

# **The Endowment Effect and the Willingness to Accept-Willingness to Pay Gap: Subject Misconceptions or Reference Dependence?**

Weining Koh and Wei-Kang Wong<sup>1</sup>

## Abstract

Plott and Zeiler (2005) found that the WTA-WTP gap disappeared when they adopted all procedural controls for misconceptions. Consequently, they attributed the gap to subject misconceptions. We replicated their procedures and findings without any lottery rounds. But we found that increasing sample size improved the gap's statistical significance. We then modified subtle details of their procedures to strengthen the reference states, keeping all key controls for misconceptions. These changes produced a moderate but highly statistically significant gap. Thus, experimental procedures might have turned reference dependence and hence the gap on and off by inducing strong or weak reference states.

Keywords: Endowment Effect, WTA-WTP Gap, Prospect Theory, Reference State, Subject Misconceptions

JEL Classification: C91, D46

---

<sup>1</sup> Koh and Wong (Corresponding Author): Department of Economics, National University of Singapore, AS2, 1 Arts Link, Singapore 117570, Republic of Singapore. Phone: (65) 65166016. Fax: (65) 67752646. Email: [ecswong@nus.edu.sg](mailto:ecswong@nus.edu.sg). We especially thank Jack Knetsch for very valuable comments. We also thank Stefano DellaVigna and the anonymous referees for valuable comments, and Brigitte Fang, Ting Zeng, Cliff Kuo for research assistance. The usual disclaimer applies. We gratefully acknowledge the support of the Staff Research Support Scheme of the NUS Faculty of Arts and Social Sciences.

## **Introduction**

Endowment effect or the increased value of a good to an individual when the good becomes part of the individual's endowment is often thought to be a manifestation of loss aversion or reference dependence – changes in the domain of losses short of the reference state are valued more than commensurate changes in the domain of gains beyond the reference state (Kahneman and Tversky, 1979; Thaler, 1980). Consequently, people commonly ask for more compensation to give up a good than they are willing to bid to acquire it, leading to an observed gap between willingness to accept (WTA) and willingness to pay (WTP) (for example, documented by Kahneman, Knetsch, and Thaler, 1990; and reviewed in, Samuelson and Zeckhauser, 1988; Rabin, 1998; Horowitz and McConnell, 2002; DellaVigna, 2009). A key, but largely implicit, pre-requisite for finding a WTA-WTP gap is that sellers must perceive “having the good” as their reference state and evaluate “giving it up” as a loss, whereas buyers must perceive “not having the good” as their reference state so that they evaluate “acquiring the good” as a gain.

Some researchers have questioned whether these valuation disparities are really due to asymmetry in preferences over gains and losses. Most notably, Plott and Zeiler (2005), or PZ henceforth, ask: if they adopt the union of all previous experimental procedures that aim to control for subject misconceptions – including ensuring subject anonymity, using an incentive-compatible elicitation mechanism, carefully explaining to subjects why truthful revelation of personal values is the optimal response under the mechanism, and providing subjects with practice and training on the elicitation mechanism – will they still observe a WTA-WTP gap? It turned out that they observed no gap in the WTA and WTP elicited for mugs even without any paid practice lottery rounds, as long as the other controls were present. Consequently, they interpret this finding as evidence that the gap is due to some subject misconceptions rather than reference dependence.

We investigate whether subtle details of PZ's procedures – such as buyers having possession of the good at the time of decision and having to give it up if they fail to buy, not always framing buy/sell in gain/loss terms, and asking subjects to answer a series of yes/no questions to plausible values for the good being valued during training – may have weakened the reference states of buyers and sellers, thereby making gains and losses less salient, and it is weak reference states rather than the elimination of subject misconceptions that have caused the WTA-WTP gap to

disappear. If the gap has disappeared because of weak reference states, then we should observe a gap if we modify their procedures to strengthen the reference states while keeping all key controls for subject misconceptions. In contrast, if the gap is due to subject misconceptions, then reference states are irrelevant and we should continue to observe no gap despite procedural modifications to strengthen the reference states.

## **Experimental Design and Results**

Twelve experimental sessions were carried out at the National University of Singapore with 345 students recruited from an advertisement posted on the university's Virtual Learning Environment, a website that publishes all university course materials.<sup>2</sup> Sessions 1-4 were run in 2009 whereas sessions 5-12 were run in 2011. Table 1 summarizes the major characteristics of each session and the results.<sup>3</sup> The advertisement announced that:

“We are looking for participants for an experiment in individual decision-making. The participants will be asked to perform some valuation tasks, which will take about 30 minutes to complete. The participants are not personally identified, but are instead given ID; the responses made by participants will be anonymous” and “The subjects will be reimbursed \$10 per hour of participation in the research.”<sup>4</sup>

All payments and transactions were in Singapore dollars.<sup>5</sup> All sessions were completed in about half an hour and every subject was eventually paid \$5 for taking part in the experiment. Subjects were asked to state their offers: maximum WTP for buyers and minimum WTA for sellers, for a coffee mug with the university insignia imprinted on it that was specially ordered at a cost of \$3.75 and not otherwise available. Subjects were not informed about the cost or the special order. All

---

<sup>2</sup> Every university student could view and respond to the advertisement. Results from an earlier, and unrelated, series of experiments carried out in Canada and in Singapore showed very similar preferences and choices, suggesting little reason for any lack of generality in the present findings (Knetsch et al., 2001).

<sup>3</sup> To ensure consistency across different sessions over time, the same experimenter had conducted all sessions. Because sessions run at different times of day may attract different types of subjects, to prevent selection bias, the experimenter had scheduled all sessions during 2-4pm on weekdays except one session that occurred an hour earlier during 1-2pm. See Table 2A for the date and time of each session.

<sup>4</sup> To avoid any confusion about the reimbursement, when the participants were selected and notified, they were explicitly told that since the experiment was expected to last for 30 minutes, the participation fee was expected to amount to \$5.

<sup>5</sup> The Singapore dollar was worth about USD 0.70.

transactions were real and conducted with an incentive-compatible Becker-DeGroot-Marschak's (1964) random price auction. If a buyer's offer was more than or the same as a randomly determined fixed offer then the buyer bought the mug and paid the fixed offer. If a seller's offer was less than or the same as a randomly determined fixed offer then the seller sold the mug and received the fixed offer. Sellers who did not sell at the fixed offer took home the mugs they were endowed with. To generate a fixed offer, instead of using a random number generator like PZ (which subjects might be unfamiliar with), we simply told the subjects that:

“Your offer will be compared to a fixed offer. The fixed offer was determined in the following way:

We prepared 21 pieces of identical papers, with 0, 0.5, 1, 1.5, 2, 2.5, ..., 8.5, 9, 9.5, 10 written on them. We folded the papers and dropped them into a bag. Just before the session, we randomly drew a number from the bag and sealed the number into the envelope you see here. We redraw the number for every session. This randomly drawn number in the envelope will be the fixed offer.

As you can see, the fixed offer will be completely unrelated to your offer and to the offers of all other persons in the room.”

The fixed offer was revealed after all subjects had stated their offers on their record sheets. Based on their offers and the fixed offer, they had to circle the transaction outcome: BUY or NO BUY for buyers and SELL or NO SELL for sellers, on both their record sheets and a separate sheet on their written instructions. They would then drop their record sheets (containing their offers, the fixed offer, and their transaction outcomes) into a box located in front of the room. Finally, they would take the written instructions containing their transaction outcomes (but not their offers) to the cashiers located outside the room, who would execute the transactions accordingly, in addition to paying them their participation fees.<sup>6</sup> To ensure anonymity, subjects were not personally identified but were randomly assigned an identification number. Thus, the subjects' offers would be known only to themselves. We explained these procedures to the subjects before they stated their offers.

All sessions maintained the key procedural controls that PZ deemed necessary to prevent subject misconceptions, which included ensuring subject anonymity, using an incentive-compatible elicitation mechanism, explaining to subjects how they would

---

<sup>6</sup> Before the subjects left the room, the experimenters informed the cashiers located outside the room what the fixed offer was.

work out their maximum WTP and minimum WTA and why truthful revelation was the optimal response using illustrative numerical examples, but excluded the lottery rounds.

We excluded paid lottery rounds because in PZ's experiment 2, the mug round came before paid lottery rounds, yet they observed no WTA-WTP gap in the mug round – their result 2. They emphasised that “the main point, however, is that paid practice rounds seem unnecessary in the presence of other procedures thought to control for subject misconceptions” (p.541) and that “experience with the lotteries could not have played a crucial role in the disappearance of the gap” (p.543). We excluded unpaid lottery rounds because Isoni, Loomes, and Sugden (2011), or ILS henceforth, showed that despite using unpaid lottery rounds and the same rigorous experimental controls that PZ used for the mug round, the WTA-WTP gap remained throughout all fourteen paid lottery rounds and showed no tendency to decline in their replication of PZ. As we shall see, since we also find no significant WTA-WTP gap in our replication of PZ without any lottery rounds for sample size comparable to PZ, this suggests that lottery rounds are not essential for PZ's results, as PZ themselves argued.

Because each of our experimental sessions consisted of only buyers or only sellers, buyers only received instructions for the buying task whereas sellers only received instructions for the selling task. In addition to simplifying the experimental instructions and illustrations, this between-subject design is desirable because switching between the roles of buyer and seller in a within-subject design may induce a mental frame and reference state of trading, instead of reference states of endowment versus no endowment; Brown (2005) shows that switching between the roles of buyer and seller triggered a buy-low sell-high good-deal seeking mentality that was unrelated to personal valuation and this misconception was prevalent despite the adoption of an incentive-compatible BDM mechanism. Another advantage of having only buyers or only sellers in each session is that it eliminates potential misconceptions arising from comparison between buyers and sellers.

