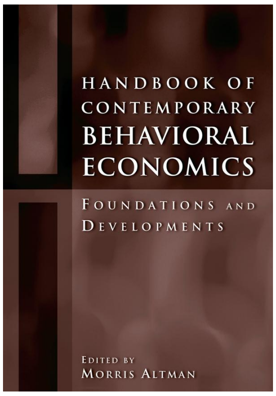

Handbook of Contemporary Behavioral Economics
Foundations and Developments
Morris Altman

Classroom Experiments in Behavioral Economics

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9781315703879.ch19
Gerrit Antonides, Fergus Bolger, Ger Trip
Published online on: 15 Jul 2006

PART 4

EXPERIMENTS AND IMPLICATIONS

Page Intentionally Left Blank

CHAPTER 19

CLASSROOM EXPERIMENTS IN BEHAVIORAL ECONOMICS

GERRIT ANTONIDES, FERGUS BOLGER, AND GER TRIP

Economic experiments have been popular ever since they were instrumental in the discovery of some famous economic paradoxes, such as the Allais, Ellsberg, and St. Petersburg paradoxes. Later experiments have been extended to problems outside the area of risk and uncertainty. Also, current economic experiments tend to involve real money or products, rather than hypothetical choices.

Economic experiments have become a tool for educational purposes. The earliest classroom experiments were conducted by Edward Chamberlin (1948), who studied market equilibria for buyers and sellers of hypothetical goods. Modern versions of Chamberlin's experiments are reported in Smith 1962, Holt 1996, and Fels 1993. Classroom experiments are but one type of experiment. Nowadays, several experimental setups can be distinguished, including laboratory experiments, classroom experiments (DeYoung 1993), and Internet experiments (Anderhub, Müller, and Schmidt 2001). Also, software is easily available for use in economics classes (e.g., Charles Holt's Web page, http://www.people.virginia.edu/~cah2k/home.html), and even a textbook for teaching economics by conducting experiments exists (Bergstrom and Miller 1999).

Experimental economics has become an industry. Laboratories for experimental economic research exist around the globe; the *Journal of Experimental Economics* has existed since 1998; and a *Handbook of Experimental Economics* has appeared (Kagel and Roth 1995). A good overview of activities, names, and Web sites in the industry is provided on Alvin Roth's Web page (http://www.economics.harvard.edu/~aroth/alroth.html).

There is a difference in focus between experimental and behavioral economics. Experimental economists usually test economic theories in market environments (i.e., auctions, rent seeking, provision of public goods, etc.). Several Web sites offer experimental setups as illustrations of economic theory (e.g., how to elicit a demand curve in class). Experimental economics aims at using insights from experiments to change market conditions in order to achieve efficient outcomes (Varian 2002). Behavioral economics refers more to the individual behavior of economic agents and subsequent research into the determinants of anomalous behavior, that is, behavior that is left unexplained by neoclassical economics. Our essay is more in line with behavioral economics.

The authors teach classes in the areas of economics, consumer behavior, and psychology. We use classroom experiments to illustrate the development of theories in these areas. We believe that students will be more interested and remember the courses better if they have personal experience with the working of the theories considered. Some of our classroom experiments were also used as pilot experiments for scientific research.

When we conduct experiments for research, we may have to use other than our own classes for two reasons. First, our own classes may be "framed" because they already may know something about the theories we are interested in, possibly leading to demand effects (Orne and Scheibe 1964). Second, our own classes usually are too small for experiments including different groups. Sometimes splitting a larger class into different groups is not feasible either—for example, if one group should not be aware of the experimental manipulations in the other group. In the laboratory, we usually assign the participants to different groups randomly, in order to avoid selection bias. When we have to use different classes we pay attention to the type of students in each. However, sometimes classes are formed according to the alphabetic order of the students' names. Such classes are ideal for use as random groups in an experiment.

