of "the community experiment."	Changing labeling of the game from neutral to "community game" affected first and second order beliefs and contributions. Contributions and beliefs about sums of contributions are lower in the labeled game.
Harbring and Irlenbusch (2011)	Contest	Baseline treatment uses "players" of different "types" choosing either a "high" or "low transfer" and a "spread," or "numbers A" or "B." Framing treatment uses "employer" choosing "employment contracts" with a "high wage" or "low wage," and an "employee" choosing a "work intensity" and a "sabotage level."	Employees chose lower sabotage levels in the Framing treatment.
Kirchler et al. (2012)	Asset market	Regular treatments used "stocks" being traded and "pay dividends." Context treatment used "stocks of a depletable gold mine," in which "gold is mined" and "finding gold" is possible with some probability.	Bubbles are significantly smaller in the Context treatment, and participants predict fundamental value dynamics more accurately.

Pevnitskaya and Ryvkin (2013)	Environmental	Neutral treatment used "common stock" and "common stock maintenance cost." Environmental treatment used "pollution" and the "cost of environmental damage."	In the Context treatment participants produce less, pollute less and get higher payoffs. The effect is reduced with experience.
Eriksson and Strimling (2014)	Public good/Common pool	Unlabeled treatment used "a game." After the game participants were asked to associate it with either teamwork or paying taxes. Labeled treatment referred to a game as "The Teamwork Game" or "The Paying Taxes Game." No association was elicited.	Without labels contributions and beliefs about contributions of others were higher among people who associated the game with teamwork rather than paying taxes. In the Labeled treatment contributions and beliefs about others' contributions were higher in the teamwork game. Contributions between associated teamwork and labeled teamwork are similar.
Avrahami et al. (2014)	Contest	Neutral treatment used "players" allocating "points" to different "bins," and a "number" being drawn to determine who wins the prize. Meaningful Context treatment used "managers" operating "tourism companies" that "offered tours" and invested "money in dressing up the tours," and a "consumer" who would choose the most dressed up tour.	All players left more bins empty in the Context treatment, i.e., context helped weak players and hindered strong players. Weak players won more frequently in the Context treatment.

B. Change of context did not change behavior

Alm et al. (1992)	Tax evasion	Neutral treatment used "payment," "check," "disclosed money," and "shortfall." Loaded treatment used "taxes," "audit," "reported income," and "penalty."	No difference in behavior between the loaded and neutral treatments. Participants in the neutral treatment did not associate it with tax game.
Abbink and Hennig-Schmidt (2006)	Bribery	Neutral treatment used "player 1," "player 2," "transfer," choose Y," "choose X." Loaded treatment used "firm," "public official," "private payment," "grant the permission," "do not grant the permission," resp., and gave a real-life situation modeled.	Average bribes were lower, and average rejections were higher in loaded treatment, but both results are not significant.
Banerjee (2014)	Bribery	Neutral treatment used "Participant A" and "Participant B," "transfer" and "ask." Bribery treatment used "Citizen," "Public Official," "bribe" and "demand," resp.	The use of context produced no difference in terms of frequency of bribes demanded or amounts of bribes.

C. Change of context changed behavior in certain cases

Griggs and Cox (1982)	Wason's Selection Task	Experiment 1 employed essentially the same language as in Wason and Shapiro (1971). Thematic treatment in Experiment 3 used the cards with drinking beverages and ages, the rule was presented as "If a person is drinking beer, then the person must be over 19 years of age."	No significant differences between Thematic and Abstract treatments in Experiment 1. In Experiment 3 participants in the Thematic treatment were more likely to select the correct cards.
Cooper et al. (1999)	Ratchet Effect Game	The Generic treatment used "players A" and "B" and different "types" deciding between "X" and "Y" or making "choices." The Context treatment used "planners" deciding between "easy" or "tough production targets" in response to "manager's" choice of "output," where firms can be of "high" or "low productivity."	Firm managers both in the roles of managers and planners developed strategic play faster in the Context treatment. The context had a stronger impact on planners' strategic play. The context had no significant impact on students' behavior.
Krajcova and Ortmann (2008)	Bribery	Benchmark treatment called participants "Participant X" and "Participant Y," actions were labeled "Transfer the endowment," "Not transfer the endowment" and "choice . . .," realizations were labeled as "outcome A" and "outcome B." Context treatment used "Entrepreneur" and "Bureaucrat," "Pay bribe," "Not pay bribe," "Denounce," "Do nothing," "Provide the favor," "Corruption has been detected" and "Corruption has not been detected."	Context had no effect on bribing frequency in the aggregate data. For the low endowment males increased bribing in the Context treatment, females did the opposite. For the high endowment males reduced bribing and females increased bribing in the Context treatment (though not significantly).
Cooper and Kagel (2009)	Limit pricing	Abstract treatment used "players" of different "types" choosing "numbers" or actions "X" or "Y," no reference to the situation modeled was given. Context treatment used "existing firm" with "high" or "low costs" choosing "output level," and "other firm" choosing to "enter this or other industry."	Meaningful context enhanced cross-game learning through higher strategic play. When combined with "superficial" changes to the game structure the effect is the opposite. Context had no effect for teams of two players.
Ellingsen et al. (2012)	Prisoner's Dilemma	In Treatment 1 the game was called the "Stock Market Game," in Treatment 2 the game was called the "Community Game."	Participants cooperated more frequently in the "Community" than in the "Stock Market" game. The effect disappeared when a computer chose the actions of other participants.

numbers wins. The weakly dominant strategy is to choose 0. In the second treatment, the game is transformed into a simple isomorphic game and presented in a familiar context (a battle on a hill with two opponents deciding how high to locate their troops). The participants in the second treatment are significantly more likely to choose weakly-dominant strategies than the participants who were presented with the standard form, even when the standard form was supplemented with a hint of what is the best action.