### The Manipulation Treatments

Our manipulation treatments (referred to as KW) made three small modifications to PZ's procedures while maintaining all key controls that PZ deemed necessary to prevent misconceptions on the part of participants. The three

modifications were: first, strengthening buyers' reference state by not giving buyers possession of the mugs at the time of decision so that they evaluate a purchase as a gain rather than as an avoidance of a loss;<sup>7</sup> second, strengthening sellers' reference state by framing a sale as a relinquishment of their possessions for money or a loss; and third, removing possible suggestion of value and anchoring effect in the numerical examples given. We discuss each modification in turn.

First, in PZ, "sellers were told that they owned the mug" whereas "buyers were told that they could inspect the mug but they did not own it" (p.539), but both buyers and sellers had possession of the mugs at the time of decision. Having possession regardless of ownership weakens the reference states because it creates conflicting reference states and evaluations for buyers. On the one hand, because buyers did not own the mug, they should evaluate the mug as a gain when they acquired it. On the other hand, because buyers had possession of the mug, they could evaluate the mug as a loss when they failed to acquire it because they had to literally give it up; so loss aversion would induce higher valuation from buyers if they wanted to avoid parting with the mugs. Having possession may be a more salient status quo when ownership is given instantaneously and arbitrarily by the experimenters; buyers may feel no less entitled to owning the mugs than sellers.<sup>8</sup>

To strengthen the reference states and to avoid any ambiguity in the framing of transactions, only sellers had possession of the mugs at the time of decision in our manipulation treatments; buyers passed around the mugs for inspection and the mugs were removed from them before they were asked to value the mugs. We paid special attention to make sure that buyers were given ample time to inspect the mugs until they were fully satisfied with the inspection. Like PZ, sellers were told that they could inspect the mug and they now owned the mug in their possession, whereas buyers were told that they could inspect the mug but they did not own it. In addition, to clarify what an instantaneously created ownership in the laboratory meant to sellers and to make sure that they understood that ownership was real, sellers were told that "the mug you are inspecting is yours to take home with you at the end of the experiment, if you decide to keep it later."

---

<sup>7</sup> "Possession" means "actual holding or occupancy with or without rightful ownership."

<sup>8</sup> In other words, if sellers feel a much greater sense of entitlement to the good than buyers, then possession or non-possession may be less important to the establishment of the reference states.

Second, PZ referred to “the item” or “the mug” for both buyers and sellers. Moreover, they sometimes did not frame a sale as a relinquishment of a seller’s possession for money. To strengthen the reference state of possession for sellers and to induce a gain/loss mental frame, we referred to “the item” or “the mug” for buyers, but replaced the neutral determiner “the” with a possessive determiner such as “my” or “your” for sellers. We also consistently framed a sale as a transaction where the sellers have to give up their items or their mugs for money. For example, we told buyers “to determine the maximum you would be willing to pay for the mug”. But instead of telling sellers “to determine the minimum you would be willing to accept for the mug”, we told sellers “to determine the minimum you would be willing to accept to give up your mug.”

Third, to illustrate how subjects would work out the minimum WTA and maximum WTP, PZ’s instructions asked subjects to answer a series of exploratory Yes/No questions on WTA and WTP in response to specific dollar figures that were between 0 and 10 (i.e., plausible values for a coffee mug), finally arriving at a minimum WTA of \$6.50 and a maximum WTP of \$5.25 in the examples. To remove possible anchoring effect or suggestion of value, we replaced these dollar figures with numerical values that were clearly implausible for a coffee mug (by using constant multiples, typically 100, of the original values that were between 0 and 10) and we emphasised that these examples were purely hypothetical.<sup>9</sup>

## The Results

Sessions 1 & 5 and 2 & 6 replicated the verbal and written instructions of PZ for buyers and sellers, respectively.<sup>10</sup> Similarly, sessions 3 & 7 and 4 & 8 implemented the manipulation treatments of KW for buyers and sellers, respectively. In what follows, we first report the results for 2009 (sessions 1-4) and 2011 (sessions

---

<sup>9</sup> For instance, PZ gave the following numerical example to illustrate how one would work out the minimum one is willing to accept [with our modifications in square brackets]:

Would I accept \$10 [\$1000] to give up Item B [my item]? Yes.

Would I accept \$8 [\$800] for B [to give up my item]? Yes.

Would I accept \$7 [\$700] for B [to give up my item]? Yes.

Would I accept \$6 [\$600] for B [to give up my item]? No, not \$6 [\$600]. So I need to increase.

Would I accept \$6.50 [\$650]? I don’t care whether I end up with \$6.50 [\$650] or Item B [keeping my item]. Then that is the minimum I’d be willing to accept for Item B [to give up my item].

<sup>10</sup> We adopted a few very minor modifications suggested in ILS to make the instructions more consistent and complete. These minor changes could not have accounted for any change in the experimental results, as they did not change the experimental results in ILS’s replication of PZ in the mug experiment.

5-8) separately before pooling them. Table 2A reports the date and time of each session, individual subject data and their summary statistics.

In our replication of PZ in 2009, we find a median WTP of \$2.5 (mean of \$2.57) and a median WTA of \$3 (mean of \$4.1), leading to a median WTA-WTP ratio of 1.2 (mean of 1.60).<sup>11</sup> In 2011, we find a median WTP of \$1.8 (mean of \$2.75) and a median WTA of \$3.5 (mean of \$3.61), leading to a median WTA-WTP ratio of 1.94 (mean of 1.31).

In our manipulation treatments in 2009, we find a median WTP of \$2 (mean of \$2.56) and a median WTA of \$4 (mean of \$3.8), leading to a median WTA-WTP ratio of 2 (mean of 1.48). In 2011, we find a median WTP of \$2.5 (mean of \$2.50) and a median WTA of \$4.25 (mean of \$4.54), leading to a median WTA-WTP ratio of 1.7 (mean of 1.82).

Table 2B reports the summary statistics based on pooled data from 2009 and 2011. In our replication of PZ, we find a median WTP of \$2.5 (mean of \$2.66) and a median WTA of \$3.25 (mean of \$3.85), leading to a median WTA-WTP ratio of 1.3 (mean of 1.45). In contrast, in our manipulation treatments, we find a median WTP of \$2 (mean of \$2.53) and a median WTA of \$4.25 (mean of \$4.16), leading to a median WTA-WTP ratio of 2.13 (mean of 1.64).

Table 3 reports the results of statistical tests to determine whether the data support the hypothesis that WTP is different from WTA. Panel A reports the results using all values elicited. We perform the Wilcoxon-Mann-Whitney test, which tests for whether the WTP and WTA samples were drawn from identical distributions, and median test, which tests for whether the WTP and WTA samples were drawn from distributions with identical medians.<sup>12</sup>

The results show that when we consider the 2009 and 2011 sessions separately, we cannot reject the null hypotheses of identical distributions or identical medians in our replication of PZ at the conventional significance levels. In other words, when PZ's procedures were used, there was no evidence that WTP was significantly different from WTA even without the lottery rounds. Specifically, using WTP and WTA from 2009, the Wilcoxon-Mann-Whitney rank sum test results in a  $z$

---

<sup>11</sup> If we exclude two potential outliers for WTA in session 2 that exceeded \$10, then our replication of PZ in 2009 yields a median WTA of \$3 as before (but a lower mean of \$3.39), leading to a median WTA-WTP gap of 1.2 as before (but a lower mean of 1.32). None of the offers in the remaining sessions exceeded \$10.

<sup>12</sup> Following PZ, we did not perform tests for the equality of means; means could be more sensitive to outliers.

value of -1.484 ( $p=0.1378$ ) and the median test results in a Pearson  $X^2$  statistic of 0.5880 ( $p=0.443$ ).<sup>13</sup> Similarly, using the values from 2011, the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -1.411 ( $p=0.1583$ ) and the median test results in a Pearson  $X^2$  statistic of 0.1528 ( $p=0.696$ ).

In contrast, when PZ's procedures were modified to strengthen the reference states, there was strong evidence that WTP was different from WTA. Even when we consider the 2009 and 2011 sessions separately, we can always reject the null hypotheses of identical distributions or identical medians in our manipulation treatments at the five percent level. Specifically, using WTP and WTA from 2009, the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -2.495 ( $p=0.0126$ ) and the median test results in a Pearson  $X^2$  statistic of 8.2127 ( $p=0.004$ ). Similarly, using the values from 2011, the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -2.832 ( $p=0.0046$ ) and the median test results in a Pearson  $X^2$  statistic of 6.0146 ( $p=0.014$ ).

Because the sessions in 2009 and 2011 yield similar results, it is natural to pool them. Not surprisingly, with pooled data, there was even stronger evidence that WTP was different from WTA in our manipulation treatments; the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -3.654 ( $p=0.0003$ ) and the median test results in a Pearson  $X^2$  statistic of 11.4718 ( $p=0.001$ ). We can reject the null hypotheses of identical distributions and identical medians in our manipulation treatments at the one percent level. But somewhat surprisingly, with pooled data, it turns out that we can also reject the null hypothesis of identical distributions for WTP and WTA in our replication of PZ at the five percent level, although we still cannot reject the null hypothesis of identical medians at the conventional levels; the Wilcoxon-Mann-Whitney rank sum test now results in a  $z$  value of -2.102 ( $p=0.0356$ ), whereas the median test results in a Pearson  $X^2$  statistic of 0.9921 ( $p=0.319$ ).<sup>14</sup>

Panel B of Table 3 reports the results when potential outliers – offers exceeding \$10 in session 2 – are excluded. When we exclude potential outliers from

---

<sup>13</sup> Based on the individual data in Table 2A, it is natural to ask to what extent including potential outliers with WTA exceeding \$10 in session 2 might have affected our conclusions. Panel B of Table 3 reports the results when these potential outliers are excluded. It is clear that excluding these high offers would only make the WTA-WTP gap even less significant in our replication of PZ in 2009 and we still would not be able to reject the null hypotheses of identical populations or identical medians for WTP and WTA; specifically, the Wilcoxon-Mann-Whitney rank sum test now results in a  $z$  value of -1.080 ( $p=0.2803$ ), whereas a median test results in a Pearson  $X^2$  statistic of 0.2414 ( $p=0.623$ ).