Sometimes the types of student vary across different classes. In such cases there is the probability of selection affecting our results. For example, it is known that game theoretic classes may behave differently than social science classes in experiments on cooperative behavior (Frank, Gilovich, and Regan 1993). Some other factors that may selectively influence our results are gender, age, income, intelligence, ethnicity, and residence. Some of these variables may be included as co-variates in the analysis of results to assess their possible influence.[1]

Yet another type of classes we use are from Dutch secondary schools. Partly as a promotion for Wageningen University, mobile laboratories on economics, physics, chemistry, agriculture, and food are taken to secondary school classes, where pupils participate in the experiments. Since in this essay we report on several experiments from the mobile economics laboratory, we provide a brief overview of this project next.

MOBILE LABORATORY ON ECONOMICS

Wageningen University has developed several mobile laboratories for education in secondary schools in the Netherlands, including physics and chemistry laboratories. In 2002 the idea of creating a mobile laboratory in economics came up. The objectives of the project were twofold: stimulating scientific interest in economics and promoting Wageningen University. The basic idea was that by presenting interactive experiments derived from behavioral economics pupils would experience the richness of the economic discipline. Active participation of pupils was stimulated by using real products and real money. A provocative title—"Adam Smith Was Wrong"—was chosen, mainly to attract the attention of the teachers (most pupils do not know who Adam Smith was). This title was derived from the movie *A Beautiful Mind* about the life of John Nash. A clip of this movie was actually included in the laboratory.

In the academic year 2003–4 the economics laboratory was presented in approximately eighty classrooms all over the Netherlands. Each laboratory was presented by two students of Wageningen University, who were trained for two weeks and then went out touring for six weeks. Then the next pair was trained and went out to the schools.

The final form of the mobile economics laboratory consisted of a ninety-minute program. In exceptional cases, part of the program could be presented within a forty-five-minute framework. The full program was as follows: introduction, ultimatum game, framing experiment, endowment experiment, beauty contest (i.e., guess-the-number game), prisoner's dilemma experiment, and conclusion. The clip of the movie *A Beautiful Mind* shows four young men—one of them John Nash—who meet five young women in a café, one being blond and her friends being dark-haired. The young men prefer the blonde, but Nash makes clear that if they all go for the blonde they "will block each other," and after the men are rejected by the blonde, the dark-haired women will also lose interest "because nobody likes to be second choice." So some form of cooperation is needed to achieve the common goal, which is finding a girl for the night.

The main theme of the laboratory is the concept of economic rationality. If economic rationality is defined as a short-term maximization of own profit, regardless of the interests of others, then what can be concluded from the experiments in this laboratory? The pupils are invited to think about this key question, maybe inventing and elaborating their own experiment. Some lessons from the mobile laboratory that the pupils should take into account are that people do care about the interests of others, people behave inconsistently, and even if one is fully rational it is wise to take into account the irrationality of others. These lessons are well known from the behavioral approach to economics but have not reached the regular introductory textbooks. In the words of Kahneman: "A search through some introductory textbooks in economics indicates that if there has been any change, it has not yet filtered down to that level: the same assumptions are still in place as the cornerstones of economic analysis" (2003a, 162). The cornerstones Kahneman refers to are selfishness, rationality, and unchanging tastes (or consistency).

ENDOWMENT EFFECT

An important topic in behavioral economics is the idea that utility is not derived from total assets and levels of consumption but rather from changes with respect to these entities (Kahneman 2003b). Kahneman and Tversky's work on prospect theory (1979, 1992) points to the asymmetric evaluation of changes in the current state of affairs. The current state of affairs serves as a reference point for evaluating the changes. In particular, positive changes are evaluated less positively than negative changes are evaluated negatively. This has led to the popular credo that losses loom larger than gains. Because of this result, people in general are more eager to avoid losses than to acquire gains, which is called loss aversion.

Loss aversion has been investigated in different contexts. In finance, it has been observed that investors realize their gains too early and are reluctant to take their losses (Shefrin and Statman 1985; Odean 1998). In consumer behavior, people dislike product alternatives that in some respect deviate negatively from the products they currently use (Tversky and Kahneman 1991; Johnson et al. 1993). This phenomenon appears so strong that people in general seem to prefer the status quo over alternatives (Samuelson and Zeckhauser 1988). For example, when trading in their cars, consumers value a high trade-in price for their old car more than a discount on the new car, indicating loss aversion for their old car (Purohit 1995). Also, the sunk cost effect—that is, taking into account past investments when making current decisions—points to the psychological importance of lost assets or past expenses (Thaler 1980).