The importance of understanding an experimental task and the role of context in it becomes clear in cases when reduced confusion challenges a well-established result. Kirchler et al. (2012) report that a change in language produces a large and significant reduction in bubbles in the experimental asset markets. Their setting is the canonical asset-market design of Smith et al. (1988), in which participants are endowed with cash and assets that pay stochastic dividends in every period. Participants trade assets over the course of several periods. The fundamental value of an asset declines over time, since this value consists only of the (stochastic) dividends to be generated in each remaining period. Numerous studies find that bubbles occur, i.e., assets are traded well above their fundamental value, followed by crashes.

In their baseline treatment, Kirchler et al. use the standard language of "assets" and "fundamental value." Motivated by the fact that most participants have difficulties with understanding the concept of a declining fundamental value, they introduce another treatment, in which the assets are framed as "gold mines," and the amount of "gold" in them "depletes" over time. This change in the language of the instructions produces a large decrease in the formation of bubbles and subsequent crashes. This effect is robust to changes that have been shown to rekindle the bubbles, such as increasing cash-to-assets ratios or dividend variances.

The use of meaningful context can be beneficial in strategic settings, since these often require sophisticated logic to make the right choice. The use of context is especially beneficial in signaling games. The equilibrium solution for these games relies on forward induction, which requires the ability to think strategically.

Cooper and Kagel (2003) provide evidence for this in an experiment that uses a signaling game modeling limit-pricing behavior. There are two players in the game: an incumbent monopolist and a potential entrant. The monopolist has either high or low costs. It makes sense for the entrant to enter the market only if the monopolist has high costs. The entrant, however, cannot observe the costs directly, but can observe a signal given by the monopolist's output. Knowing this, the monopolist can strategically overproduce (use limit pricing) to make the entrant believe that she has low costs. Cooper and Kagel compare the two treatments, one of which uses generic terms such as "A player," "B player," "A1 type," "A2 type." The other treatment uses meaningful terms such as "Existing firm," "Other firm," "High cost," "Low cost," decision to enter "this industry or some other industry," and explains the situation modeled. In the treatment with richer context, high-cost monopolists use limit pricing more frequently, and more low-cost monopolists try limit pricing at least once, which is interpreted as the learning effect of context. The effect is substituted for experience as the game progresses.

Cooper and Kagel (2009) take the point further by looking at cross-game learning between two signaling games. The games are similar to the one in Cooper and Kagel (2003), except for a few small changes. The parameters of the first game generate a pooling equilibrium, while the parameters of the second one generate a separating equilibrium. The study looks at how the experience with the first game translates into strategic behavior in the second one. The participants are given either abstract or meaningful contexts, which use the same language as above. The study finds that the participants who are given the meaningful context show higher levels of strategic play in the second game, as compared to the participants who are given the abstract context.

Avrahami et al. (2014) provide another example of the benefit of meaningful context in strategic settings. They study a contest setting with asymmetric players and find that the weak players made better choices in the labeled treatment. In their game, two players select how many points to allocate among several bins. One of the bins is selected at random, and the player who has more tokens in this bin wins. The study introduces an asymmetry between the players regarding their total number of tokens, so that there are weak and strong players. It is optimal for the weak players to leave more bins empty. The baseline treatment used the abstract terms such as "bins and points," while the labeled treatment described a situation of two "tour companies" who were competing for clients by presenting the "tours" of different attractiveness. The study finds that the meaningful terms help the weak players, who leave more bins empty and win more frequently.

The effect of context is not always positive. There are cases when it leads to worse choices. In the Avrahami et al. (2014) study, the use of meaningful context makes the strong players leave more bins empty, which they should not do. As a result, strong players in the labeled treatment lose more often to weak players.

The second experiment in Cooper and Kagel (2009) looks at what happens when a change in context is combined with "superficial" changes to the game structure. The participants with experience in the first game (in abstract context) are then given another game with different parameter values, and the choices are made over "prices" rather than "quantities". These changes leave the structure of the game identical to the first game. In one treatment, they present the participants with a new game using the same abstract terms; in a second treatment, the game is labeled. The results indicate that the participants in the second case developed strategic play more slowly than the participants in the first case. By the end of the session, however, both groups show similar levels of strategic play. This finding points towards the danger of changing the context within a session.

In some cases, meaningful context does not have any effect on the understanding of a task or strategic thinking. Griggs and Cox (1982) report that in two of three experiments they failed to replicate the earlier results of the use of thematic materials in the Wason selection task described above. In their first experiment, they used American cities (Miami and Atlanta) and car and plane as modes of transportation. A typical claim was stated as "Every time I go to Miami I travel by car." In their second

experiment, they used envelopes that were either sealed or not and stamps of different value. A typical claim was stated as "If a letter is sealed, then it has a 15-cent stamp on it." Neither thematic material improved participants' performance significantly.

In a strategic setting, Cooper et al. (1999) do not find that meaningful context leads to better choices in a game designed to model the "ratchet effect," when the participants were Chinese students. The two players are a firm and a central planner. A firm can have either a low or a high productivity. The central planner does not observe the firm's type, but can try to infer it from the firm's output. A high-productivity firm can anticipate it and try to mimic a low-productivity firm by producing less in the first period to avoid ratcheting up the production targets in the second period. The game was presented in either abstract ("player A," "choose X") or meaningful ("manager," "planner," "choose production target") terms. The use of meaningful context did not have any significant effect on the students' behavior.

