

Clarifying The Self-Serving Bias

Todd R. Kaplan and Bradley J. Ruffle

Ben-Gurion University

Beer Sheva, Israel

In their article “Explaining Bargaining Impasse: The Role of Self-Serving Biases”, (*JEP*, Winter 1997), Babcock and Loewenstein (hereafter BL) invoke the notion of self-serving biases to account for behavior in a rich variety of settings, for instance, the failure to reach an agreement in laboratory bargaining games, the behavior of job searchers, and the discrepancy between plaintiffs’ and defendants’ assessments of tort cases.

This comment deals with three aspects of BL’s paper. First, we offer a more general definition of the self-serving bias. Second, we show how much of their evidence of a bias from previous bargaining experiments is open to alternative interpretations. Third, we debate their methodological choice of conducting contextually rich experiments. We conclude by introducing our own experiment which tests for the presence of a self-serving bias in a context-free environment.

A self-serving bias exists where an individual’s preferences affect his beliefs in an optimistic direction, one favoring his own payoff. Beliefs may be about one’s own ability, the environment, another player’s type or about what is a fair outcome. Note that this definition is more general than the one implied by BL. They refer to the self-serving bias as a tendency “to conflate [blend together] what is fair with what benefits oneself.” They go on to explain that “people tend to arrive at judgments of what is fair or right that are biased in the direction of their own self-interests”. We feel that this definition is too restrictive. It excludes many of their own examples apparently explained by the self-serving bias. The “above average” effects (p.111), the low take-up rate of unemployment insurance (p.120), and many of the discrepancies between judges’ and lawyers’ survey responses (p.121) are all unrelated to fairness. These examples appear misplaced given BL’s definition. Their inclusion suggests to us that BL intended a more general definition of the self-serving bias. Our expanded definition serves to capture the above phenomena.

Of the examples which do fit BL's definition, the interpretation of the ones reported in the section "Reinterpreting Findings from Previous Bargaining Experiments" is flawed. The bargaining impasse they try to explain with the self-serving bias can equally be explained with subjects *ex ante* having different notions of fairness.

Consider the Roth and Murnighan (1982) (RM) study (p.117) where two subjects bargain over the division of 100 lottery tickets. The subject holding the winning lottery ticket receives either \$5 or \$20. RM report that 12 (22) percent of the subjects failed to reach an agreement when neither (both) subject(s) knew who would receive which payoff. BL view this difference of the non-agreement rates as evidence of the self-serving bias: "both sides viewed as fair the focal settlement that benefitted themselves, so the \$20-prize player was likely to hold out for half of the chips, while the \$5-prize player demanded equal expected values." In fact, there exists an alternative explanation for RM's finding. When subjects bargain before knowing who receives which payoff, an equal division of the tickets is the unique focal outcome.¹ When they bargain after knowing who receives which payoff, equal expected payoffs arises as a second focal division. Let us illustrate that disagreement must necessarily increase as the number of focal points increases from one to two. Suppose a fraction x of the total population hold *ex ante* beliefs that 50/50 (equal chance of winning) is a fair split of the tickets, while $1 - x$ believes 20/80 (equal expected payoffs) is fair. In the first treatment, subjects are unable to condition their offers on who is assigned to which prize. And so the 20/80 types cannot express their true beliefs about a fair outcome, but instead are likely to agree to the only focal point, an equal split of the tickets. The second treatment increases the probability of disagreement because the 20/80 types can now choose their preferred division. Where subjects are randomly paired, the probability of conflict increases by $x(1 - x)$. This term has a maximum value of 0.25. A potential 25 percent rise in conflict

¹Schelling (1960) introduced the term "focal point" to suggest which equilibria might actually be reached in coordination games where multiple Nash equilibria exist. We, along with others in the experimental bargaining game literature, must therefore apologize for our (ab)use of the term to refer to outcomes in games of conflict which appear more salient to players.

is significantly larger than the 10 percent rise that BL cite as evidence of the self-serving bias!