<sup>14</sup> Nevertheless, it is worth noting that the rank sum test is generally thought to be more powerful than the median test.

pooled data in our replication of PZ, the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -1.831 ( $p=0.0671$ ), whereas the median test results in a Pearson  $X^2$  statistic of 1.0129 ( $p=0.314$ ). Increasing sample size still improves statistical significance in our replication of PZ, such that the  $z$  value from rank sum test is now statistically significant at the ten percent level.

In summary, by pooling the data from 2009 and 2011, our replication of PZ now yields some evidence that WTP was different from WTA, even though we found no statistically significant difference when we considered them separately. This might seem surprising, especially since we adopted all key procedures PZ deemed necessary to control for subject misconceptions. Nevertheless, the hypothesis of weak reference states would in fact predict it. Weak reference states would cause statistical tests to have low power, i.e., low probability of rejecting the null hypothesis of no endowment effect when there is in fact an endowment effect. Increasing sample size is expected to increase the power of tests.

It is worth emphasising that when considered separately, our 2009 and 2011 experiments are comparable in sample size to the original experiments in PZ and like them, we also cannot reject the null hypothesis of no endowment effect at the conventional levels.<sup>15</sup> So we did replicate PZ's results without the lottery rounds.<sup>16</sup> Thus, it seems natural that the evidence for endowment effect that emerged in the larger pooled sample should be interpreted as revealing the alternative hypothesis of weak reference states.

## Discussion

For sample size comparable to PZ, we found statistically insignificant valuation disparities between buyers and sellers in our replication of PZ without any practice lottery rounds, but highly statistically significant disparities when we modified subtle details of their procedures to strengthen the reference states, keeping all key procedural controls for subject misconceptions. These modifications are not

---

<sup>15</sup> There are two treatments to control for subject misconceptions in PZ (2005). PZ's experiment 2, where the mug round came before paid lottery rounds, consisted of 26 subjects with 12 buyers and 14 sellers, respectively. The other treatment in PZ, where the mug round came after paid lottery rounds, consisted of two sessions conducted at two different universities, and a total of 48 subjects with 24 buyers and 24 sellers, respectively. In other words, that made a total mixed subject pool of (26+48=) 74 across two different treatments, with a total of 36 buyers and 38 sellers, respectively.

<sup>16</sup> Nevertheless, we should point out that unlike us, PZ had found values for WTP and WTA that were much closer, sometimes with WTP exceeding WTA. But it is difficult to interpret these results with confidence because of the small sample size in these sessions.

expected to reduce the extent of control for misconceptions. They are also exempt from classical incentives that Plott and Zeiler (2007) argue can induce asymmetric choices in an exchange experiment where subjects are endowed with one good but given the opportunity to trade for an alternative good, such as different transaction costs for traders and non-traders, possible signals of relative value of the endowed good versus the alternative good, cascade effects due to public revelation of choices, as well as non-classical behavioural motives such as subjects' expectations of experimenter's altruism in the choice of endowment; in all of our treatments, both buyers and sellers incurred basically the same transaction costs, valued the same good privately and anonymously after being given ample time to inspect the good being valued.<sup>17</sup> Thus, the results suggest that although different experimental procedures

---

<sup>17</sup> In an exchange experiment where two goods (the endowed good and the alternative) are involved, Plott and Zeiler (2007) propose a few hypotheses on how physical possession of one good but not the other may introduce some classical incentives that promote valuation disparities for reasons unrelated to reference dependence. These arguments seem much less plausible in our WTA-WTP experiments where only one good was involved and buyers were separated from sellers. Nevertheless, we briefly discuss them in turn.

First, could physical possession change the transaction cost of buying versus not buying? In our experimental setup, there is no meaningful difference in the marginal transaction cost of buying versus not buying whether there is physical possession or not. Furthermore, Plott and Zeiler (2007) have ruled out this explanation themselves.

Next, could giving physical possession to sellers but not buyers be construed as a signal of either the experimenters' intention for sellers to keep the mugs or the mugs' value because of their proximity? A variant of this argument appears under the enhancement effect theory in PZ (2011). Generally, the enhancement effect theory holds that experimental procedures may add features to the goods being valued and these features may have independent values, e.g., if the experimenter told the subjects that the endowment was a gift. Specifically, in the current context, PZ (2011) argue that the proximity of mug may have created additional value for reasons unrelated to loss aversion.

It is worth emphasising that for every enhancement effect proposed, there is an equally plausible effect working in the opposite direction. For example, by giving buyer possession of the mug he does not own, could this be viewed as a subtle hint by the experimenter that the buyer should keep it?

Furthermore, enhancement effect is more likely to be invoked by experimental procedures when the treatment is unexpected, but not when the treatment is the expected, because it is the unexpected that registers in the mind of subjects, consciously or subconsciously. If so, possession without ownership could create an enhancement effect for buyers in PZ's treatment. In other words, upon close examination, the hypothesis of enhancement effect in fact does not have a clear prediction in this context and could very well work in the opposite direction, as opposed to what PZ hypothesised.

Furthermore, we think that PZ's (2007) arguments are untenable for a few other reasons.

First, to the extent that a signal arises from a comparison between having physical possession versus no possession, there was no signalling in our experiments because in any session, either all subjects have possession or they all have no possession. PZ (2007) argue that placement of the goods might signal relative value, and a signal of relative value can only arise when subjects get to compare, either between themselves or between multiple goods. Neither of these comparisons is possible in our experiments.

Second, there is no classical preference theory that links the location of a good to the perception of value; one could just as easily argue that subjects may treasure goods in their possession less than those that are not in their possession, for example because "the grass is always greener on the other side".

Third, the offers recorded were blind to the experimenters and so the experimenters' intention was irrelevant. Furthermore, the classical preference theories assume that people maximize self-interest without other-regarding preferences.

could very well turn the WTA-WTP gap on and off by inducing or eliminating various subject misconceptions or classical incentives, they could also cause the on-off results because they strengthened or weakened the reference states from which gains and losses were evaluated.

Our three procedural modifications were motivated by theory and evidence in the existing literature. One literature investigates the determinants of the reference state with the following findings. First, actual and physical possessions matter and people seem unaware of changes in their tastes that come with possessions. Loewenstein and Adler (1995) show that sellers who were asked to imagine having the mugs but not given actual and physical possessions<sup>18</sup> underestimated their valuations when given possessions.<sup>19</sup> Second, while ownership is generally an important factor in determining people's reference, ownership alone is neither necessary nor sufficient because other determinants of reference state such as expectations could result in the adoption of a different reference state. For example, Köszegi and Rabin (2006) propose a model in which the reference states are people's expectations of outcomes and suggest that one interpretation of PZ's findings is that "they have successfully decoupled subjects' expectations from their initial ownership status" (p.1142).<sup>20</sup>

Researchers have found some empirical support for expectation-based reference-dependent preferences: Abeler, Falk, Goette, and Huffman (2011) find that subjects' earnings expectations determine their effort provision in real-effort experiment, whereas Ericson and Fuster (2010) find that subjects who have a lower exogenous probability of being able to trade their item (and who therefore expect to keep it) are less likely to choose to trade when trade opportunity arises in an exchange experiment. Furthermore, consistent with Köszegi and Rabin's interpretation, Knetsch and Wong (2009) find that the effect of ownership on the reference state can be undermined by other manipulations that weaken the perception of this otherwise natural reference state. Using exchange experiments, they show that to establish

---

<sup>18</sup> A mug was displayed at the front of the room.

<sup>19</sup> Each subject was actually given a mug. To describe possession, Loewenstein and Adler (1995) used phrases like "obtain the mug", "receive a mug", or "get a mug" but never used words like "own" or "ownership" in their experimental instructions.

<sup>20</sup> The reference state is basically left unspecified in Kahneman and Tversky (1979) and has generally been assumed to be the status quo. But as Köszegi and Rabin (2006) point out, "in virtually all experiments interpreted as supporting this assumption, subjects plausibly expect to keep the status quo, so these studies are also consistent with the reference point being expectations" (p.1142).

reference state and endowment effect, ownership is neither necessary nor sufficient.<sup>21</sup> Nevertheless, Knetsch and Wong's (2009) results give no indication of the monetary strength of the endowment effect. This paper extends their work to measure the monetary strength of the endowment effect while controlling for misconceptions over complex elicitation procedures and other classical incentives such as transaction costs and cascade effects.

Closely related is the literature on framing. A central finding in behavioural economics is that framing matters, where "extensionally equivalent descriptions lead to different choices by altering the relative salience of different aspects of the problem" (Kahneman, 2003, p.1458). Rabin (1998) argues that "an important and predictable influence of framing on choice relates to loss aversion... Because losses resonate with people more than gains, a frame that highlights the losses associated with a choice makes that choice less attractive" (p.36). This literature shows that the reference state is determined by how choices are framed and presented, providing the motivation for framing in gain/loss terms in our manipulation treatments.

