Tom Sawyer and the Construction of Value

Dan Ariely, George Loewenstein, and Drazen Prelec

Abstract:

This paper challenges the common assumption that economic agents know their tastes. After reviewing previous research showing that valuation of ordinary products and experiences can be manipulated by non-normative cues, we present three studies showing that in some cases people do not even have a pre-existing sense of whether an experience is good or bad – even when they have experienced a sample of it.

JEL Classifications: D00 microeconomics, general
Keywords: preferences, preference uncertainty, coherent arbitrariness

Dan Ariely is Luiz Alvarez Renta Professor of Management Science at MIT and a visiting scholar at the Federal Reserve Bank of Boston's Research Center for Behavioral Economics and Decision-Making. His email address is ariely@mit.edu. George Loewenstein is Professor of Economics and Psychology in the Department of Social and Decision Sciences at Carnegie Mellon University. Drazen Prelec is Professor of Management Science at MIT.

This version: July 2005

In a famous passage of Mark Twain's novel, *Tom Sawyer,* Tom is faced with the unenviable job of whitewashing his aunt's fence in full view of his friends who will pass by shortly and whose snickering promises to add insult to injury. When his friends do show up, Tom applies himself to the paintbrush with gusto, presenting the tedious chore as a rare opportunity. Tom's friends wind up not only paying for the privilege of taking their turn at the fence, but deriving real pleasure from the task—a win-win outcome if there ever was one. In Twain's words, Tom "had discovered a great law of human action, without knowing it—namely, that in order to make a man or a boy covet a thing, it is only necessary to make the thing difficult to attain."

There are no mysteries in what painting a fence entails. Hence, Tom's "law" challenges the intuition that whether a familiar activity or experience is pleasant or unpleasant is a self-evident matter—at least to the person participating in that activity. If true, Tom's law would pose a fundamental challenge to economics. In a world where people don't reliably know what they like, it cannot be assumed that voluntary trades will improve well-being or that markets will increase welfare.

Recent research by psychologists and behavioral economists suggests that Twain's notions about human nature may be on the mark, at least in some situations [Frederick and Fischhoff (1998), Hsee, Loewenstein, Blount, and Bazerman (1999), Kahneman, Ritov, and Schkade (1999), Slovic (1995), Sunstein, Kahneman, Schkade, and Ritov (2002)]. In a set of previous experiments that document a phenomenon we labelled "coherent arbitrariness," we [Ariely, Loewenstein, and Prelec (2003)] showed that valuations of goods and experiences have a large arbitrary component—"arbitrariness." Yet after one valuation has been made, people

provide subsequent valuations that are scaled appropriately relative to the first—"coherence." In one study that illustrates the concept, we sold consumer products ranging in value from $10 to $100 (computer equipment, wine bottles, chocolate boxes, books) to postgraduate business students. Students were presented with one product at a time and were asked whether they would buy it for a price obtained by converting the last two digits of their Social Security number (an essentially random identification number) into a dollar figure—for example, 34 became $34. After this yes/no response, which we intended to serve as an "anchor" for their later responses [Chapman and Johnson (1999), Johnson & Schkade (1989)], we elicited their maximum willingness to pay for the product, using a standard incentive-compatible procedure.

Although students were reminded that the Social Security number is a random quantity conveying no information, those who happened to have high Social Security numbers were willing to pay much more for the products. For example, students with Social Security numbers in the bottom 20 percent of the Social Security number distribution priced on average a '98 Cotes du Rhone wine at $8.64, while those with Social Security numbers in the top 20 percent of the distribution priced on average the same bottle at $27.91. Because the assignment of Social Security numbers to students is random, we can regard the two groups as identical with respect to their underlying tastes and knowledge of wine. Evidently, the same person can value a given item at $10 or at $30, depending on historical accidents such as answering questions about randomly generated prices.