The somewhat-mixed results on the effect of meaningful context lead some researchers to the conclusion that the "better-understanding" effect of meaningful context works through memory cues or similarities to past experiences.[4] Participants who are experienced in a particular field can benefit from the use of context if it links their expertise to the content of an experiment.

Griggs and Cox (1982) argue that the failure to replicate the positive effect of thematic materials is caused by the participants' inability to relate their experience to the content of the Wason selection task. In Experiment 3, they make use of the language and situations that are familiar to their subject pool, represented by undergraduate students in Florida 18–22 years of age. The situation was presented in terms of drinks and age. The four cards were labeled as "drinking beer," "drinking coke," "16 years of age," and "22 years of age." The claim to be validated was presented as "If a person is drinking beer, then the person must be over 19 years of age." Participants in this treatment were much more likely to identify the correct cards than in the abstract treatment. Griggs and Cox attribute the success to the familiarity with the situation. In Florida at the time, it was against the law to drink alcoholic beverages for people under 19 years of age, which students were very likely to know. This language allowed the participants to invoke their experience quickly and solve the problem. On the other hand, the environment of letters and stamps used in Experiment 2 of the study was not familiar to students, and therefore did not produce any effect.

Cooper et al. (1999) arrive at a similar conclusion by comparing the strategic play of students versus managers. The meaningful context did substantially improve the strategic play among managers. Perhaps this was because managers were familiar with the situation as described and could invoke past experience to guide their actions in the game. In fact, Cooper et al. use this result to claim that Chinese managers could have been subject to ratcheting from the central government. On the other hand, students were very unlikely to be familiar with the situation of ratcheting, and therefore could not use memory cues to facilitate the strategic thinking.

The "memory cues" hypothesis raises some important concerns about the use of meaningful context. Ideally, the use of meaningful terms should make a decision task more accessible for participants, so that they make better choices. Replacing abstract terms with meaningful ones, however, can have unintended consequences. The context may replace logic with a merely mechanical application of familiar terms to the unfamiliar situation. Manktelow and Evans (1979) raise this point in relation to a large body of studies on the Wason selection task, and the mixed evidence on the success of context in them. They argue that context is a double-edged sword, as it can either facilitate understanding or impede thinking.

Whether or not this becomes a problem depends on the research question. This might create a problem if a researcher is interested in participants' deliberate decisions, as is the case in the studies on reasoning problems, such as Wason's selection task. This does not create a problem, however, if a researcher is, in fact, interested in participants' past experiences. For example, this is the case in the Cooper et al. (1999) study, in which one of the research questions is whether Chinese managers were participant to ratcheting.

3. Changing social behavior

Another application of context is to affect the social component of games, as perceived by participants. In such settings as pollution, bribery, or financial decision-making, the use of abstract terms may eliminate important social considerations that are inherent to them in naturally-occurring situations. The games, which model these scenarios but do not use their language, may yield results that have low external validity.

The importance of the social component of games, as opposed to their underlying payoff structure, has been demonstrated in various studies. The most well-known example is, perhaps, the Prisoner's Dilemma game studied by Liberman et al. (2004), who found that a mere change in the game name produced significant changes in behavior. In an identical Prisoner's Dilemma game, participants were much more likely to cooperate when the game was labeled "The Community" game than when it was labeled "The Wall Street" game. In Burnham et al. (2000), the participants showed significantly higher rates of trusting and trustworthy behavior in a trust game when they were called "partners," than when they were called "opponents."

The way the social component affects behavior depends on participants' perceptions of the language used to frame it. In the previous example the title "Community" had a pro-social connotation, while the title "The Wall Street" connoted greed.

---

[4] The idea of similarity in decision-making is developed formally in Gilboa and Schmeidler (1995) as an alternative theory of choice under uncertainty. Jehiel (2005) develops an equilibrium concept based on similarity and shows that it can explain behaviors in a variety of multistage games.

However, in general the connotation of a particular title may depend on participants' cultural background. In some instances, the word "community" has a negative connotation, as the next example shows. Dufwenberg et al. (2011) study the effect of labeling in common pool/public good games. Both types of games are conducted in a standard fashion, but the parameters are chosen such that the games are payoff-equivalent. The only variation in the language of the instructions is that the experiment is referred to as simply "the experiment" (neutral treatment) in one case and as "the community experiment" (labeled treatment) in the other. The study finds that labeling affected both participants' beliefs and their contributions, with both lower in the labeled treatment. This result conflicts with the Liberman et al. (2004) result. Dufwenberg et al. explain this by the negative connotation the word "community" has in their subject pool (undergraduate students in Germany). They report that conducting the same experiment in Switzerland reversed the results.[5]

The presence of the social component and its effect on participant behavior depend on the way the social interaction is organized, and not on the structure of the game itself. Elimination of the possibility for participants to interact may eliminate the social component as shown by Ellingsen et al. (2012), who also study the behavior in prisoner's dilemma. Following the previous research, they change only the name of the game, which is either a "Stock market game" or a "Community game." The difference is, however, that participants in a pair have asymmetric roles. Only one participant in a pair makes an active decision of cooperation versus defection, while the other one is passive, with his or her actions chosen by a computer, based on the actual frequency of choices in other treatments. The study finds no difference in cooperation rates between the two treatments. This result is used to argue that labeling might work not through the change of social preferences, but through the change of beliefs about what others are likely to do.