Another literature investigates anchoring effect (Tversky and Kahneman, 1974) in valuation tasks. Johnson and Schkade (1989) show that valuation of lotteries may be affected by the anchoring effect. For example, asking subjects whether their certainty equivalent for a lottery is above or below an anchor value influences subsequently stated certainty equivalents. Ariely, Loewenstein, and Prelec (2003) apply the anchoring manipulation to the valuation of products and hedonic experiences. They show that an incentive-compatible Becker-DeGroot-Marschak's (1964) mechanism and full information conditions fail to remove anchoring effect in the WTA elicited. Moreover, Jacowitz and Kahneman (1995) document an anchoring-like effect in responses to the Yes/No question itself: the mere consideration of a proposition tends to increase the plausibility of that proposition itself, especially if subjects reason that a value mentioned in the question is unlikely to be absurd. Green, Jacowitz, Kahneman, and McFadden (1998) find similar effects

---

<sup>21</sup> In particular, Knetsch and Wong (2009) show that "in a treatment that provides essentially all of the procedural controls that are deemed necessary to control for incentives recognized in standard theory but does not a priori create a large dichotomy between the perceived reference state and the initial entitlement, the participants exhibit a strong reluctance to trade away initial entitlements despite a lack of ownership of this entitlement", whereas "in a treatment that includes procedures that may control for other incentives but can also be expected to create a large dichotomy between the perceived reference states and the initial endowments, the participants show no such reluctance to trade despite having ownership of the original entitlement" (p.408).

in the WTP for public goods from Yes/No referendum responses and open-ended responses.

### **Robustness Check**

Our first two procedural modifications strengthen the reference states of buyers and sellers through possession and framing, while the last removes possible anchoring or suggestion of value by changing the values in the numerical examples. We now investigate to what extent possession and framing alone could have contributed to the gap's on-off results. We ran another four experimental sessions (Sessions 9-12) at the National University of Singapore on two consecutive days in 2011. We recruited another 105 students using similar procedures as before.<sup>22</sup> In these sessions (referred to as KW2), we implemented only the possession and framing manipulations, but used the same numerical examples as PZ. Furthermore, like PZ, sellers in KW2 were simply told that: "You can inspect the mug. You now own the mug."<sup>23</sup>

For completeness, we present the results of statistical tests for all possible pairings of the individual sessions for buyers and sellers. Nevertheless, because of the small sample size, we only discuss the results based on pooled data in what follows. Table 2B shows that in our manipulation treatments of KW2, we find a median WTP of \$2 (mean of \$2.36) and a median WTA of \$4.5 (mean of \$4.15), leading to a median WTA-WTP ratio of 2.25 (mean of 1.76).

Table 3 shows that in our manipulation treatments of KW2, there was again strong evidence that WTP was different from WTA at the conventional significance levels; the Wilcoxon-Mann-Whitney rank sum test results in a  $z$  value of -4.12 ( $p=0.0000$ ) and the median test results in a Pearson  $X^2$  statistic of 14.8974 ( $p=0.000$ ). Thus, possession and framing could account for the contrasting results between our replication of PZ and our manipulation treatments. The bottom line seems to be that the finding of no endowment effect is not robust to small changes in experimental details.

### **Further Checks**

---

<sup>22</sup> Sessions 9 and 11 were for buyers, whereas 10 and 12 were for sellers.

<sup>23</sup> In other words, sellers in KW2 were not told that: "the mug you are inspecting is yours to take home with you at the end of the experiment, if you decide to keep it later."

Our primary goal is to investigate the on-off results of the WTA-WTP gap. Nevertheless, it may be interesting to check whether there is a statistically significant difference between the WTPs in our replication of PZ and our manipulation treatments, and similarly for the WTAs. Table 4 reports the results of Wilcoxon-Mann-Whitney test and median test using pooled data. Panel A reports the results using all offers. Panel B reports the results when we exclude potential outliers – WTA that exceeded \$10 in session 2.

The results suggest that our manipulation treatments did reduce buyers' WTP and raise sellers' WTA relative to our replication of PZ. Nevertheless, the differences between WTPs are not statistically significant at the conventional levels, and likewise for the differences between WTAs. At face value, these results mean that while our procedural modifications cause only statistically insignificant changes to WTA and WTP, respectively, their cumulative effects on the WTA-WTP gap are jointly statistically significant. This interpretation reflects the subtlety of our procedural modifications. But these results could also reflect the limitations of our experimental setup: because the item used for valuation was a mug, this had limited the range of values elicited. For buyers, there was little room for downward adjustment given that the median WTP in our replication of PZ was already at \$2.5. For sellers, there were very few procedural differences between our manipulation treatments and our replication of PZ except some instances of framing. Collectively, these features in our experimental setup may have limited individual differences that are identifiable across treatments.

It is worth emphasising again that the primary purpose of this paper is to investigate why valuation disparities (i.e., the gap between WTA and WTP) appear or disappear with different experimental procedures. We design our experiments specifically for this purpose. To investigate how individual procedures affect the WTP and WTA is beyond the scope of this paper because it requires a completely different experimental design. For example, we will need sharper difference between the procedures for sellers across different treatments. Consequently, the results in this section should only be viewed as preliminary exploration.

### **Further Discussion and Conclusions**

ILS ask what makes experiments that use goods like mugs to elicit valuations different from experiments that use lotteries, “such that when the same experimental

procedures are applied to both, WTP-WTA disparities can persist in lottery valuations but not in mug valuations” (ILS, p.19)?<sup>24</sup> We have shown some evidence of disparities in mug valuations in large sample. Still, if weak reference states explain the results as we argue, it is natural to ask whether there are any reasons why the reference states may be stronger in lottery valuations than mug valuations when PZ’s procedures are used. We argue that for lottery tasks involving uncertain money outcomes, when there is feedback on the realizations of lottery outcomes for lottery forgone, such as in PZ’s and ILS’s experiments, anticipatory regrets (Loomes and Sugden, 1982) may be stronger for sellers who have sold the lotteries they are endowed with than for buyers who have not bought the lotteries. Anticipated regret aversion triggers stronger loss aversion, resulting in more significant WTA-WTP disparities in lottery valuations than in mug valuations.<sup>25</sup>

The strength of the reference states may depend subtly on the economic environment and its complexity. Procedurally-induced weak reference states may be an especially important issue in laboratory experiments where the experimenter gives ownership *arbitrarily* and *instantaneously* without creating any sense of entitlement to the newly endowed good. In these experiments, the determinants of reference state such as ownership, status quo, expectations, and possession may sometimes not coincide with the establishment of the same reference state, resulting in weak reference states.

We acknowledge that it may be difficult to specify a set of sufficient conditions for the reference state to be potent once we accept that the reference state

---

<sup>24</sup> ILS themselves suggest two structural differences between mug and lottery valuation tasks. First, in the lottery tasks but not the mug task, “the response-mode units are also used in specifying the objects that are being valued” and so “lottery tasks may prompt respondents to use ‘anchoring’ heuristics that are not applicable to other tasks,” such as the mug task. Second, tasks involving lotteries with monetary outcomes induce additional types of misconceptions, for which the PZ procedures do not control. This is the view that PZ endorse, as they argue that “for the lottery rounds we made no attempt to apply the revealed theory method” (2011, p.1018). Of course, these two structural differences may very well be the reasons why valuation disparities remain in the lottery tasks but not the mug task when PZ’s procedures are used.

<sup>25</sup> Empirical evidence suggests that providing feedback on ex-post outcomes for gambles forgone makes the reference states of “what would have been had the subjects made a different choice” more salient, inducing anticipated regrets (Zeelenberg, Beattie, van der Pligt and de Vries, 1996) and a higher WTA for sellers of lottery (Humphrey, Mann, and Starmer, 2005). There is also some evidence that actual ownership is not necessary to evoke anticipated regret as long as there is some sense of entitlement to the lottery even when one is not playing. Specifically, Zeelenberg and Pieters (2004) show that in the Netherlands, people anticipate more regret over not playing the Postcode Lottery, where one’s postcode is the ticket number and so non-participants may feel a sense of endowment to the postcode lottery and they may still find out that they would have won had they played, than in a traditional State Lottery.

has more than one determinants; what suffices in a simple and intuitive environment may be inadequate in a complicated environment that bears little resemblance to real life situations, as more complicated procedural controls are included to rule out potential confounding factors. But this difficulty only highlights the need for more future research to uncover the relative importance of individual determinants of reference state and their interactions with the economic environment.

In summary, PZ found that the WTA-WTP gap disappeared when they adopted all known procedural controls for misconceptions. Therefore they conclude that the gap arises from subject misconceptions rather than asymmetric valuation of gains and losses relative to the reference state. Our results suggest that the evidence is not robust enough to draw this conclusion. First, when we replicated their procedures, keeping all necessary controls for misconceptions but excluding the lottery rounds, we found that the gap was always statistically insignificant in samples of comparable size to PZ. Thus, we replicated PZ's finding of no gap without any lottery rounds. This supports PZ's claim that experience with the lottery rounds are not essential for their results. We also found that increasing sample size improved the gap's statistical significance. This evidence is consistent with the hypothesis of weak reference states: weak reference states cause statistical tests to have low power and increasing sample size is expected to raise the power of tests. Second, when we modified subtle details of their procedures to strengthen the reference states of buyers and sellers, again keeping all key controls for misconceptions, we found a moderate but highly statistically significant gap. Thus, experimental procedures could turn reference dependence and hence the gap on and off by inducing strong or weak reference states without necessarily changing the extent of control for misconceptions. The strength of reference states also seems to be a more parsimonious explanation of the gap's on-off results than subject misconceptions or other classical incentives.