Probably the strongest illustration of loss aversion is the endowment effect, basically implying that goods in one's possession are valued higher than before they were possessed (Knetsch and Sinden 1984; Knetsch 1995). Ownership of a good seems to change the value placed on the good. The Coase theorem in standard economic theory claims that the value of a good should be independent of one's entitlement to the good (Coase 1960). The endowment effect is easily shown by randomly distributing two different goods, say A and B, among a number of people (Knetsch and Sinden 1984). Standard economic theory assumes that people would prefer either A or B or are indifferent. Hence, the standard assumption is that about half of the people have obtained the nonpreferred good and would be willing to exchange it for the other good. However, when asked for their willingness to exchange, in fact only 10 percent of the people want to exchange. This result substantially deviates from the standard economic expectation. Similar results were obtained by asking nonowners of a good for their willingness to pay (WTP) for the good. Kahneman, Knetsch, and Thaler (1990) report an average WTP of $2.21 for a mug. Likewise, owners of the good were asked for their willingness to accept (WTA) the loss of the good in exchange for a monetary compensation. The average monetary compensation required (WTA) was $5.78. The WTA was 161 percent higher than WTP, indicating the effect of loss aversion for the owners of the good.

How can we know that the endowment effect is due to loss aversion rather than "acquisition aversion" (resulting in lower WTP)? Kahneman, Knetsch, and Thaler (1990) compared product valuations of three groups: buyers, choosers, and sellers. Buyers' average WTP for a mug amounted to $2.87, whereas sellers' average WTA was $7.12. The WTA/WTP ratio of 2.5 clearly shows the endowment effect. Choosers neither owned a mug nor were asked to pay for the mug. They indicated for a number of different cash amounts whether they preferred the mug or cash. The amount at which choosers were indifferent between the mug and cash, $3.12 on average, indicated their value of the mug. Since the choosers' valuations were very close to the buyers' evaluation, the WTP/WTA disparity can hardly be explained by reluctance to pay for the mug but should be explained from loss aversion.

A number of other factors appear to influence the size of the endowment effect, including:

1. Reduction of the cognitive dissonance created by possible incompatibility between one's prior opinions concerning a good and the ownership of the good
2. Mere exposure, i.e., repeated exposure to a good tends to increase one's liking for the good
3. Mere possession, i.e., possessing a coupon or gift certificate for a good increases preference for the good
4. Mere ownership, i.e., people tend to judge their own possessions as more attractive than the possessions of others
5. Attachment, i.e., relatively high evaluation of products consistent with one's self-image, and products obtained by one's own effort rather than by chance
6. Transaction demand, i.e., the eagerness to buy or sell may reduce the endowment effect
7. Duration of ownership, i.e., the longer one owns a good, the stronger the endowment effect tends to be
8. Product-related factors: substitutability of goods tends to reduce the endowment effect and hedonic goods seem to be preferred in a forfeiture task, whereas functional goods seem to be preferred in an acquisition task (DeGroot 2003)

In our own research into factors influencing the endowment effect we frequently use classroom experiments. Our research shows how classroom experiments can be used both to replicate the endowment effect and to design relevant variations of the classical experiments.

Classroom Experiments on the Endowment Effect

Cognitive Dissonance Effect

Above we mentioned cognitive dissonance as a factor contributing to the endowment effect. Cognitive dissonance theory predicts that attitudes and opinions that are inconsistent with the actual situation will be changed in accordance with the situation (Festinger 1957; Cooper and Fazio 1984). For example, students who had to debate an issue (e.g., abortion) from a standpoint opposite to their own developed a more positive attitude toward the issue than before (Scott 1957). In this case, the situation was the actual defense of the opposite standpoint. In the case of the endowment effect, the situation is the legal entitlement to the good. So being endowed with a good might change one's attitude toward the good.