If consumers' valuations of goods are so malleable, then why does one observe stable demand curves in the marketplace? A second aspect of the study provides a clue. If one looks across the different goods that were sold, one can see that, while absolute values were surprisingly malleable, subjects did seem to have a sensible idea of the relative values of the

different goods. For example, because the information that was provided about the wine (based on *The Wine Advocate* magazine) made it obvious which wine was superior, all students priced the relatively fancy bottle of wine—the '96 Hermitage Jaboulet "La Chapelle"—higher than the '98 Cote du Rhone. The students did not know how much they valued either bottle, as demonstrated by the impact of the arbitrary Social Security number, but they did know that the superior wine was worth more than the inferior wine, and they priced them accordingly. A researcher who looked at our data but did not know about the Social Security number manipulation would conclude that these consumers were behaving perfectly in line with economic theory—the more valuable products were indeed priced higher than the less valuable ones. But the impact of the Social Security number manipulation shows that valuations, in fact, incorporated a high degree of arbitrariness.

Although the effect of the arbitrary Social Security number on valuations was dramatic, one could argue that the result crucially exploited subjects' uncertainty about what the goods were worth, either to them or in the market (for resale perhaps). A stronger test required an experience that could not be traded, but that could be experienced, and hence fully understood before the valuation task. For this purpose, many of the experiments reported in the same paper elicited subjects' willingness to accept compensation in exchange for listening to aversive sounds delivered to subjects through headphones. The benefit of using sounds is that subjects can be given a sample that provides full information about the experience. In an experiment representative of several reported in the paper, we told subjects that they were about to hear an unpleasant sound played over headphones and asked them to consider, hypothetically, whether they would be willing to listen to the sound for 300 seconds in exchange for an amount that they composed from the last three digits of their Social Security number (for example, 287 = $2.87).

After hearing a sample of the sound and making this hypothetical choice, subjects then stated the smallest amount of money they would accept to actually hear the sound for 100, 300, and 600 seconds. After each response, they endured the sound and received payment if their stated minimum fell below a randomly drawn price for that duration. Even in such a transparent setting, we found that valuations followed the 'coherently arbitrary' pattern: Subjects demanded about one and a half times as much to hear the 300 second sound as to hear the 100 second sound, and half again more to hear the 600 second sound. However, subjects with lower ending Social Security numbers demanded much less compensation than those with higher numbers.

These results showed that individuals did not seem to have a pre-existing personal dollar value for ordinary products and experiences. Taking these findings as a starting point, the present paper asks a more basic question: *Do people even have a pre-existing sense of whether an experience is good or bad?* Tom's "law" suggests that they do not—that the exact same experience can be desired or avoided, depending on context and presentation. The three experiments presented here show that individuals can be made to exogenously classify some experiences as either positive or negative, depending on whether the preceding question asked them if they would pay or needed to be paid for the experience in question. Experiment 1 demonstrates the basic effect of "Tom's law. The results of Experiment 2 show, further, that after one such arbitrary response is given (as either positive or negative), other responses follow in a seemingly coherent fashion. Finally, to rule out two competing explanations—that the effect can be attributed to the demands of the situation or that subjects use the form of the initial anchoring question to infer the quality of the experience—Experiment 3 demonstrates that "Tom's law" holds even when the random assignment of individuals to either the 'pay' or 'be paid' conditions

was made transparent (the assignment was set by the last digit of their Social Security Number, and both possibilities were presented on the elicitation form).

Experiment 1. The purpose of the first experiment was to test whether there are experiences that individuals can perceive as either positive or negative. The study was conducted with 146 undergraduate students enrolled in a marketing class at the University of California at Berkeley. At the end of class, respondents were told that, in a week's time, their professor (Ariely, with whom they were familiar) would be conducting a 15-minute poetry reading from Walt Whitman's "Leaves of Grass." Half of the respondents (N=75) were then asked whether, hypothetically, they would be willing to pay $2 to listen to their professor recite poetry. The other half of the respondents (N=71) were asked whether, hypothetically, they would be willing to accept $2 to listen to their professor recite poetry. After answering one of these hypothetical questions, all respondents were told that the poetry reading scheduled for next week was going to be free and were asked to indicate if they wanted to be notified via email about its location and time. The goal of this question was to test whether the initial hypothetical question affected whether respondents viewed the experience as positive (meaning that they would like to attend if it was free) or negative (meaning that they would prefer not to attend if it was free).