Finally, it should be noted that Plott and Zeiler (2011) conjecture that any procedural changes – such as the location of the mugs or framing in gains and losses – may have led to very subtle enhancement effects and demand effects rather than loss aversion. Nevertheless, the exact mechanism through which this enhancement is supposed to take place is unclear (see footnote 17). Furthermore, with this conjecture, it is unclear how the hypothesis of subject misconceptions is rejectable because no procedural changes that lead to the finding of an endowment effect can be used to

reject the hypothesis.<sup>26</sup> While demand effects can certainly affect experimental findings, neither this conjecture nor the opposite conjecture can be rejected using the data from our experiments or PZ's experiments. The opposite conjecture is that for every enhancement effect or demand effect hypothesised, there is an equally plausible demand effect working in the opposite direction: from the elaborate measures the experimenters took to ensure that there are few differences between having ownership and no ownership, it is natural for subjects to infer that they are supposed to behave as if ownership is irrelevant.<sup>27</sup> In fact, it seems plausible that demand effects are more likely when subjects observe unexpected treatments or apparent deviations from the more intuitive treatments – such as buyers having possession of the mugs they do not own and framing buy/sell as similar valuation task – because these features register more saliently in the mind of subjects, not the other way around.

At least part of the controversy may be due to different beliefs on how the endowment effect experiments should be run, including how to adequately control for subject misconceptions or to establish reference states. But when there is no consensus on how the experiments should be run, sensitivity to minor procedural changes and slight relaxation of strong behavioural assumption – that the knowledge of ownership is sufficient to determine the reference state regardless of the other conditions – are valid robustness issues. Basically, ownership needs to be internalized before it will become a reference state that affects decision and behaviour.<sup>28</sup> The bottom line is that the evidence is not robust enough to warrant the strong conclusion that subject misconceptions are what cause the gap and the gap has nothing to do with loss aversion or prospect theory. Apart from the interpretation, the results also have implications on the conditions under which valuation disparities are likely to be observed: if no valuation disparities are observed only under very singular

---

<sup>26</sup> It is also unclear how one can test the external validity of the hypothesis. For example, ILS apply PZ's procedures to the valuation of lotteries, and find an endowment effect. But PZ (2011) argue that this evidence cannot be used to reject the hypothesis of subject misconceptions because for the lottery task, a different set of procedural controls for subject misconceptions should be used.

<sup>27</sup> In fact PZ (2005) made a similar point: "A third interpretation is that the procedures themselves involve a type of demand effect in which the subjects perceive that the experimenter wants to strip from responses any special value of ownership... the conjecture itself cannot be rejected using the data from our experiments" (p.543).

<sup>28</sup> The "internalization" of the endowment or thinking about the good as part of the individual's endowment may require more than cognitively understanding the meaning of ownership or the probability of endowment, especially in a complicated environment that bears little resemblance to real life situations.

experimental setup with little tolerance for small deviations, then valuation disparities are expected to prevail under more general conditions in the real world.

The bottom line is that our results show that what really matter for the gap's on-off results may be the subtle procedural details, such as possession and framing in gain/loss terms, rather than the general procedural controls for misconceptions, such as ensuring subject anonymity, using an incentive-compatible elicitation mechanism, explaining to subjects why truthful revelation is the optimal response. These pivotal procedural details are not unambiguously or purely related to control for misconceptions. Knowing which procedures positively affect the gap's on-off results is crucial for future research and discussion.

## References

- Abeler, J., Falk, A., Goette, L. and Huffman, D.** 2011. "Reference Points and Effort Provision." *American Economic Review*, 101(2): 470-492.
- Ariely, Dan, George Loewenstein, and Drazen Prelec.** 2003. "Coherent Arbitrariness: Stable Demand Curves without Stable Preferences." *Quarterly Journal of Economics*, 118(1): 73-105.
- Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak.** 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science*, 9(3): 226–232.
- Brown, Thomas C.** 2005. "Loss Aversion without the Endowment Effect, and Other Explanations for the WTA-WTP Disparity." *Journal of Economic Behavior & Organization*, 57(3): 367-379.
- DellaVigna, Stefano.** 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315-372.
- Ericson, Keith M. Marzilli, and Andreas Fuster.** 2010. "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." *Working Paper*.
- Genesove, David, and Christopher Mayer.** 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics*, 116(4): 1233-1260.
- Green, Donald, Karen E. Jacowitz, Daniel Kahneman, and Daniel McFadden.** 1998. "Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods." *Resources and Energy Economics*, 20(2): 85–116.
- Horowitz, John K., and Kenneth E. McConnell.** 2002. "A Review of WTA/WTP Studies." *Journal of Environmental Economics and Management*, 44(3): 426-447.
- Humphrey, Steven J., Paul Mann, and Chris Starmer.** 2005. "Testing for Feedback-Conditional Regret Effects Using a Natural Lottery." CeDEx Discussion Paper No. 2005-07, University of Nottingham.

- Isoni, Andrea, Graham Loomes, and Robert Sugden.** 2011. "The Willingness to Pay-Willingness to Accept Gap, the 'Endowment Effect', Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment." *American Economic Review*, 101(2): 991-1011.
- Jacowitz, Karen E., and Daniel Kahneman.** 1995. "Measures of Anchoring in Estimation Tasks." *Personality and Social Psychology Bulletin*, 21(11): 1161–1166.
- Johnson, Eric J., and David A. Schkade.** 1989. "Bias in Utility Assessments: Further Evidence and Explanations." *Management Science*, 35(4): 406–424.
- Kahneman, Daniel.** 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review*, 93(5): 1449-1475.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler.** 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy*, 98(6): 1325-1348.
- Kahneman, Daniel, and Amos Tversky.** 1979. "Prospect Theory: An Analysis of Decisions under Risk." *Econometrica*, 47(2): 263-291.
- Knetsch, Jack L., Fang-Fang Tang, and Richard H. Thaler.** 2001. "The Endowment Effect and Repeated Market Trials: Is the Vickrey Auction Demand Revealing?" *Experimental Economics*, 4(3): 257-269.
- Knetsch, Jack L., and Wei-Kang Wong.** 2009. "The Endowment Effect and the Reference State: Evidence and Manipulation." *Journal of Economic Behavior & Organization*, 71(2): 407-413.
- Kőszegi, Botond, and Matthew Rabin.** 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133-1165.
- Loewenstein, George, and Daniel Adler.** 1995. "A Bias in the Prediction of Tastes." *The Economic Journal*, 105(431): 929-937.
- Loomes, Graham, and Robert Sugden.** 1982. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *The Economic Journal*, 92(368): 805-824.

- Plott, Charles R., and Kathryn Zeiler.** 2005. “The Willingness to Pay – Willingness to Accept Gap, the ‘Endowment Effect,’ Subject Misconceptions, and Experimental Procedures for Eliciting Valuations.” *American Economic Review*, 95(3): 530-545.
- Plott, Charles R., and Kathryn Zeiler.** 2007. “Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?” *American Economic Review*, 97(4): 1449-1466.
- Plott, Charles R., and Kathryn Zeiler.** 2011. “The Willingness to Pay – Willingness to Accept Gap, the ‘Endowment Effect,’ Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Reply” *American Economic Review*, 101(2): 1012-1028.
- Rabin, Matthew.** 1998. “Psychology and Economics.” *Journal of Economic Literature*, 36(1): 11-46.
- Samuelson, William, and Richard Zeckhauser.** 1988. “Status Quo Bias in Decision Making.” *Journal of Risk and Uncertainty*, 1(1): 7-59.
- Thaler, Richard H.** 1980. “Toward a Positive Theory of Consumer Choice.” *Journal of Economic Behavior & Organization*, 1(1): 39-60.
- Tversky, Amos, and Daniel Kahneman.** 1974. “Judgment under Uncertainty: Heuristics and Biases.” *Science*, 185(4157): 1124–1131.
- Zeelenberg, Marcel, Jane Beattie, Joop van der Pligt, and Nanne K. de Vries.** 1996. “Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making.” *Organizational Behavior and Human Decision Processes*, 65(2): 148-158.
- Zeelenberg, Marcel, and Rik Pieters.** 2004. “Consequences of Regret Aversion in Real Life: The Case of the Dutch Postcode Lottery.” *Organizational Behavior and Human Decision Processes*, 93(2): 155-168.

**Table 1 – Design Features of Experiment Treatments**

Treatment	PZ (Replication)	KW (Manipulation)	KW2 (Manipulation)
Statistically significant WTA-WTP gap?	Mostly No	Yes	Mostly Yes
Buying Task	N=31 (Session 1) N=30 (Session 5)	N=30 (Session 3) N=31 (Session 7)	N=28 (Session 9) N=24 (Session 11)
Selling Task	N=29 (Session 2) N=31 (Session 6)	N=30 (Session 4) N=28 (Session 8)	N=23 (Session 10) N=30 (Session 12)
Possession at time of decision	Yes for sellers Yes for buyers	Yes for sellers No for buyers	Yes for sellers No for buyers
Framing or reference formation	Use “the mug” for both buyers and sellers.  Sometimes frame selling as in terms of “giving up” their mugs but not always.	Use “the mug” for buyers, but use possessive determiners for sellers (e.g., “my mug”).  Always frame selling as in terms of “giving up” their mugs.	Use “the mug” for buyers, but use possessive determiners for sellers (e.g., “my mug”).  Always frame selling as in terms of “giving up” their mugs.
Possible anchoring or suggestion of value	Examples use numerical values that are plausible for a	Examples use numerical values that are absurd for a coffee mug and emphasize	Examples use numerical values that are plausible for a coffee mug

	coffee mug	that they are hypothetical	
Controls for Misconceptions			
1. Incentive-compatible elicitation device	Yes	Yes	Yes
2. Numerical examples explaining why it is not optimal to underbid or overbid	Yes	Yes	Yes
3. Anonymity in decision			
4. Use of lotteries	Yes	Yes	Yes
5. Buyers and sellers are different subjects	No Yes	No Yes	No Yes

Note: Sessions 1-4 were conducted in 2009. Sessions 5-12 were conducted in 2011.