Barr and Serra (2009) show that the use of meaningful context makes the social component salient, which is not the case when the game is framed in abstract terms. They study a bribing scenario, in which bribes impose negative externalities on other players. In their game, there are three player types: "citizens," "public officials," and "other members of society". Citizens can offer bribes to public officials, and in doing so they incur a fixed cost regardless of whether a bribe is accepted or not, and gain monetary value if a bribe is accepted. Upon receiving a bribe offer, public officials can either accept or reject it. Accepting a bribe implies supplying a corrupt service. Each other member of society incurs costs that depend on the number of successful instances of bribing. Bribing, therefore, creates a negative externality, which is common information. If players are selfish, the equilibrium prediction is universal bribing in the game, which leaves other members with the highest negative externality. If, however, agents incur psychological costs of bribing caused by the negative externality imposed on others, this could lead to less bribing. The study varies the psychological costs by changing the language of the game. The framed treatment used the meaningful terms as above. The abstract treatment used terms such as "Player A" and "offers." As predicted, participants were less likely to offer bribes with the richer context.

Pollution is another scenario in which an activity may cause negative externalities. Pevnitskaya and Ryvkin (2013) show that the use of meaningful context can be successful in manipulating participants' psychological costs associated with negative externalities in this scenario as well. In their dynamic game, each participant chooses between production and consumption. Production yields more benefits than consumption, but it also adds to the pollution level in every period. Only a fraction of this level is carried over across the periods, however, since the environment partially regenerates each period. Each participant bears the costs of the environmental damage as a fraction of the accumulated pollution. The study compares participants' production choices across the two treatments that differ in their language. The environmental treatment used the terms as above. The neutral treatment replaced these terms with "common stock" and "common stock maintenance cost." In the environmental treatment participants produced (and hence, polluted) less and had higher payoffs than in the neutral treatment.

Harbring and Irlenbusch (2011) provide another example of higher saliency of psychological costs associated with negative externalities in the meaningful context. They study the effect of context in a principal-agent setting with sabotage. In their game, a principal offers wage contracts, which specify the amount that a winner and the non-winners would get, based on who delivers the highest output. The agents can participate in two activities, one of which is productive (exerting effort) and the other is destructive (sabotage). Exerting effort leads to higher output for the exerting agent, while sabotage reduces outputs for all other players, thus imposing a negative externality. Participants' choices of sabotage levels are compared across the two treatments that vary the language of the game. The framed treatment used the meaningful language described above, while the baseline treatment used completely abstract terms. The study finds that the meaningful language (calling sabotage by its name) significantly reduces sabotage.

In some cases, the social component may be salient enough even with the abstract language. If a game has a simple structure, participants may infer the situation being modeled and thus implicitly associate the game with a particular scenario. In these cases, a change of language produces only a weak effect, if any at all. Abbink and Hennig-Schmidt (2006), who investigate the role of context in a bribery setting that differs from the one in Barr and Serra (2009), demonstrate this point. In their game, which has a resemblance to trust and reciprocity relations, there are two players: "a firm" and "a public official." The firm moves first by deciding whether to send "private payment" (i.e., a bribe) and in what amount. A bribe offer is associated with a small fixed cost, which does not depend on the public official's decision. If the public official accepts the bribe, it is subtracted from the firm's account and after being tripled is added to the official's account. Accepting the

---

[5] Dufwenberg et al. (2011) argue that a subtle interplay between framing, beliefs and choices exists; frames may affect beliefs, and beliefs shape motivation and choice. Context may induce cues that affect choices and behavior.

bribe leads to a tiny chance that both players will be terminated from the game, which represents being caught. Finally, regardless of whether the bribe was offered or not, the public official decides whether to grant the firm a "permission to build a factory" or not. A favorable decision leads to small costs that are incurred by every participant in the session, which represents a negative externality. The framed treatment used the above language, while the neutral treatment replaced all the terms with abstract ones, such as "player 1," "transfer," and "choose X." The study does not find significant differences in bribe offers and permission decisions between the treatments.

Banerjee (2014) also shows that a change of language does not influence behavior in a simple game with a clear structure. The study also looks at a bribery setting based on an ultimatum game, but offers a variation on how the situation is modeled. In this variant, the two players are a "citizen" and a "public official." The citizen first completes a real-effort task to earn income. The public official judges the outcome of the task. If the citizen reaches a specified performance standard, the public official has an option to demand a bribe, so that the citizen could then receive his or her earnings. If the citizen declines the demand, he or she receives nothing; otherwise, the citizen receives the prize for the successful completion of a task less the bribe amount. The public official receives a fixed wage and a bribe if the demand is accepted. The neutral treatment replaces all of the above terms with abstract ones, such as "participant A," "transfer," and "ask." The frequency of bribe demands and bribe amounts are not significantly different between the labeled and neutral treatments.

A similar result is found in settings other than bribery. Alm et al. (1992) look at participants' behavior in a tax-compliance game. After receiving an endowment, each participant decides how much of his or her "income" to report to the regulator. "Reported income" is subject to "taxes" that are subtracted from it. "Audit" is performed randomly, and "penalty" applies in the case of underreporting. The labeled treatment used the language as above, while the baseline treatment replaced it with still meaningful, but less tax-related words, such as "payment," "check," "disclosed money," and "shortfall." The study reports that there are no significant differences in tax compliance rates between participants in the baseline and labeled treatments.

A change of language may fail to change behavior for another reason. If the social component is simply absent from a game, it does not matter whether the game is framed in abstract or meaningful terms. Context is not a magic bullet that changes behavior by definition; it must be supported by the underlying game structure. Recall that in Barr and Serra (2009) the social component works through the psychological costs of negative externalities imposed on others. If these negative externalities are not part of the game, there is no way to make the psychological costs more salient.