The results in Table 1 show that the poetry reading did not have great appeal: Only 3 percent of the respondents were willing to pay $2 to listen to Ariely recite poetry. However, most (59 percent) respondents were willing to endure the recital for $2. More important for our purpose was the response to the second question. The percentage of respondents willing to attend the free poetry recitation was 35 percent when they had first been asked if they would pay to attend the recital, but only 8 percent when they had first been asked whether they would attend

the recital in exchange for pay, t(144) = 4.0, p < 0.001. The first response clearly influences whether individuals view the experience as positive or negative.

--- Table 1 ---

Experiment 2. Experiment 2 (conducted with a large undergraduate class at MIT) was designed to replicate Experiment 1, while also examining consistency within an individual across responses. Half of the respondents (N=91) were first asked whether, hypothetically, they would be willing to pay $10 to listen to Professor Ariely recite poetry for 10 minutes, followed by a request to indicate their monetary valuations for 1, 3, and 6 minutes of poetry reading. The other half (N=73) were first asked whether, hypothetically, they would be willing to accept $10 to listen to Professor Ariely recite poetry for 10 minutes, followed by a request to indicate the minimum they would be willing to accept for 1, 3, and 6 minute of poetry reading. After indicating their prices, respondents were asked similar questions about their willingness to participate in a study of decision-making, which was described to them by giving an example of the Wason card selection task. Subjects who had been asked about their willingness to pay to listen to poetry in the first part of the study were asked in this part of the experiment about their hypothetical willingness to accept $10 for a 10 minutes participation in such a study. Subjects who had been asked about their willingness to accept compensation to listen to poetry were asked about their hypothetical willingness to pay $10 for a 10 minutes participation in such a study. After this initial hypothetical question, subjects were asked to indicate their monetary valuations for participating in such a study for 1, 3, and 6 minutes.

As is evident from Figure 1, valuations were strongly influenced by the initial question. Individuals in the pay condition were, on average, willing to pay for the experience, while individuals in the accept payment condition, on average, required compensation to undergo the same experience. Furthermore, respondents consistently indicated higher sums of money for longer durations, whether it was a matter of paying for or being paid for the experience. Respondents did not have a pre-existing sense whether the poetry reading (or participating in a decision-making experiment) was a good or bad experience, but they knew that either way "more" of the experience warranted greater payment.

---- Figure ---

Experiment 3. Our third and final experiment addresses the concern that the subjects might have taken the initial question as a cue about the quality or value of the poetry reading event. To that end we modified the procedure in two ways. First, before providing any response, subjects heard a one-minute sample of poetry reading from their professor (the experiment took place at a large undergraduate class at MIT). Direct exposure to a sample of the experience should logically diminish the significance of any indirect cues. Second, the instructions made explicit that there were two different conditions and that the assignment to one or the other condition was random.

The experiment began with the professor announcing that in a few days he would be conducting a reading of poetry from Walt Whitman's *Leaves of Grass*. For about a minute, he then read the first few verses of "Whoever you are holding me now in hand."

> *Whoever you are holding me now in hand,*
> *Without one thing all will be useless,*
> *I give you fair warning before you attempt me further,*

I am not what you supposed, but far different.

Who is he that would become my follower?
Who would sign himself a candidate for my affections?

The way is suspicious, the result uncertain, perhaps destructive,
You would have to give up all else, I alone would expect to be your
        sole and exclusive standard,
Your novitiate would even then be long and exhausting,
The whole past theory of your life and all conformity to the lives
        around you would have to be abandon'd,
Therefore release me now before troubling yourself any further, let
        go your hand from my shoulders,
Put me down and depart on your way.

After reading the sample the professor added that: "The full reading will last fifteen minutes. The number of seats is fixed, but I would like to get something close to a full house. I am not sure how many of you are interested in attending this, so I have developed a scheme for allocating the seats to those who are most interested in the reading."