We conducted several experiments in which students randomly received one of a pair of goods. In one study, we used rolls of Top Drop or Top Gum (two types of licorice); in another study, we used Toblerone or Milka chocolate bars. We told the students that the product they had received was theirs to keep. When the students were offered the possibility of exchanging their good for the alternative, less than 20 percent wanted to trade (thus showing the endowment effect). Then we asked all students to justify their decisions.[2] Of those who did not want to trade, a large majority stated that they preferred the candy they had in their hand to the alternative, even though the initial distribution had been random. Clearly, simply receiving some candy had the effect, for many people, of making it their "most preferred."

Type of Good

Substitutability. The type of good may influence the size of the endowment effect. Hanemann (1991) suggested that substitutability of the goods would increase the willingness to trade. Chapman (1998) offered owners of a good the opportunity to trade their goods for both identical goods and similar goods (not exactly identical). Only part of the sample was willing to trade identical goods. Similar (not identical) goods were traded somewhat more easily when her participants received a small compensation (5 cents) for exchange, but only for those participants who were willing to trade the identical goods. In other circumstances the willingness to trade hardly differed across similar and dissimilar goods. Van Dijk and Van Knippenberg (1998) found even less willingness to trade wines from different countries than wines from the same country. In this particular case, similar goods were exchanged more than dissimilar goods. Our experience with a variety of snacks, pens, mugs, and postcards is that the endowment effect usually is quite strong, even for similar goods.

Evaluability. Hsee (1996) developed the idea that the ease of evaluating a good may influence a consumer's willingness to pay for the good under different circumstances. Easy-to-evaluate product attributes (e.g., broken dinnerware or damaged book covers) were found to be more important in situations where the good was evaluated in isolation. Product attributes that were hard to evaluate in isolation (e.g., number of entries in a dictionary) turned out to be more important when comparisons with similar goods were possible. The willingness to pay for (or the willingness to exchange) a hard-to-evaluate product may be lower than for an easy-to-evaluate good, leading to a larger endowment effect for the former than for the latter. We tested this hypothesis in a classroom setting.[3]

We randomly distributed Pentel fine-line pens and opaque drinking glasses among a group of twenty-nine law and economics students. Each student rated both products with respect to ease of evaluation and stated both WTA for the good in possession and WTP for the alternative good. Then a random price was drawn and transactions were made. For the pen, the average WTP was €0.53 and the average WTA was €0.92; for the glass, the average WTP was €0.74 and the average WTA was €1.13. The differences between WTA and WTP were significant for both goods ($p < .01$), in agreement with the endowment effect. The effect of the product was not significant, and neither was the product × price interaction effect. Despite higher ratings of evaluability for the glass than for the pen, the nonsignificant interaction effect indicated that the size of the endowment effect was not affected by evaluability.

Hedonic versus functional goods. Since hedonic goods can be defined as providing affective and sensory experiences of aesthetic or sensory pleasure, fantasy, and fun (Hirschman and Holbrook 1982), these goods may lead to more psychological attachment than functional goods, whose consumption is more cognitively driven and goal-oriented and which accomplish a functional or practical task (Strahilevitz and Loewenstein 1998). Hence the willing-

Table 19.1

Willingness to Exchange Different Goods and Money (%)

	Hedonic goods	Functional goods	Money
Mobile laboratory: peppermints, pens	22	47	
Knetsch 1989: mugs, chocolate	10	11	
Knetsch 1995: mugs, pens, money		10	16
Dhar and Wertenbroch 2000: M&Ms, glue sticks	15	85	

ness to exchange may be lower for hedonic than functional goods. Further, since money is supposed to lead to even less psychological attachment, willingness to exchange money will be higher than for goods.