Krajcova and Ortmann (2008) illustrate this consideration in another variation on a bribery scenario. The "entrepreneur" moves first by deciding whether to "pay a bribe" to the "bureaucrat." The bureaucrat then has three options to choose from: to "accept a bribe", "do nothing" or "denounce the bribe." Denouncing the bribe calls for an "audit", which detects the corruption with a certain probability. If the corruption is detected the entrepreneur pays the full "fine" amount, while the bureaucrat pays the reduced fine. Doing nothing passes the choice to the entrepreneur who has the options of doing nothing or denouncing. Denouncing works as before, but this time, the entrepreneur incurs reduced fines. If the bureaucrat accepts the bribe, the turn goes to the entrepreneur, who has the same to options, as before. If the entrepreneur does nothing (in either case), the audit can still be conducted with some probability. The entrepreneur benefits from accepting the bribe, in which case "the favor is provided," however, no externalities are imposed. The neutral treatment replaces all the meaningful terms described above with the abstract ones such as "participants" and "transfers." The study finds that the labeling manipulation does not produce significant differences in bribing choices, although there are some gender differences.

In cases when the social component is absent from a game, the use of meaningful terms may even lead to behavior that appears anti-social. Cason and Raymond (2011) investigate the effect of labeling in an environmental setting based on trading pollution permits. Each participant represents "a manager of a power plant" that produces a fixed amount of "electricity." Production generates "pollution," which can be "abated" at a cost. Participants differ in their abatement costs. Each participant is endowed with pollution permits, which can be traded in a continuous double-auction market. Pollution above the level allowed by the permits holdings must be abated. The heterogeneity in abatement costs creates gains from trade so that high-cost types have an incentive to buy permits from low-cost types. After the trade, participants select their abatement levels and report them to the "regulator." Each participant is monitored with a certain probability, in which case a penalty applies if under-reporting is found.

The environmental treatment uses the language above, while in the neutral treatment participants trade "coupons," "chose numbers" and report them to the "inspector." The study finds that participants comply less in the environmentally-labeled treatment. They under-report their levels of pollution to the regulator, demonstrating lower levels of environmentally-friendly behavior, which is the opposite of the effect found in Pevnitskaya and Ryvkin (2013). The important difference is, however, that participants do not have negative externalities (that is, other players' payoffs are not reduced) from pollution in Cason and Raymond (2011), so that social preferences would not appear to apply.

These papers in ensemble suggest that there are two competing explanations for the "pro-social" effect of context. The first relates to triggering pro-social mode among participants, i.e., context works through a change of preferences. The use of such words as "cooperative," "partner," or "community" primes participants to think of the game in more pro-social terms. In fact, even subtle priming before the main task was demonstrated to have an impact on choices. Elliott et al. (1998) asked participants to read several news briefs that talked about the benefit of either cooperative or competitive approaches to business. Participants who were primed cooperatively were more likely to contribute to a public good. The participants in Kay and Ross (2003) first unscrambled sentences with cooperative or competitive-sounding words and were then asked to associate the prisoner's dilemma with a few given labels, and also to state their likely action.

Table 2
Summary of studies that are explicit about their choice of context.

Study	Setting	Context Description	Justification
Cox and Oaxaca (1989)	Job search	"Avoid emotive terms"	Obtain results with unambiguous interpretation
Brandts and Cooper (2006)	Coordination game	"Corporate context without strong connotations"	Enhance understanding of instructions
Landeo et al. (2007)	Signaling ultimatum game	"Neutral labels"	Realistic labels would not help understanding of the game and introduce noise through home-grown values
Alatas et al. (2009)	Bribery	"Emotive terms"	Simulate real-life situation
Lambsdorff and Frank (2011)	Bribery	"Non-neutral"	Enhance understanding of instructions
Berninghaus et al. (2013)	Coordination game	"Unloaded frame"	Avoid framing effects
Engel et al. (2013)	Bribery	Quasi-loaded	Not interested in framing effects
Nosenzo (2013)	Gift-exchange game	"Labor market framing"	More natural
Li et al. (2013)	Organ donation	"Descriptive framing"	Evoke psychological costs associated with organ donation
Brick et al. (2013)	Environmental	"Climate change context"	Encapsulate dimensions of real-life choice
Rivas (2013)	Bribery	"Neutral terms"	Tradition
Gibson et al. (2013)	Preferences for truthfulness	"Accounting earnings management"	Simulate real-life situation
Abbink et al. (2014)	Bribery	"In-context language"	Evoke emotional and moral responses, enhance external validity
Thöni and Gächter (2014)	Gift-exchange game	"Buyer-seller terminology"	More neutral than labor relations context
Bremzen et al. (2015)	Principal-agent	"Framed in terms of labor market"	Enhance learning

Cooperatively-primed participants were more likely to associate the game with cooperative labels, such as "Community Game," and select cooperation.

The second explanation holds that the language used in the instructions serves as a coordination device among participants. In other words, context works not through a change in preferences but through a change in beliefs about what others are likely to do. Participants in a "Community Game" are more likely to think that others will cooperate and therefore cooperate themselves, preferring mutual cooperation to unilateral defection. Ellingsen et al. (2012) provide direct evidence in support of this hypothesis in the Prisoner's Dilemma. When active participants were playing with passive participants, whose actions were chosen by a computer based on the previous play, the pro-social effect of the cooperative frame disappeared. Dufwenberg et al. (2011) provide evidence that labeling changes participants' first and second-order beliefs when their public-good game is labeled as a "community experiment." Participants contributed less and believed that others would also contribute less in the labeled treatment. Eriksson and Strimling (2014) show that Prisoner's Dilemma labeled as the "Teamwork Game" induces both higher contributions and beliefs about others' contributions among participants.