The scheme was explained on the instructions sheet. Subjects were first asked to write down the last digit of their Social Security number; if the number was odd, to answer the first question from the two questions below; and if the number was even, to answer the second question from the two questions below. Both questions were printed on the same page, one immediately above the other. In the first question (odd digits only), they first inserted the last digit of their Social Security number into the sentence: "Would you attend the poetry reading for a payment of $_____," and then answered it with either a YES or a NO. They also answered the same question with the dollar value increased by 50¢ and decreased by 50¢. In the second question (even digits only), they inserted the last digit of their Social Security number into the

sentence: "Would you pay $_____ in order to attend the poetry reading," and then answered with a YES or a NO. They also answered the question with a 50¢ increase and decrease.

Subjects then turned to the back page of the questionnaire, which elicited a two-sided demand curve, ostensibly to help the experimenter determine whether he "will need to pay subjects for attending." Here, the format was identical for both groups. Subjects indicated whether they "would attend the recital if paid $10, $9, $8, .., $1," whether they "would attend the recital if it was for free," and whether they "would pay $1, $2, …, $10" to attend the recital. In all, subjects answered 21 pricing questions, which specified their personal value (positive or negative) of the recital to within +/- 50¢, within the range of +$10 to -$10.

The results are summarized in Table 2 below. On the initial question, a majority of subjects (63 percent) were willing to attend the recital if paid the dollar equivalent of their digit, but only 20 percent were willing to pay that equivalent to attend. More importantly, the subsequent valuations of the reading were strongly influenced by the initial question: Of those who initially considered whether they would attend if paid, only 9 percent were willing to attend for free, in contrast to the 49 percent who would do so among those who initially considered whether they would pay for the reading, $t(79) = 4.48$, $p < .0001$. The mean value of the reading was negative for both groups, but much more so for those who had been asked if they would attend for pay, who required on average $4.46 in compensation to attend the reading, as opposed to the $1.13 required by those who were asked if they would pay to attend, $t(79) = 4.28$, $p <. 0001$. Interestingly, in this study the actual digit of the Social Security number had no effect on valuations nor on willingness to attend for free, after controlling for whether the digit was odd or even. It seems that the main impact of the anchoring question was to determine whether the subjects perceived the experience as positive or negative, not *how* positive or negative.

--- Table 2 ---

The results of Experiment 3 are consistent with the previous two experiments and effectively rule out an inferential explanation for the "Tom Sawyer" effect. It is hard to imagine how the subject could infer anything about the quality of the poetry reading from a transparently random assignment. The experiment also rules out a demand effect explanation, because any conceivable demand effects would surely be present in exactly the same form across both experimental conditions.

Discussion. Looking around, we see people making a myriad of choices, ranging from the trivial to the profound. People decide whether or not to purchase Big Macs, to smoke, run red lights, take vacations in Patagonia, listen to Wagner, slave away at doctoral dissertations, marry, have children, live in the suburbs, vote Republican, and so on. The apparent orderliness in these choices, their stability for a given individual, and the generally correct directional response to changing incentives encourage the belief that the choices are firmly rooted in personal likes and dislikes—in fundamental values.

We suggest, in contrast, that correct directional responses to changing incentives do not provide strong support for fundamental valuation, but can follow from the fact that people try to behave in a sensible manner when it is obvious how to do so. If earlier choices are recalled, the next time a similar choice situation arises, decision-makers will attempt to behave in a fashion that doesn't violate obvious rules of consistency. If one was willing to pay $x for company Y's stock yesterday, then today's announcement of an unexpectedly profitable quarter should make one willing to pay more than $x. Such a sensible decision-making heuristic, however, tells us nothing about whether yesterday's valuation was reasonable. Economists observe responsiveness to incentives and conclude that individuals are making choices based on fundamental valuation,

much as they would if they observed our experiment without awareness of the initial manipulation. However, as Summers (1986) points out in a seminal paper, such responsiveness is a necessary, but not sufficient, condition to establish fundamental valuation. People can respond sensibly to changes in conditions even when they do so from arbitrary baseline levels.