In the mobile laboratory we studied the endowment effect for a hedonic good (peppermint) versus a functional good (pen). It appeared that willingness to exchange the hedonic goods was lower than for the functional good. However, Knetsch (1989) found hardly any difference in willingness to exchange across the two types of good. Knetsch (1995) used goods versus goods and goods versus money. It appeared that money was exchanged more easily than goods, although the result was not significant. Dhar and Wertenbroch (2000) found a strong difference in choices for giving up M&Ms or glue sticks when individuals were endowed with both goods. The willingness to give up the glue stick was far greater than for the M&Ms.[4] The size of the endowment effect for different types of goods is shown in Table 19.1.

If we accept WTA as the measure of willingness to exchange, endowment effects appear even larger than in goods exchanges. In a class of fifteen Ph.D. students, participants could buy or receive a box of chocolates and a flashlight (retail price €3.50 each). The average WTP was €6.15 for chocolates and €3.10 for flashlights, while the average WTA was €2.55 and €0.69, respectively. The endowment effect was significant ($p < .01$), and chocolates were valued higher than flashlights ($p < .01$) despite equal retail prices. However, no significant interaction effect occurred, so the endowment effect appeared about equally strong for chocolates as for flashlights.

Endowment effect for imagined transactions. The endowment effect also worked when students just imagined that they could acquire or relinquish an object. We used an elementary economics class of fifty students, half of whom were told that a plant would be given to one of them as a gift. Each member of this group had to state the minimum WTA in case the plant was given to him or her. The other half of the class stated their maximum WTP for the plant in order to buy it from the owner of the plant. Then one student from each group was drawn randomly. If WTP of the buyer exceeded the WTA of the new owner of the plant, the plant would change hands; otherwise the owner took the plant home.

The average WTP for the plant was €2.55; the average WTA was €4.31 ($p < .05$), thus showing the endowment effect for an object that was not really owned and which could be obtained with only a very small chance.

PRISONER'S DILEMMA

The prisoner's dilemma is a cooperation game frequently studied in the social sciences. The game deals with a district attorney who wants two prisoners to confess their joint crime. The district attorney tells each prisoner: "If you both confess, you will each go into jail for three

Table 19.2

Payoff Table of a Prisoner's Dilemma

		Prisoner II	
		Deny	Confess
Prisoner I	Deny	(–1, –1)	(–10, 0)
	Confess	(0, –10)	(–3, –3)

years. If you both deny, you will each go into jail for one year. If only you confess, you will be free and the other person will get ten years in prison." Communication between the prisoners is not allowed. Regardless of the other prisoner's behavior, it is always better to confess, as is shown in the payoff matrix in Table 19.2. The matrix shows the outcomes of the prisoners' choice combinations (left entries between parentheses for prisoner I, right entries for prisoner II). For "confess" the outcomes for prisoner I (0 or –3) are better than for "deny" (–1, –10, respectively), and vice versa for prisoner II. This makes both prisoners confess, leading to a worse outcome than under mutual denial. Denial is indicated as the cooperative strategy, confession as the defective strategy.

By systematically varying the payoffs, different motives for playing the game can be investigated. For example, by changing player I's payoffs while keeping player II's payoff constant, the effect of player I's individualistic motive can be shown. By changing player II's payoff while keeping constant player I's payoff, player I's altruistic motive can be shown. Competition can be shown by changing the payoff difference between players I and II. Charness and Rabin (2002) showed the existence of both cooperative motives and a motive for avoiding very low outcomes of the other player.

The prisoner's dilemma can be extended to multiple players in different ways. Variations of prisoner's dilemma games and free rider problems in public economics can be found in, for example, Kagel and Roth 1995. Dawes (1980) considers the "take some game," in which each player can either choose to receive $1 (cooperative choice) or choose to receive $3, in which case everyone is fined $1 for that choice (defective choice). If everyone cooperates, each player will receive $1. If everyone defects, each player will receive $3 minus $1 times the number of players.

In the "give some game" (Dawes 1980) each player may choose either to keep $8 received from the experimenter (defective choice) or give $3 from the experimenter to each of the players (cooperative choice). If everyone cooperates, each player receives $3 times the number of players. If everyone defects, each player will receive $8.