**Table 2A – Individual Subject Data and Summary Statistics**

Session	Treatment	Individual responses (in S\$)	N	Mean	Median	Std. dev.
1	PZ.Buyer: WTP 17/2/2009, 2-3pm	0, 0, 0.45, 1, 1, 1, 1.2, 1.5, 1.8, 2, 2, 2, 2, 2.5, 2.5, 2.5, 2.5, 2.5, 3, 3, 3.45, 3.5, 3.55, 3.9, 4, 4, 4, 4, 4.5, 4.5, 5.75	31	2.57	2.50	1.42
2	PZ.Seller: WTA 18/2/2009, 2-3pm	0, 0, 0.5, 1, 1, 1, 1.5, 2, 2, 2, 2, 2.5, 2.5, 3, 3, 3.5, 4, 4.5, 4.5, 4.5, 5, 5.5, 5.5, 6, 6, 8.5, 10, 12.5, 15	29	4.10	3.00	3.60
3	KW.Buyer: WTP 17/2/2009, 3-4pm	0, 0, 0, 0.5, 0.5, 1, 1, 1, 1, 1, 1.5, 1.5, 2, 2, 2, 2, 2, 2, 2.5, 2.5, 2.5, 2.5, 2.5, 3, 3.5, 4.5, 5.5, 8, 8.9, 10	30	2.56	2.00	2.51
4	KW.Seller: WTA 18/2/2009, 1-2pm	0, 0, 0, 0.5, 1.5, 2, 2, 2, 2.5, 2.5, 2.5, 3, 3, 3.5, 3.5, 4.5, 4.5, 4.5, 4.5, 5, 5, 5, 5, 5.5, 5.5, 5.5, 6.5, 7, 7.5, 10	30	3.80	4.00	2.36
5	PZ.Buyer: WTP 20/1/2011, 2-3pm	0, 0, 0.5, 0.5, 0.5, 0.8, 1, 1, 1, 1, 1, 1.2, 1.5, 1.5, 1.6, 2, 3, 3.5, 3.5, 3.5, 3.5, 4, 4, 4.5, 4.9, 5, 5.5, 5.5, 7, 10	30	2.75	1.80	2.35
6	PZ.Seller: WTA 20/1/2011, 3-4pm	0, 0, 0.5, 0.5, 0.5, 0.5, 1, 1.5, 2, 2, 2, 2, 2, 3, 3, 3.5, 3.5, 4.5, 4.5, 4.5, 5, 5.5, 5.5, 5.5, 6, 6, 7, 7, 7, 8, 8.5	31	3.61	3.50	2.54
7	KW.Buyer: WTP 21/1/2011, 2-3pm	0, 0, 0, 0, 0.5, 0.5, 0.5, 0.5, 1, 1, 1, 2, 2, 2, 2.5, 2.5, 2.5, 2.5, 3, 3, 3, 3, 3.5, 3.5, 3.5, 4, 5, 5, 6.5, 6.5, 7	31	2.50	2.50	2.00
8	KW.Seller: WTA	0, 0, 0.5, 1, 2, 2, 2, 2.5, 2.5, 3, 4, 4, 4, 4, 4.5, 4.5, 5, 5, 5, 6.5, 7, 7, 7.5,	28	4.54	4.25	2.83

	21/1/2011, 3-4pm	7.5, 8, 9, 9, 10				
9	KW2.Buyer: WTP 27/1/2011, 2-3pm	0, 0.49, 0.5, 1, 1.1, 1.5, 1.5, 1.9, 2, 2, 2, 2, 2.5, 2.5, 2.5, 2.5, 3, 3, 3.5, 3.5, 4, 4, 4.9, 5, 5, 5, 7, 8.5	28	2.94	2.50	1.96
10	KW2.Seller: WTA 27/1/2011, 3-4pm	0.5, 0.5, 1, 1.5, 2, 2, 2.5, 3, 3, 4, 4.5, 4.5, 4.5, 5, 5, 5.5, 6, 6.5, 6.5, 7, 7.5, 7.5, 8.5	23	4.28	4.50	2.39
11	KW2.Buyer: WTP 28/1/2011, 2-3pm	0, 0, 0, 0, 0, 0, 1, 1, 1.5, 1.5, 1.5, 1.5, 2, 2, 2, 2, 2, 2, 2.5, 2.5, 2.8, 3, 3.9, 5.75	24	1.69	1.75	1.39
12	KW2.Seller: WTA 28/1/2011, 3-4pm	0, 0, 0, 1, 2, 2, 2.5, 3, 3, 3, 3, 3.5, 3.5, 3.5, 3.5, 4, 4.5, 4.5, 4.5, 4.5, 4.5, 4.95, 5, 5.5, 6, 6.5, 7, 8, 8.5, 10	30	4.05	3.75	2.41

Note: Sessions 1-4 were conducted in 2009. Sessions 5-12 were conducted in 2011. All Sessions were conducted during 2-4pm on weekdays except Session 4, which was conducted during 1-2pm. All Sessions were completed in about 30 minutes within the indicated time slots.

**Table 2B –Summary Statistics – Pooled Data**

Sessions	Treatment	N	Mean	Median	Std. dev.
1 & 5	PZ.Buyer: WTP	61	2.66	2.50	1.92
2 & 6	PZ.Seller: WTA	60	3.85	3.25	3.08
3 & 7	KW.Buyer: WTP	61	2.53	2.00	2.25
4 & 8	KW.Seller: WTA	58	4.16	4.25	2.60
9 & 11	KW2.Buyer: WTP	52	2.36	2.00	1.82
10 & 12	KW2.Seller: WTA	53	4.15	4.50	2.38

Note: Sessions 1-4 were conducted in 2009. Sessions 5-12 were conducted in 2011.

**Table 3 – Statistical Test Results: WTP vs. WTA**

	Wilcoxon-Mann-Whitney rank sum test (Null hypothesis: identical distributions)			Median test (Null hypothesis: populations have identical medians)		
	<i>z</i>	<i>p</i> -value	Conclusion ( $\alpha = .05$ )	Pearson $X^2$	<i>p</i> -value	Conclusion ( $\alpha = .05$ )
<u>Panel A: All Offers</u>						
<u>PZ treatment</u>						
Sessions 1 vs. 2	-1.484	0.1378	Can't reject null	0.5880	0.443	Can't reject null
Sessions 5 vs. 6	-1.411	0.1583	Can't reject null	0.1528	0.696	Can't reject null
Sessions 1 & 5 vs. 2 & 6	-2.102	0.0356	Reject null	0.9921	0.319	Can't reject null
<u>KW treatment</u>						
Sessions 3 vs. 4	-2.495	0.0126	Reject null	8.2127	0.004	Reject null
Sessions 7 vs. 8	-2.832	0.0046	Reject null	6.0146	0.014	Reject null
Sessions 3 & 7 vs. 4 & 8	-3.654	0.0003	Reject null	11.4718	0.001	Reject null
<u>KW2 treatment</u>						
Sessions 9 vs. 10	-2.059	0.0395	Reject null	2.2771	0.131	Can't reject null
Sessions 9 vs. 12	-2.005	0.0450	Reject null	3.3833	0.066	Can't reject null
Sessions 11 vs. 10	-3.720	0.0002	Reject null	9.3725	0.002	Reject null

Sessions 11 vs. 12	-4.025	0.0001	Reject null	16.8750	0.000	Reject null
Sessions 9 & 11 vs. 10 & 12	-4.120	0.0000	Reject null	14.8974	0.000	Reject null
<u>Panel B: Exclude Offers &gt; 10</u>	<i>z</i>	<i>p</i> -value	Conclusion ( $\alpha = .05$ )	Pearson $X^2$	<i>p</i> -value	Conclusion ( $\alpha = .05$ )
<u>PZ treatment</u>						
Sessions 1 vs. 2	-1.080	0.2803	Can't reject null	0.2414	0.623	Can't reject null
Sessions 5 vs. 6	NA	NA	NA	NA	NA	NA
Sessions 1 & 5 vs. 2 & 6	-1.831	0.0671	Can't reject null	1.0129	0.314	Can't reject null

Note: The Pearson  $X^2$  statistics were corrected for continuity.

“NA”: There are no offers > 10 for PZ treatment for the sessions in 2011.

**Table 4 – Statistical Test Results: WTP vs. WTP; WTA vs. WTA**

	Wilcoxon-Mann-Whitney rank sum test (Null hypothesis: identical distributions)			Median test (Null hypothesis: populations have identical medians)		
	<i>z</i>	<i>p</i> -value	Conclusion ( $\alpha = .05$ )	Pearson $\chi^2$	<i>p</i> -value	Conclusion ( $\alpha = .05$ )
<u>Panel A: All Offers</u>						
<u>Buyer's WTP: PZ vs. KW</u> Sessions 1 & 5 vs. 3 & 7	0.897	0.3698	Can't reject null	0.1311	0.717	Can't reject null
<u>Buyer's WTP: PZ vs. KW2</u> Sessions 1 & 5 vs. 9 & 11	0.824	0.4102	Can't reject null	0.7882	0.375	Can't reject null
<u>Seller's WTA: PZ vs. KW</u> Sessions 2 & 6 vs. 4 & 8	-0.963	0.3353	Can't reject null	0.2929	0.588	Can't reject null
<u>Seller's WTA: PZ vs. KW2</u> Sessions 2 & 6 vs. 10 & 12	-1.096	0.2730	Can't reject null	0.7098	0.400	Can't reject null
<u>Panel B: Exclude Offers &gt; 10</u>						
<u>Seller's WTA: PZ vs. KW</u> Sessions 2 & 6 vs. 4 & 8	-1.310	0.1903	Can't reject null	1.6897	0.194	Can't reject null

<u>Seller's WTA: PZ vs. KW2</u>						
Sessions 2 & 6 vs. 10 & 12	-1.439	0.1502	Can't reject null	1.0664	0.302	Can't reject null

Note: The Pearson  $\chi^2$  statistics were corrected for continuity. There are no offers > 10 by PZ's buyers.