4. Not all contexts are created equal

Table 2 describes recent papers that are explicit about their use of context. The table reports the term that was used to describe a context of an experiment and a justification for it. Unlike abstract description, the terminology used is not restricted.

We distinguish between three levels (or types) of context. These levels differ in their strength and applicability for particular needs. The first level, which can be called "abstract," employs terms like "player A," "type X," "choose B," and so on: The language is abstract and is not related to any experience outside of the lab. This level is sometimes referred to in the literature as "neutral."

The second level, which can be called "meaningful," employs terms that can be related to a real-life situation, such as "buyers and sellers," or "trading." These terms have a meaning, but do not evoke strong emotions or connotations. Meaningful context can enhance understanding of an environment and reduce confusion among participants, making choices more consistent or strategic. The use of meaningful terms is particularly beneficial when a task requires sophisticated reasoning. Games that rely on forward induction, such as signaling games, or dynamic games, such as asset market games, are good examples for this. Meaningful terms may prove beneficial when dealing with complicated instructions. In this case, the context may yield less confused responses. While there are rare instances when the use of meaningful language produces more confusion among participants, the overall effect is positive. Often the meaningful language will produce no change in behavior.[6]

When applying meaningful terms for the purpose of reducing confusion, a few important considerations are worth keeping in mind. Meaningful and familiar language is useful if the researcher is interested in deliberative decisions, but there is a danger of invoking automatic responses due to subjects' past experiences. Care should be taken to avoid this.

---

[6] Of course, the instructions used should always be readily available, so that the reader can ascertain whether contextual instructions are used in a manipulative way.

The third level, which can be called "evocative," includes terms that not only relate to a real-life situation, but also can evoke strong emotions. Examples of such a language would be a "bribe," "climate change," or "organ donation". The line between the second and the third levels is clearly thin. It is hard to argue that a particular choice of a meaningful context does not resonate strongly with any participant under any circumstances. For example, in the gift-exchange game a "neutral" context is often used that presents an experiment as an interaction between "buyers and sellers," while in reality it tells a story of a principal-agent relation. It is not obvious that a "buyer-seller" relation will evoke less emotional response than a "worker-employer" relation. But certain contexts are more likely to evoke emotional responses – for example, people may well be more emotionally engaged in a context of "climate change" than in a "buyer-seller" relationship. And we suspect it could matter whether one faces an "opponent" or a "partner" in the game, as was found in Burnham et al. (2000).

While evocative framing may alter the results relative to an abstract one, this change may increase the external validity of the results by increasing the emotional charge in the direction of interest for the experimenter, making the conclusions more applicable to the situation under study. Examples of such cases include bribery, pollution, financial decision-making, coordination games and common pool/public good games. All of these cases typically feature the social component in the form of negative externalities, which might involve psychological costs. Using evocative terms makes the psychological costs more salient and can lead to behavior that better reflects what happens in naturally-occurring environments.

Consider the following example by Dohmen et al. (2009), who study risk attitudes using a representative survey and a complementary experiment conducted with a representative subject pool. They find that asking people about their willingness to take risks "in general" is a good predictor of actual behavior overall. However, the best predictor of specific behaviors comes from specific questions; if one wants to predict smoking behavior, asking him or her about willingness to take risks in health matters, rather than the general risk question or questions incorporating different contexts, results in the strongest correlation.

The use of evocative terms is more likely to be irrelevant in simple games, whose structure is suggestive of the situation being modeled. Particular forms of bribery games and tax-compliance games have very clear structures that leave no doubt about the underlying scenario, even when the instructions are framed in abstract terms. However, it might be hard to know *a priori* whether the participants interpret the game "correctly." As mentioned earlier, leaving the context of the game to the participants' may in fact imply losing some degree of experimental control (Harrison and List, 2004). Extensive evidence suggests that the way participants play the game is affected by their perception of it. For example, Yamagishi et al. (2013) conducted a large-scale study of the Japanese population using different two-player games. Participants were more likely to behave cooperatively in these games when they viewed them as similar to cooperative environments. Eriksson and Strimling (2014) show that cooperative behavior among the participants who faced "The Teamwork Game" in a public-goods game was similar to that of the participants who implicitly (spontaneously) associated the game with teamwork. Engel and Rand (2014) find that play in a Prisoner's Dilemma framed using abstract context is similar to play in the game framed in cooperative context, suggesting that participants' implicit association of the game was that of a cooperative nature.

Evocative terms should be used carefully, since they can interact with the participants' observable characteristics. It is worthwhile to take account of the participants' cultural background, since similar labels have different connotations in different cultures. Additionally, there is evidence that women can be more sensitive to the context of an experiment than men (Croson and Gneezy, 2009).

5. Conclusion: reasons (not) to use context

The traditional use of abstract context in experimental economics is commonly viewed as a way to achieve experimental control, since using meaningful context can affect behavior through participants' home-grown values. Abstract context might also help in replicating results across cultures, languages, and populations in which contextual instructions might have a different meaning.

Yet, there are advantages to using context-framed instructions. For example, participants may implicitly frame the same game in varying ways, resulting in very different outcomes. The use of explicit association via context, therefore, can help gain experimental control, since participants' implicit associations are rarely elicited in economic experiments. In addition, experimental participants may often bring their own context into the laboratory. Since this likely to vary across the population, an abstract context may potentially weaken the experimenter's control over the environment.