The degree of coherence and arbitrariness in any set of choices is likely to depend on many factors. Arbitrariness is enhanced by ambiguity in a good or experience. We deliberately selected somewhat ambiguous experiences in the three experiments presented here; clearly, some experiences are unambiguously good or bad. At the same time, many of the most important decisions that people make—about marriage, education, emigration, jobs and vacations—involve streams of heterogeneous experiences that are arguably even more difficult to assess, and hence are even more vulnerable to arbitrary influences and conventions, than our simple poetry-reading proposition [Ariely and Carmon (2003)]. Is a vacation that includes peaceful hours of reading on the beach, delicious meals, but also screaming children, money worries, and stressful transportation a good thing or a bad thing, on balance? This is a difficult question to answer.

Coherence will depend on how easy it is to spot behavioral inconsistencies, which depends in turn on whether numerical scales or indices are available, on how close in time choices are made, and on whether connections among them are obvious. In general, we should expect to see the greatest degree of coherence when an individual who has made an initial choice, and is now faced with a second choice, remembers the first, and is aware of how he or she must resolve the second choice so as to be consistent with the first. These requirements are regularly satisfied in a narrow range of decisions, of which financial decisions are perhaps the best example.

But even in highly rationalized financial markets, we have no assurance that the absolute levels of prices are 'sensible'. In financial markets, a stock's value is supposed to reflect individual investors' estimates about that company's expected stream of future dividends, appropriately discounted, but estimates of future dividend streams are inherently uncertain, and there is no agreed-upon discount rate. When we look at short-term fluctuations in prices, we do indeed see stock prices responding appropriately to good or bad news about individual companies or about the economy as a whole. In the long term, however, markets exhibit wide fluctuations that are completely out of line with historical fluctuations in dividend streams, as the economist Robert Shiller has argued persuasively (Shiller, 1998). Day by day, investors can see the coherence in short-term market responses, but nothing signals so clearly whether the market is over- or under-priced as a whole.

The arbitrariness inherent in people's preferences means that decisions can be influenced by nonnormative factors. One such factor is whether the decision-maker is exposed to alternatives simultaneously or one at a time [see Hsee *et al.* (1999)]. Economics generally assumes that all economic decisions are choices between alternatives. Thus, choosing whether or not to eat out at a restaurant tonight is implicitly a choice between eating at the restaurant versus spending the money in other ways. However, the phenomenon of coherent arbitrariness suggests that whether these choices are implicit or explicit may make a difference.

The distinction between explicitly considering alternative options and not doing so has important implications for empirical, and especially experimental, research in economics. One important design issue, when conducting experiments, is whether to compare different experimental treatments between- or within-subject—whether to expose individual subjects to different treatments or compare treatments across subjects. This decision is often viewed mainly

as an issue of time constraints and statistical power. However, considerable research suggests that people can behave quite differently in these two experimental setups. For example, Frederick and Fischhoff (1998) elicited willingness to pay (WTP) for two different quantities of common market goods (for example, toilet paper, applesauce, and tuna fish) using both a between-subjects design (in which respondents valued either the small or large quantity of each good) and a within-subjects design (in which respondents valued both the small and large quantity of each good). The difference in WTP was in the right direction in both designs, but it was much greater (2.5 times as large) in the within-subjects condition, which explicitly manipulated quantity. The usual finding is that within-subject manipulations produce larger effects than between-subject manipulations [for example, Fox and Tversky (1995), Keren and Raaijmakers (1988)], though there are situations in which the opposite occurs.[1]

Why does it matter whether economic decisions are determined by fundamental values? First, coherent arbitrariness violates the basic economic assumptions about how the 'general equilibrium' of an economy comes into existence. Modern economics assumes that exogenous consumer preferences interact with 'technologies' and initial endowments to produce equilibrium states of the economy—prices and production levels. This analysis falls apart if preferences are themselves influenced by the very equilibrium states that they are presumed to create. Indeed, in the domain of economic decision-making, the most salient and potentially powerful anchors may well be the public parameters of the economy itself—the relative prices and scarcities of different commodities. By posting a price for a new product, for instance, a firm invites consumers to consider whether they would purchase at that price and so replicates the anchoring manipulation as conducted in our experiments. If prices and other economic parameters function like public anchors, then consumer tastes no longer exist independently of prices but are

endogenous to the economy. In that case, the equilibrium price and production levels of the economy are no longer uniquely determined by its physical and human resources and characteristics. Rather, a certain price level may prevail because of collective anchoring, triggered by historical accidents or manipulations.