## Not for Publication

### Appendix: Summary of Experimental Instructions

This appendix contains a summary of the experimental instructions to facilitate easy comparison. Actual experimental instructions used by the experimenters and written instructions received by the subjects are available upon request.

The instructions for PZ replication are shown in black, whereas the instructions for our manipulation treatments (KW) are shown in square brackets and in red, e.g., [red], when they deviate from the instructions for PZ replication. Underlined text preceded by Note to Experimenter is meant for the experimenters only and not in the written instructions received by the subjects.

#### General Instructions for All Experiments

Note to Experimenter: Prior to the beginning of the experiment, the experimenter placed an example of record sheet on the overhead projector. The subject's offer, fixed offer price, and result calculation were left blank. As the participants come in, tell them that they will get their participation fees at the end.

WRITTEN INSTRUCTIONS – Give to Subjects (Also read aloud)

Please follow my instructions carefully. Do not get ahead of my verbal instructions.

This is an experiment in individual decision-making. Our purpose is to study technical issues involved in decision-making. A research grant has provided funds for this research.

The instructions are simple and your final outcome will depend on the decisions you make.

You have received a record sheet with an individual identification number on it. This is your private information. Do not share it with anyone. We ask that you do not communicate with other people during the experiment. Please refrain from verbally reacting to events that occur during the experiment. This is very important.

If you have any questions, please raise your hand and wait for the experimenter to attend to you.

#### Instructions for Seller

Note to Experimenter:

Possession condition: give a mug to each participant. Leave the mug with the participant. At the end of the experiment, each participant will bring the mug to the person outside (sequentially) and settle the trade – which means either give the mug back in exchange for the fixed offer or walk away with the mug. In the possession condition, make sure that the mugs are out of their boxes while in possession.

As the mugs were being distributed, say: “You can inspect the mug. You now own the mug.” [As the mugs were being distributed, say: “You can inspect the mug. You now own the mug. The mug you are inspecting is yours to take home with you at the end of the experiment, if you decide to keep it later.”]

The experimenter will offer to buy the mug that you own. Your task is to make an offer for the mug [your mug] and record it on your record sheet.

## Not for Publication

As you will see, your best strategy is to determine the minimum you would be willing to accept for the mug [to give up your mug] and offer that amount. It will not be to your advantage to offer more than this amount, and it will not be to your advantage to offer less. Simply determine the minimum you would be willing to accept and make that amount your offer.

Your offer will be compared to a fixed offer. The fixed offer was determined in the following way:

We prepared 21 pieces of identical papers, with 0, 0.5, 1, 1.5, 2, 2.5, ..., 8.5, 9, 9.5, 10 written on them. We folded the papers and dropped them into a bag. Just before the session, we randomly drew a number from the bag and sealed the number into the envelope you see here. We redraw the number for every session. This randomly drawn number in the envelope will be the fixed offer.

As you can see, the fixed offer will be completely unrelated to your offer and to the offers of all other persons in the room.

If your offer is less than or the same as the fixed offer then you sell the mug [your mug]. You had the low offer, so you are the seller. But, here's the interesting part. **You do not receive your offer.** Instead, you receive the fixed offer, a price equal to or higher than your offer.

[Note to Experimenter: say "As a hypothetical example..."]

Example: if you offer 1,000 and the fixed offer is 1,020, you have the low offer. You sell the item [your item] and you receive the fixed offer of 1,020.

If your offer is more than the fixed offer then you do not sell the item [your item]. You keep the item [your item].

Example: if you offer 1,000 and the fixed offer is 950, you do not have the low offer. Therefore, you do not sell the item [your item].

Note to experimenter: Illustrate how the record sheets will be used in the two examples above. Ask "any questions?"

As a seller, you should offer the **minimum amount you would be willing to accept** in exchange for [to give up] the mug you own.

Remember, there are no advantages to strategic behaviour. Your best strategy is to determine your personal value for the mug [your mug] and record that value as your offer. There is not necessarily a "correct" value. Personal values can differ from individual to individual.

The following example illustrates how you work out what's the minimum you are willing to accept.

"Imagine that I am a seller and I own Item B. How do I know what amount is the minimum I'd be willing to accept to give up Item B [my item]?"

Start with \$100 [\$1000]. Would I be willing to give up Item B [my item] in exchange for \$100 [\$1000]? If so, then decrease the amount to \$95 [\$950]. If I'm willing to accept \$95 [\$950] to give up Item B [my item], then decrease further. I keep decreasing until I come to an amount that makes me indifferent between keeping Item B [my item] and getting the money.

[Note to Experimenter: say "As a hypothetical example..."]

Example:

Would I accept \$10 [\$1000] to give up Item B [my item]? Yes.

Would I accept \$8 [\$800] for B [to give up my item]? Yes.

## Not for Publication

Would I accept \$7 [\$700] for B [to give up my item]? Yes.

Would I accept \$6 [\$600] for B [to give up my item]? No, not \$6 [\$600]. So I need to increase.

Would I accept \$6.50 [\$650]? I don't care whether I end up with \$6.50 [\$650] or Item B [keeping my item].

Then that is the minimum I'd be willing to accept for Item B [to give up my item]. I'll record that number on my record sheet.

The key to determining the minimum you'd be willing to accept is remembering that you will not receive the amount you ask for. Instead, if you receive anything, you will always get the fixed offer.

Why is my best strategy to bid the minimum I'd be willing to accept? Let's go back to the Example:

[Note to Experimenter: say "As a hypothetical example..."]

Say I decide that the minimum I'd be willing to accept for Item B [to give up my item] is \$6.50 [\$650].

What happens if I ask for less than \$6.50 [\$650]? Say I ask for only \$6 [\$600].

If the fixed offer is, say, \$6.25 [\$625], then I have to sell my item. I lose out because I have to give up Item B [my item] which I think is worth \$6.50 [\$650], but I only get \$6.25 [\$625] in exchange.

What happens if I ask for more than \$6.50 [\$650]? Say I ask for \$7 [\$700].

If the fixed offer is \$6.75 [\$675], then I do not sell. But, had I bid \$6.50 [\$650], I would have sold the item [my item] and received \$6.75 [\$675] for an item that I think is worth only \$6.50 [\$650]. I lose out.

Note to experimenter: Ask "any questions?"

At the end of the exercise, please detach your Record Sheet, fold it and drop it into the box in front. Then take the remaining portion of the written instructions containing your ID number to the cashier outside the room, who will execute the transactions (SELL or NO SELL at the fixed offer) and pay you your participation fees. Note that the cashier outside will not see your offer. The experimenter will not be able to link any specific subject to a subject ID number and his/her offer. Thus, your offer will be known only to yourself.

Note to Experimenter: say "All transactions (SELL or NO SELL) are real. If you SELL, then you will receive the fixed offer (not your offer) from the cashier outside in exchange for the mug [but you will give up your mug]. If NO SELL, then you will not receive the fixed offer but you will take home the mug [but you will keep your mug]."

Encourage and address questions from subjects. After answering questions, start the experiment.

After the experiment, remind subjects to copy only the SELL or NO SELL decision in step 3 of the Record Sheet to the previous page. Ask the subjects to drop their record sheets into the box in front. Then ask them to leave sequentially. The person outside pays the participation fees, settles the trade and collects the ID number back. The person outside will not see the subjects' offers.

**Not for Publication**

ID Number: \_\_\_\_\_

Keep this page to yourself. Show your ID number to the person outside this room after the experiment. At the end of the experiment, based on your decision in Step 3 on the Record Sheet, Circle **SELL** or **NO SELL**

## Record Sheet

ID Number: \_\_\_\_\_

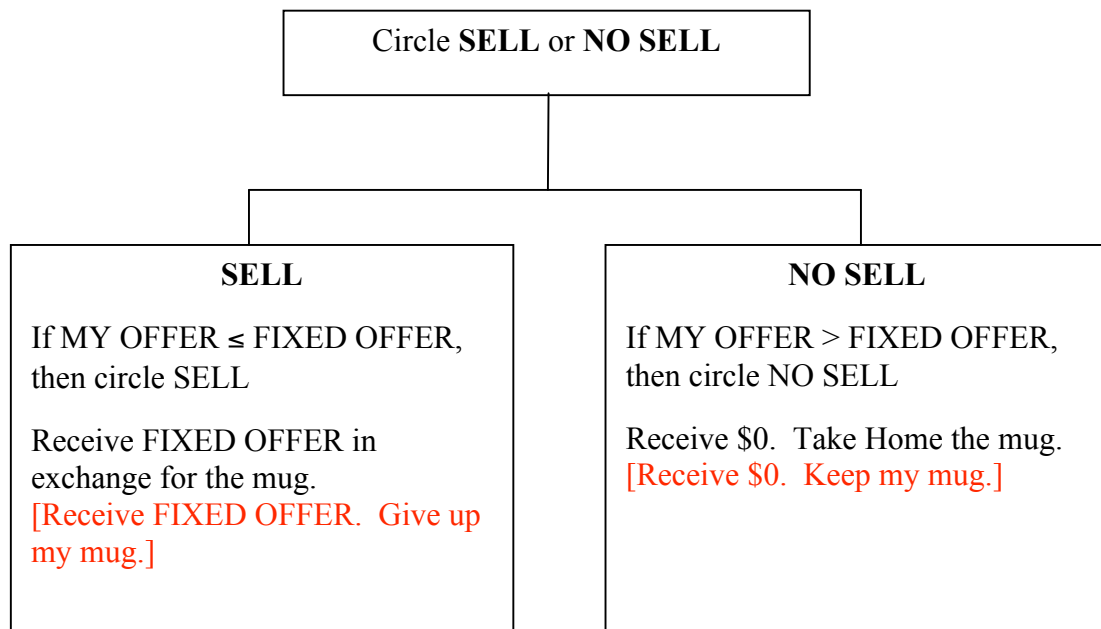
Step 1: decide on my offer.