And of course, not all contexts are created equal. The "meaningful" context can help in understanding the situations and might be related to the question of interest, such as "buyers and sellers," or "trading." While stepping away from the abstract description, this type of context seems natural and "innocent" with respect to the researcher choosing to use it. Most of our conclusions refer to this type of context. On the other hand, what we called the third type of context, the "evocative" one, presents a challenge in the sense that different framing of the same problem can evoke strong emotions. Researchers who choose to use such language should be ready to explain why it was the most appropriate for the research question. The risk is that allowing different framing of the emotional aspect of the instructions increases the degrees of freedom of the researchers, and may be used to manipulate results. Therefore, researchers should be clear about their choice of context and the justification for their choice, and an outside observer should be able to verify it by having access to participant instructions. Nevertheless, our overall conclusion from the literature is that using evocative language either does not affect behavior or affects it in a desirable way by evoking the desired emotional response.

This additional degree of freedom that context provides, when used properly and transparently, can be a useful tool for experimental design. Context can create the desired control group for subsequent treatments. For example, a researcher who wants to start with the baseline of high cooperation may use cooperative labels in a public goods game, and then introduce treatments of interest that affect the level of cooperation. Whether to use abstract or meaningful context may depend on the goals of the researcher. In some sense, we are advocating an expansion of the experimenter's toolkit. We do not contend that one approach or the other should always be used, but the issue of context should be an important part of the design process.

To summarize, the recommendation from surveying the literature is for experimental economists to consider using meaningful and even evocative language in the instructions in many experimental environments. When studying a gift-exchange environment, calling participants "Player A" and "Player B" might result in the lower understanding of the environment relative to "Employer" and "Worker." It might also provoke a different type of preferences than the one the researcher is interested in: If the less abstract words do affect behavior, it is not clear that this effect is undesirable—after all, the goal of the experiment is to study an employer-worker relationship!

References

Abbink, Klaus, Hennig-Schmidt, Heike, 2006. Neutral versus loaded instructions in a bribery experiment. Exp. Econ. 9 (2), 103–121.
Abbink, Klaus, Dasgupta, Utteeyo, Gangadharan, Lata, Jain, Tarun, 2014. Letting the briber go free: an experiment on mitigating harassment bribes. J. Public Econ. 111, 17–28.
Alatas, Vivi, Cameron, Lisa, Chaudhuri, Ananish, Erkal, Nisvan, Gangadharan, Lata, 2009. Gender, culture, and corruption: insights from an experimental analysis. South. Econ. J. 75 (3), 663–680.
Alm, James, Gary McClelland, H., William Schulze, D., 1992. Why do people pay taxes? J. Public Econ. 48 (1), 21–38.
Avrahami, Judith, Kareev, Yaakov, Todd, Peter M., Silverman, Boaz, 2014. Allocation of resources in asymmetric competitions: how do the weak maintain a chance of winning? J. Econ. Psychol. 42, 161–174.
Banerjee, Ritwik, 2014. On the interpretation of bribery in a laboratory corruption game: moral frames and social norms. Exp. Econ. 2, 1–28.
Barr, Abigail, Serra, Danila, 2009. The effects of externalities and framing on bribery in a petty corruption experiment. Exp. Econ. 12 (4), 488–503.
Berninghaus, Siegfried K., Haller, Sven, Krüger, Tyll, Neumann, Thomas, Schosser, Stephan, Vogt, Bodo, 2013. Risk attitude, beliefs, and information in a corruption game – an experimental analysis. J. Econ. Psychol. 34, 46–60.
Brandts, Jordi, Cooper, David J., 2006. Observability and overcoming coordination failure in organizations: an experimental study. Exp. Econ. 9 (4), 407–423.
Bremzen, Andrei, Khokhlova, Elena, Suvorov, Anton, Van de Ven, Jeroen, 2015. Bad news: an experimental study on the informational effects of rewards. Rev. Econ. Stat. 97 (1), 55–70.
Brick, Kerri, Visser, Martine, Van der Hoven, Zoe, 2013. Cooperation and climate change: can communication facilitate the provision of public goods in heterogeneous settings? Environ. Resour. Econ., 1–23.
Burnham, Terence, McCabe, Kevin, Smith, Vernon L., 2000. Friend-or-foe intentionality priming in an extensive form trust game. J. Econ. Behav. Organ. 43 (1), 57–73.
Cason, Timothy N., Raymond, Leigh, 2011. Framing effects in an emissions trading experiment with voluntary compliance. Res. Exp. Econ. 14, 77–114.
Chou, Eileen, McConnell, Margaret, Nagel, Rosemarie, Plott, Charles R., 2009. The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person guessing game. Exp. Econ. 12 (2), 159–179.
Engel, Christoph, Rand, David G., 2014. What does "clean" really mean? The implicit framing of decontextualized experiments. Econ. Lett. 122 (3), 386–389.
Cooper, David J., Kagel, John H., 2003. The impact of meaningful context on strategic play in signaling games. J. Econ. Behav. Organ. 50 (3), 311–337.
Cooper, David J., Kagel, John H., 2009. The role of context and team play in cross-Game learning. J. Eur. Econ. Assoc. 7 (5), 1101–1139.
Cooper, David J., Kagel, John H., Lo, Wei, Gu, Qing Liang, 1999. Gaming against managers in incentive systems: experimental results with Chinese students and Chinese managers. Am. Econ. Rev. 89 (4), 781–804.
Cox, James C., Oaxaca, Ronald L., 1989. Laboratory experiments with a finite-horizon job-search model. J. Risk Uncertain. 2 (3), 301–329.
Croson, Rachel, Gneezy, Uri, 2009. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.
Dohmen, Thomas, Falk, Armin, Huffman, David, Sunde, Uwe, 2009. Homo reciprocans: survey evidence on behavioural outcomes. Econ. J. 119 (536), 592–612.
Dufwenberg, Martin, Gächter, Simon, Hennig-Schmidt, Heike, 2011. The framing of games and the psychology of play. Games Econ. Behav. 73 (2), 459–478.
Ellingsen, Tore, Johannesson, Magnus, Möllerstrom, Johanna, Munkhammar, Sara, 2012. Social framing effects: preferences or beliefs? Games Econ. Behav. 76 (1), 117–130.
Ellingsen, Tore, Dreber, Anna, Johannesson, Magnus, Rand, David, 2013. Social framing effects: preferences or beliefs? Exp. Econ. 16, 349–371.
Elliott, Catherine S., Hayward, Donald M., Canon, Sebastian, 1998. Institutional framing: some experimental evidence. J. Econ. Behav. Organ. 35 (4), 455–464.
Engel, Christoph, Goerg, Sebastian J., Yu, Gaoneng, 2013. Symmetric Vs. Asymmetric Punishment Regimes for Bribery, MPI Collective Goods Preprint 2012/1. Max Planck Institute.
Eriksson, Kimmo, Strimling, Pontus, 2014. Spontaneous associations and label framing have similar effects in the public goods game. Judgm. Decis. Mak. 9 (5), 360–372.
Gibson, Rajna, Tanner, Carmen, Wagner, Alexander F., 2013. Preferences for truthfulness: heterogeneity among and within individuals. Am. Econ. Rev. 103 (1), 532–548.
Gilboa, Itzhak, Schmeidler, David, 1995. Case-based decision theory. Q. J. Econ. 110 (3), 605–639.
Griggs, Richard A., Cox, James R., 1982. The elusive thematic-materials effect in Wason's selection task. Br. J. Psychol. 73 (3), 407–420.
Harbring, Christine, Irlenbusch, Bernd, 2011. Sabotage in tournaments: evidence from a laboratory experiment. Manage. Sci. 57 (4), 611–627.
Harrison, Glenn, List, John, 2004. Field experiments. J. Econ. Lit. 42 (4), 1009–1055.
Hennig-Schmidt, Heike, Rockenbach, Bettina, Sadrieh, Abdolkarim, 2010. In search of workers' real effort reciprocity – a field and a laboratory experiment. J. Eur. Econ. Assoc. 8 (4), 817–837.
Jehiel, Philippe, 2005. Analogy-based expectation equilibrium. J. Econ. Theory 123 (2), 81–104.
Kay, Aaron C., Ross, Lee, 2003. The perceptual push: the interplay of implicit cues and explicit situational construals on behavioral intentions in the Prisoners Dilemma. J. Exp. Soc. Psychol. 39 (6), 634–643.
Kirchler, Michael, Huber, Jürgen, Stöckl, Thomas, 2012. Thar she bursts: reducing confusion reduces bubbles. Am. Econ. Rev. 102 (2), 865–883.
Krajcova, Jana and Andreas Ortmann, 2008. Testing Leniency Programs Experimentally: The Impact of 'Natural' Framing, Working Paper 372, CERGE-EI 2008.
Lambsdorff, Johann Graf, Frank, Björn, 2011. Corrupt reciprocity – experimental evidence on a men's game. Int. Rev. Law Econ. 31 (2), 116–125.