Another variation of the multiple-player prisoner's dilemma game that was used in the mobile laboratory on economics is the "disappearing lottery prize," taken from Hofstaedter 1983 and Bazerman 1998. In this game, the pupils could submit up to six lottery tickets to win a prize. However, the prize was divided by the total amount of tickets that were submitted. The cooperative choice of the players is to submit only one ticket each. In this case, the prize is maximal while the chances of winning are equal for all players. However, the temptation of defective choice is strong. If one player submits six lottery tickets, the chance of winning is six times the chance of winning under cooperative choice. However, if everyone plays six lottery tickets, chances of winning are equal but the prize is six times as small as under cooperative choice. We wanted to test two hypotheses: (1) both sexes behave equally cooperatively and (2) both sexes expect the other sex to behave equally cooperatively. The results are given below.

Disappearing Lottery Prize Experiment

The experiment is best explained by following its instructions:

> *The next experiment is not just fun, it will also be used for scientific research. During this research you are not allowed to talk or discuss. If this happens we have to stop the experiment and there will be no winner. We will play two rounds. In the meantime: be quiet. No questions can be raised during the experiment. It will take approximately 5 minutes.*
>
> *In this classroom there are N pupils. Therefore the maximum amount to win will be N × 5 euros. Each pupil in the classroom will receive a sheet of paper. On this sheet you can indicate how many lottery tickets you want to play, minimum 0 and maximum 6. All sheets will be collected and one of the participating tickets will be the winner. The winning prize depends on the total number of participating lottery tickets. Make sure nobody sees how many tickets you play.*
>
> *In the instruction you will see an example: Suppose a classroom with four pupils. A plays 3 tickets, B plays 2 tickets, C plays 0 tickets, and D plays 3 tickets. Now the average number of playing tickets is 2 and the total amount to win is 8 euros. The ones who play the highest number of tickets have the biggest chance of winning; however, the higher the total number of lottery tickets played by the whole classroom, the lower the prize. In this example the maximum prize could have been 4 × 5 = 20 euros, but the actual prize will be 20 ÷ 8 = 2.50 euros. Now we will distribute the sheets of paper. The exact procedure can be read on the sheets, as a reminder.*

So the actual prize equaled the maximum possible prize divided by the number of participating tickets. After the lottery sheets from the first round were collected, a second round was played immediately thereafter. In this round boys and girls played as two subgroups, each for its own prize. The prize for the winner among the boys depended upon the total amount of lottery tickets played by the boys, and likewise for the girls. This experimental design was employed to study differences in behavior of the sexes when they played against their own sex or against the other sex.

Apart from choosing the amount of lottery tickets to participate in the lottery, the pupils also had to predict the expected average amount of lottery tickets played by the whole group (round 1), played by the boys (round 2), and played by the girls (round 2). In total, five items were gathered for each participant: the number of lottery tickets played in round 1, the expectation about average behavior in round 1, the number of lottery tickets played in round 2, the expectation for boys' behavior in round 2, and the expectation for girls' behavior in round 2.

One additional remark that was written on the answer sheet and not given in the general instructions was what would happen if everyone played 0 tickets. This situation, although being a hypothetical case, had to be addressed in order to avoid any misunderstanding, and also to prevent giving an alibi for not playing cooperatively. The solution to this was: "If everyone plays with zero lottery tickets, one pupil will be randomly drawn and will receive 10 euros. So the bonus for (everyone) playing 0 is higher than for (everyone) playing 1 ticket, since in that case the prize would be 5 euros."