A similar argument can be made concerning BL's interpretation of Knez and Camerer (KC) (1995). KC conduct ultimatum game experiments in which players have an outside option in the case of rejection. For instance, if \$10 is to be divided, and the proposer earns \$4 and the responder earns \$3 if the responder rejects the proposer's division, then there are two focal Nash equilibria: an equal division of the \$10 (\$5, \$5), or a equal division of the surplus above the outside offers, (\$5.50, \$4.50). BL (and KC themselves, p.80) conclude from observed rejection rates of nearly half of the offers – much higher than in ordinary ultimatum studies – that a self-serving bias is at work. The problem again is that they are comparing rejection rates in a game with two focal divisions to those in a game with only one. For a true test of the self-serving bias in KC's experiment, we suggest the following control treatment: conduct a treatment with outside options identical to KC's, except that each subject indicates both an offer and the minimum offer he is willing to accept (WTA). Only *after* a subject makes his decisions does he learn his randomly determined role. Having controlled for the number of focal divisions, a self-serving bias can be said to exist if offers or WTA in KC's treatment are greater than those in the control treatment.

The third point we would like to address is BL's critique of the notion of a context-free experiment. Their tort case studies, for instance, were designed "to test for the effect of the self-serving bias in a contextually rich and controlled experimental setting." (p.112) They think "the emphasis among economists on expunging context in experiments is a mistake." (p.122) This generalization strikes us as too strong. It runs counter to the most basic principles of a controlled experiment. The beauty of economics experiments is not their ability to mimic the real world, but their ability to control for it. The interpretation of real-world phenomena is confounded by a host of explanatory variables and unobservables, including individuals' connotations about the environment in which they act. The laboratory offers the opportunity to control for, or at least minimize, subjects' subjective and unobservable beliefs about the institution in which they operate.

While no experiment can ever fully expunge context, we must still strive to minimize it.

Our purpose in this note is not to say that the self-serving bias does not exist – both authors agree that it most likely does. In fact, our own more general definition of the self-serving bias suggests scope for the bias to operate in additional settings, settings void of fairness issues. Instead we find much of the evidence thus far offered in support of a self-serving bias unconvincing. These points led us to design our own test of the self-serving bias, one which is (nearly) context-free and unrelated to fairness. It is a variation of a game developed by Moulin (1986). In the guessing game, players simultaneously choose a number in the closed interval $[0, 100]$. The player whose number is closest to $p \times$ mean of all numbers chosen (where p is a parameter common knowledge to all subjects) wins a fixed amount. All other players earn zero.

The unique Nash equilibrium of this game, for $p \in [0, 1)$, is for all players to choose zero. Nagel (1995) has tested this game experimentally to investigate players' depth of reasoning. A player will choose a number greater than zero if he is irrational (zero-order beliefs), if he is rational but believes others are irrational (first-order beliefs), or, more generally, if at some level in his infinite hierarchy of beliefs, he specifies some irrationality.

Our version of the guessing game (Kaplan and Ruffle (1998)) introduces a second, additional payoff. There are 30 subjects in each session, each subject with an identity number from 1-30. Those subjects with an odd identity number receive the mean guess of all 29 other chosen numbers. Similarly, all even-numbered subjects receive 100 minus the mean guess of all 29 other subjects. The self-serving bias predicts that the guesses of the odd-numbered subjects will be greater than those with even identity numbers. After all, an odd-numbered subject should make the self-serving guess others are less rational, and therefore guess higher, than is supposed by an even-numbered subject whose lower guesses reflect their relatively optimistic self-serving beliefs about others' rationality. The existence of a self-serving bias in such a benign context in which strategic and altruistic behavior are controlled for would offer strong evidence of its pervasiveness.

References

Babcock, Linda and George Loewenstein (Winter 1997) “Explaining Bargaining Impasse: The Role of Self-Serving Biases”, *Journal of Economic Perspectives*, 11:1, 109-126.

Kaplan, Todd and Bradley Ruffle (1998) “The Self-Serving Bias and Beliefs about Rationality”, in progress, Ben-Gurion University.

Knez, Marc and Colin Camerer (1995) “Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments”, *Games and Economic Behavior*, 10, 65-94.

Moulin, Hervé, *Game Theory for the Social Sciences*, New York Press: New York, 1986.

Nagel, Rosemarie (December 1995) “Unraveling in Guessing Games: An Experimental Study”, *American Economic Review*, 85:5, 1313-1326.

Roth, Alvin E. and J. Keith Murnighan (September 1982) “The Role of Information in Bargaining: An Experimental Study”, *Econometrica*, 50, 1123-1142.

Schelling, Thomas, *Strategy of Conflict*, Harvard University Press: Cambridge, 1960.