MY OFFER : \$\$ \_\_\_\_\_

Step 2: listen for fixed offer announcement

FIXED OFFER : \$\$ \_\_\_\_\_

Step 3:



Step 4: Detach this record sheet and drop it into the box in front.

## Not for Publication

### Instructions for Buyer

Note to Experimenter: Hand out the mugs. Say: “You can inspect the mug, but you do not own it.”

Possession condition: give a mug to each participant for inspection. Leave the mug with the participant. At the end of the experiment, each participant will bring the mug to the person outside (sequentially) and settle the trade – which means either give the mug back or pay the fixed offer. In the possession condition, make sure that the mugs are out of their boxes while in possession.

[No possession condition: give a few mugs for all participants to pass around and inspect. After everybody has inspected a mug, take the mugs from the participants and leave them in the desk in front.]

The experimenter will offer a mug (the one you inspected) [just like the one you inspected] for sale. [We have enough mugs on hand should everybody decide to buy one.] Your task is to make an offer for the mug and record it on your record sheet.

As you will see, your best strategy is to determine the maximum you would be willing to pay for the mug and offer that amount. It will not be to your advantage to offer more than this amount, and it will not be to your advantage to offer less. Simply determine the maximum you would be willing to pay and make that amount your offer.

Your offer will be compared to a fixed offer. The fixed offer was determined in the following way:

We prepared 21 pieces of identical papers, with 0, 0.5, 1, 1.5, 2, 2.5, ..., 8.5, 9, 9.5, 10 written on them. We folded the papers and dropped them into a bag. Just before the session, we randomly drew a number from the bag and sealed the number into the envelope you see here. We redraw the number for every session. This randomly drawn number in the envelope will be the fixed offer.

As you can see, the fixed offer will be completely unrelated to your offer and to the offers of all other persons in the room.

If your offer is more than or the same as the fixed offer then you buy the mug. You had the high offer, so you are the buyer. But, here’s the interesting part. **You do not pay the amount you offered.** Instead, you pay the fixed offer, an amount equal to or less than your offer.

[Note to Experimenter: say “As a hypothetical example...”]

Example: if you offer 1,000 and the fixed offer is 950, you have the high offer. You buy the item but pay only 950.

If your offer is less than the fixed offer then you do not buy the item. Instead, you keep your money.

Example: if you offer 1,000 and the fixed offer is 1,020, you do not have the high offer. Therefore, you do not buy the item. You keep your money.

Note to experimenter: Illustrate how the record sheets will be used in the two examples above. Ask “any questions?”

As a buyer, you should offer exactly the **maximum amount you would be willing to pay** in exchange for the mug.

## Not for Publication

Remember, there are no advantages to strategic behaviour. Your best strategy is to determine your personal value for the mug and record that value as your offer. There is not necessarily a “correct” value. Personal values can differ from individual to individual.

The following example illustrates how you work out what’s the maximum you are willing to pay.

“Imagine that I am a buyer and Item A is up for sale. How do I know what amount is the maximum I’d be willing to pay for Item A?

Start with 1 cent [\$1]. Would I be willing to pay 1 cent [\$1] for the item? If so, then increase the amount to 2 cents [\$2]. If I’m willing to pay 2 cents [\$2], then increase further. I keep increasing until I come to an amount that makes me indifferent between keeping the money and getting Item A.

[Note to Experimenter: say “As a hypothetical example...”]

Example: Would I pay \$1 [\$100] for A? Yes.

Would I pay \$2 [\$200] for A? Yes.

Would I pay \$5 [\$500] for A? Yes.

Would I pay \$6 [\$600] for A? No, not \$6 [\$600]. So I need to decrease.

Would I pay \$5.50 [\$550]? No, not that much.

How about \$5.25 [\$525]? I don’t care whether I end up with \$5.25 [\$525] or the item.

Then that is the maximum I’d be willing to pay for Item A. I’ll record that number on my record sheet.

The key to determining the maximum you’d be willing to pay is remembering that you will not pay the amount you bid. Instead, if you pay anything, you will pay the fixed offer.

Why is my best strategy to bid the maximum I’d be willing to pay? Let’s go back to the Example:

[Note to Experimenter: say “As a hypothetical example...”]

Say that I decide that the maximum I’d be willing to pay for Item A is \$5.25 [\$525].

What happens if I bid less than \$5.25 [\$525]? Say I bid \$5 [\$500].

If the fixed offer is, say, \$5.10 [\$510], then I don’t get the item. Had I bid \$5.25 [\$525], I would have received the item and had to pay only \$5.10 [\$510] for an item that I think is worth \$5.25 [\$525]. I lose out.

What happens if I bid higher than \$5.25 [\$525]? Say I bid \$5.50 [\$550].

If the fixed offer is \$5.45 [\$545], then I have to pay \$5.45 [\$545] for an item that I really think is worth only \$5.25 [\$525]. I lose out.

Note to experimenter: Ask “any questions?”

At the end of the exercise, please detach your Record Sheet, fold it and drop it into the box in front. Then take the remaining portion of the written instructions containing your ID number to the cashier outside the room, who will execute the transactions (BUY or NO BUY at the fixed offer) and pay you your participation fees. Note that the cashier outside will not see your offer. The experimenter will not be able to link any specific subject to a subject ID number and his/her offer. Thus, your offer will be known only to yourself.

Note to Experimenter: say “All transactions (BUY or NO BUY) are real. If you BUY, then you will pay the fixed offer (not your offer) to the cashier outside in exchange for the mug. If NO BUY, then you will not pay the fixed offer but you will also not get the mug.”

## Not for Publication

Encourage and address questions from subjects. After answering questions, start the experiment. Tell the subjects that if they do not have enough money with them, we will hold the mug for them and they can come back for it later.

After the experiment, remind subjects to copy only the BUY or NO BUY decision in step 3 of the Record Sheet to the previous page. Ask the subjects to drop their record sheets into the box in front. Then ask them to leave sequentially. The person outside pays the participation fees, settles the trade and collects the ID number back. The person outside will not see the subjects' offers. If a subject has not enough money, tell him/her to come back with enough money and we will hold the mug for him/her.

ID Number: \_\_\_\_\_

Keep this page to yourself. Show your ID number to the person outside this room after the experiment. At the end of the experiment, based on your decision in Step 3 on the Record Sheet, Circle **BUY** or **NO BUY**

# Record Sheet

ID Number: \_\_\_\_\_

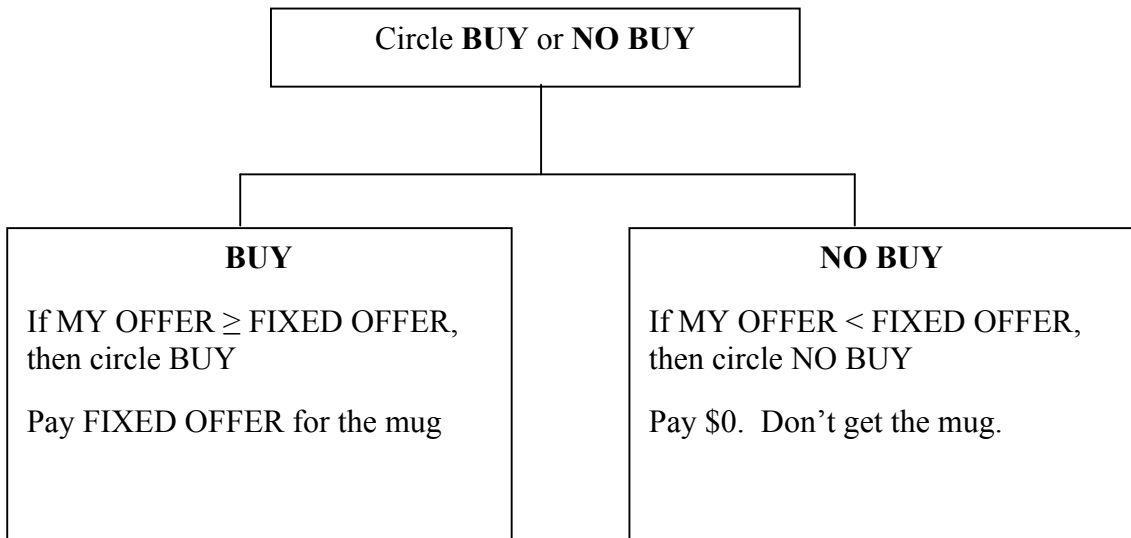
Step 1: decide on my offer.

MY OFFER : S\$ \_\_\_\_\_

Step 2: listen for fixed offer announcement

FIXED OFFER : S\$ \_\_\_\_\_

Step 3:



Step 4: Detach this record sheet and drop it into the box in front.