Results from the Disappearing Lottery Prize Experiment

The results from twenty schools visited during autumn 2003 were analyzed.[5] The experiment was conducted in the highest classes of the secondary schools that prepare for university. The

Table 19.3

Distribution of the Number of Lottery Tickets Played

	Boys (N = 148)	Girls (N = 136)	Total (N = 284)
Round 1			
0 tickets	2	1	3
1 ticket	31	46	77
2 tickets	41	65	106
3 tickets	28	19	47
4 tickets	7	1	8
5 tickets	1	0	1
6 tickets	38	4	42
Mean	3.09 (1.91)[a]	1.92 (1.00)	2.53 (1.65)
Round 2			
0 tickets	1	0	1
1 ticket	29	53	82
2 tickets	38	50	88
3 tickets	25	26	51
4 tickets	10	0	10
5 tickets	4	2	6
6 tickets	41	5	46
Mean	3.28 (1.93)	1.99 (1.15)	2.67 (1.72)

[a]The figure in parentheses is the standard deviation.

age of most pupils was approximately 17 years. The average size of the classes was 22.1 pupils; the smallest class had 14 and the largest class 34. The total number of pupils was 442: 235 boys and 207 girls.

A statistical analysis was conducted for the respondents who filled out all five items: the number of lottery tickets played in round 1, the expectation about average behavior in round 1, the number of lottery tickets played in round 2, the expectation for boys' behavior in round 2, and the expectation for girls' behavior in round 2. Quite often expectations were not filled out, probably because these questions were stated at the end of the sheet. For 284 pupils, however, 148 boys and 136 girls, a complete record was obtained.

The distribution of number of lottery tickets played is shown in Table 19.3. Most of the time 1, 2, or 3 tickets were played; 0, 4, or 5 tickets were played rarely. Boys quite often chose to play all 6 tickets; girls did this seldom. On average boys played slightly more than 3 tickets, while girls played slightly less than 2 tickets. The difference (1.18) was highly significant ($p < .001$) (see Table 19.4). The difference in behavior between the two rounds was small. Both the boys' and the girls' subgroups played slightly more tickets, but the difference with the first round was not statistically significant.

So there was in fact a large difference in behavior between boys and girls. The next question to address is whether it was foreseen by the participants. Participants indeed predicted more tickets played by boys than by girls. They expected the boys to play on average approximately 0.5 tickets more than the girls. Since the difference in reality was larger (more than 1.0), they underestimated the level of difference in behavior. Although both sexes gave accurate predictions of their own group behavior, boys underestimated the level of cooperative behavior of girls, and girls overestimated the level of cooperative behavior of boys.

Table 19.4

Average Number of Lottery Tickets Played

	Boys (N = 148)	Girls (N = 136)	Difference Boys – Girls
Round 1	3.09 (0.16)[a]	1.92 (0.09)	1.18[b] (0.18)
Round 2	3.28 (0.16)	1.99 (0.10)	1.29[b] (0.19)
Difference (Round 2 – Round 1)	0.19 (0.10)	0.07 (0.09)	

[a] The number in parentheses is the standard error of the mean.
[b] $p < .001$.

Finally, we tested whether participants played tactically. If someone plays with more tickets than he or she predicts for the whole group, that person is deliberately trying to take advantage of the cooperative behavior of others for his or her own benefit. If someone plays with less tickets than he or she predicts for the whole group, that person is deliberately playing for the benefit of the group, despite his or her self-interest. A variable, *tact,* defined as the number of lottery tickets played minus the average number of lottery tickets predicted for the whole (sub)group, was taken as a measure of tactical playing. Boys obtained scores higher than zero, indicating that they played tactically for their own self-interest (round 1). Girls obtained scores lower than zero, indicating tactical play for the group interest (round 1). In round 2, boys continued this behavior in their own subgroup, whereas girls played the same number of lottery tickets as they expected for the whole subgroup.

DUAL PROCESSING AND EVALUATION OF GOODS

There is a long tradition of thinkers ranging from Aristotle to Freud and on to modern-day writers such as Epstein (1973) and Sloman (1996) who have argued for two (or more) systems involved in thought. For example, Epstein (Denes-Raj and Epstein 1994) proposes that there are two interactive parallel systems of cognition: rational and experiential. The former is a verbally mediated and primarily conscious analytic system that functions by a person's understanding of logic and evidence. The experiential system operates in an automatic, associational, and holistic manner. While generally adaptive in natural situations, it is often maladaptive in unnatural situations that cannot be resolved on the basis of generalizations from past experience but instead require logical analysis and an understanding of abstract relations.