Landeo, Claudia M., Nikitin, Maxim, Babcock, Linda, 2007. Split-awards and disputes: an experimental study of a strategic model of litigation. J. Econ. Behav. Organ. 63 (3), 553–572.

Li, Danyang, Hawley, Zackary, Schnier, Kurt, 2013. Increasing organ donation via changes in the default choice or allocation rule. J. Health Econ. 32 (6), 1117–1129.

Liberman, Varda, Samuels, Steven M., Ross, Lee, 2004. The name of the game: predictive power of reputations versus situational labels in determining prisoners dilemma game moves. Pers. Soc. Psychol. Bull. 30 (9), 1175–1185.

Manktelow, Ken, Evans, J. St B.T., 1979. Facilitation of reasoning by realism: effect or non-effect? Br. J. Psychol. 70 (4), 477–488.

Nosenzo, Daniele, 2013. Pay secrecy and effort provision. Econ. Inq. 51 (3), 1779–1794.

Pevnitskaya, Svetlana, Ryvkin, Dmitry, 2013. Environmental context and termination uncertainty in games with a dynamic public bad. Environ. Dev. Econ. 18 (1), 27–49.

Rivas, M. Fernanda, 2013. An experiment on corruption and gender. Bull. Econ. Res. 65 (1), 10–42.

Samuelson, Charles D., Allison, Scott T., 1994. Cognitive factors affecting the use of social decision heuristics in resource-sharing tasks. Organ. Behav. Hum. Decis. Process. 58 (1), 1–27.

Smith, Vernon L., Suchanek, Gary, Arlington, W., 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica 56 (5), 1119–1151.

Smith, Vernon L., 1976. Experimental economics: induced value theory. Am. Econ. Rev. 66 (2), 274–279.

Thöni, Christian, Simon Gächter, 2014. Peer effects and social preferences in voluntary cooperation, Working Paper 4741, CESifo.

Wason, Peter C., Shapiro, Diana, 1971. Natural and contrived experience in a reasoning problem. Q. J. Exp. Psychol. 23 (1), 63–71.

Yamagishi, Toshio, Nobuhiro, Mifune, Li, Yang, Mizuho, Shinada, Hirofumi, Hashimoto, Yutaka, Horita, Arisa, Miura, Keigo, Inukai, Shigehito, Tanida, Toko, Kiyonari, et al., 2013. Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. Organiz. Behavi. Hum. Decis. Process. 120 (2), 260–271.