Second, economics as practiced today is not only a descriptive but also a prescriptive social science. Economists derive the 'welfare implications' of alternative policies, e.g., taxation or trade, where welfare is defined in terms of the degree to which a policy leads to the satisfaction of individual preferences. Economists have, of course, identified many situations in which free market exchange may not increase welfare [for example, Schelling (1978)]. Such market failures usually arise from interactions between people with asymmetric information or from situations in which people do not internalize the costs they impose on each other. The suboptimalities that arise from coherent arbitrariness, in contrast, begin at the level of the individual. If preferences have a large arbitrary component, then even strictly personal consumption choices by fully informed individuals need not maximize welfare.

Moreover, these individual-level effects can be exacerbated by social and market interaction. The literature on information cascades already shows that, when people are uncertain about the quality of consumption goods, initial choices can have big effects on market outcomes. Thus, for example, if a small number of early diners arbitrarily choose new restaurant A over new restaurant B, A can end up packed and B empty. The scope for such effects is enlarged to the degree that people are uncertain about their own preferences. Our research suggests that the degree of uncertainty may be very substantial, even when individuals have relevant experience with the objects of choice.

References:

Ariely, D., & Carmon, Z. (2003). Summary assessment of experiences: The whole is different from the sum of its parts. In G. Loewenstein, D. Read, & R. F. Baumeister (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 323-349). New York: Russell Sage Foundation.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics, 118*, 73-105.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes, 79*, 115-153.

Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics, 110*, 585-603.

Frederick, S., & Fischhoff, B. (1998). Scope (in)sensitivity in elicited valuations. *Risk Decision and Policy, 3*, 109-123.

Hsee, C. K., Loewenstein, G., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychological Bulletin, 125*, 576-590.

Johnson, E. J., & Schkade, D. A. (1989). Bias in utility assessments: Further evidence and explanations. *Management Science, 35*, 406-424.

Kahneman, D., Ritov, I., & Schkade, D. (1999). Economic preferences or attitude expressions? An analysis of dollar responses to public issues. *Journal of Risk and Uncertainty, 19*, 203-235.

Keren, G., & Raaijmakers, J. B. (1988). On between-subjects versus within-subjects comparisons in testing utility theory. *Organizational Behavior and Human Decision Processes, 4*, 233-247.

Schelling, T. C. (1978). *Micromotives and macrobehavior*. New York: Norton.

Shiller, R. J. (1998). Human behavior and the efficiency of the financial system. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (Vol. 1, pp. 1305-1340). Amsterdam: Elsevier.

Slovic, P. (1995). The construction of preference. *American Psychologist, 50*, 364-371.

Summers, L. H. (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance, 41*, 591-602.

Sunstein, C. R., Kahneman, D., Schkade, D., & Ritov, I. (2002). Predictably incoherent judgments. *Stanford Law Review, 54*, 1153-1215.

Table 1: Desire to listen to a free poetry reading in Experiment 1. The top panel summarizes the results for the participants who were first asked whether they will accept $2 to listen to poetry. The bottom panel summarizes the results for the participants who were first asked whether they will pay $2 to listen to poetry.

Accept Group	
Would you attend the recital for $2?	59% say "Yes"
Would you attend the recital for Free?	8% say "Yes"
Pay Group	
Would you pay $2 to attend the recital?	3% say "Yes"
Would you attend the recital for Free?	35% say "Yes"

Table 2: Results of Experiment 3. The top panel summarizes the results for participants who were first asked whether they would attend the poetry recital in return for the dollar equivalent of the last digit of their Social Security number. The bottom panel summarizes the results for participants who were first asked whether they would pay the dollar equivalent of the last digit of their Social Security number to attend the poetry recital.

Odd Social Security Number Digit (hypothetical question about being paid to attend) (N=46)	
% Willing to attend for $ = <u>Soc.Sec.No.</u>	63%
Would attend for free	9%
Mean valuation (st. error)	— $4.46 (.51)
Even Social Security Number Digit (hypothetical question about paying to attend) (N=35)	
% Willing to pay $ = <u>Soc.Sec.No.</u> to attend	20%
Would attend for free	49%
Mean valuation (st. error)	— $1.13 (.59)

Figure Legend

Experiment 2: Willingness to pay / accept money in US$ for different durations of poetry (right) and experiment participation (left) as a function of whether the hypothetical question was for paying (squares) or accepting payment (circles).

Figure 1
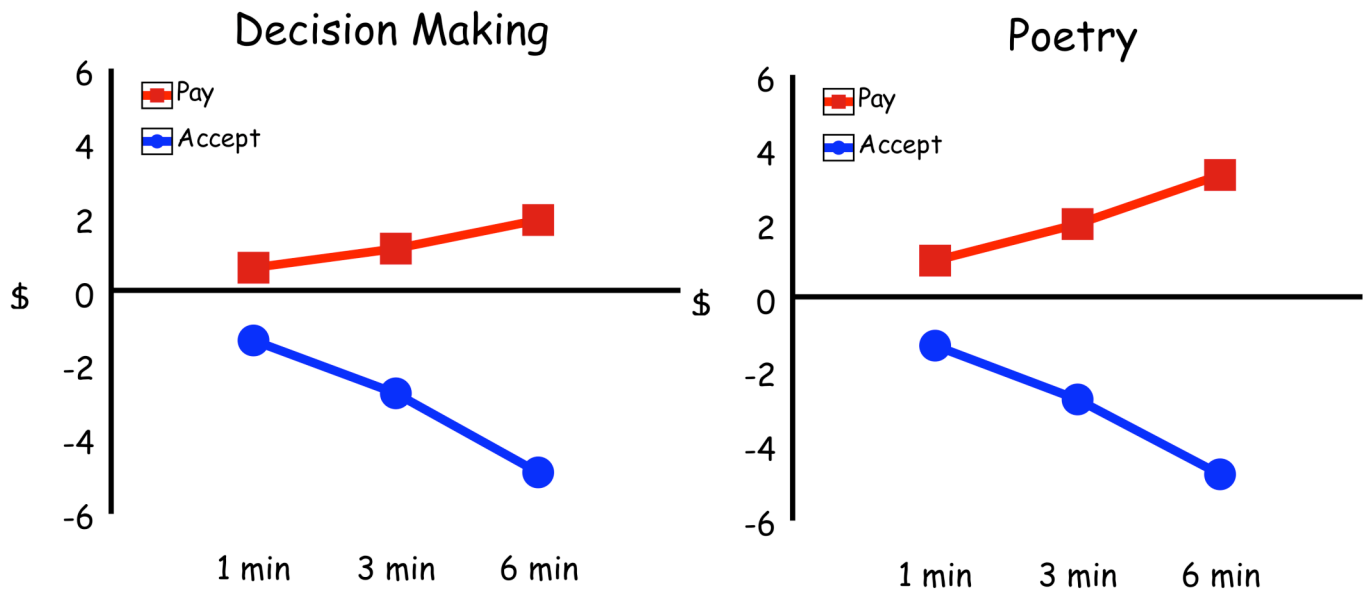
Notes

1. Between-subject experiments produce larger effects than within-subject experiments when the variable being manipulated is something that subjects do not think is normatively appropriate to respond to. Thus, for example, in a study of discrimination, one might get different responses to a black or white name on a resume if these were presented to separate subjects, but if the same subject saw both, they would be likely to recognize that race was being varied across resumes and, unless they were consciously and deliberately racist, would take efforts to evaluate both similarly.