One dual-process model that is of particular relevance to understanding economic behavior is the one proposed by Mittal (1988). The relevance of Mittal's model stems from two sources. First, it is specifically a model of consumer choice: many of the recent dual-process models are concerned with social-psychological processes more generally, for example, attitude change or social perception (see Chaiken and Trope 1999). Second, it directly addresses the relationship between processing mode and type of good, which we discussed above in relation to the endowment effect. In Mittal's model, choices can be made by means of either an *information processing mode* (IPM) or an *affective choice mode* (ACM). In IPM, product attributes are evaluated, then combined into an overall choice by means of some cognitive algebra. In contrast, in ACM, a property of the product as a whole, such as its hedonic impact or social image, determines choice. It is proposed that products can be purely functional or have both utilitarian and expressive properties to varying degrees. The *expressiveness* of a product refers to its ability to fulfill various psychosocial goals such as pleasing the senses and bolstering the ego. The more expressive a product is, the more affective processing there is. In addition, products can be more or less *involving;* in other words, there can be a greater or lesser motivation for the consumer to make the right choice (see, e.g., Mitchell 1981; Park and Mittal 1985). The more involving the product, the more information processing will take place. Finally, the reasons for choices made by affective processing are much harder to express than those for choices made by information processing.

Mittal (1988, 1994) sought empirical support for his model through two experiments. In each study participants were asked to make choices between products, then complete a questionnaire designed to assess the amount of involvement with the chosen product and the perceived expressiveness of that product, as well as the degree to which information processing and affective choice modes were used in product selection. Structural equation modeling was then used to test the relations between constructs predicted by the model. The results provide broad support for Mittal's model in that there is confirmation of the major constructs—involvement, expressiveness, ACM, and IPM—and for the proposal that ACM is positively related to expressiveness and IPM is positively related to involvement. Unfortunately, some of the predicted paths are quite weak or insignificant, some nonpredicted paths are significant, the overall model fits are far from perfect, and ACM is reported as being poorly measured. Further, the first study posed hypothetical choices (in the form of scenarios), so the external validity of the results of this study is questionable (see the methodological considerations below). Although the second study overcomes this problem by giving participants a real choice between products, which they were then subsequently allowed to keep, there is a confounding of the information given about each product and its anticipated expressiveness. For instance, products expected to be high in expressiveness were described to the participants in terms of their social and hedonic properties, whereas products low in expressiveness were described in terms of their functional and utilitarian properties. It is therefore not possible to determine whether the pattern of responses given by participants was due to their perceptions of the expressiveness of products, due to the product information, or both.

Classroom Experiments on Dual Processing

Two of us have conducted a number of classroom studies investigating the effects of processing mode. Our starting point was to conduct a partial replication of Mittal's experiments in which we tried to rectify the methodological problems outlined above (i.e., we gave participants real choices, removed the confound between product descriptions and their expressiveness, and included more measures of ACM). Despite these changes, we obtained equivocal results similar to Mittal's. We therefore decided to try a new experimental method whereby we attempted to directly manipulate processing mode, then examine the effects of this manipulation on product valuation within a choice setup. In particular, we measured the size of the endowment effect for a chosen product when the choice was made under either IPM or ACM. For reasons similar to those given above with regard to hedonic versus functional goods, we hypothesized that the endowment effect would be greater when choice of product was made under ACM than under IPM.

One hundred forty-five first-year economics undergraduates from Erasmus University Rotterdam took part during their normal classes. All participants took away with them from the experiment either a pen worth about €1.00 or a small amount of money (between €0.25 and €2.50). A two-by-two factorial design of processing by task was employed, which resulted in four independent groups of participants of approximately equal size: IPM-WTP, IPM-WTA, ACM-WTP, ACM-WTA. The products used in this experiment were two types of pen. The processing manipulation in the IPM condition was a list of ten features whereby the two pens could be differentiated: color, form, materials, form of clip, nib type, nib protection, ink color, ink permanence, writing comfort, and weight. The instructions were to rate each attribute for each pen on a five-point scale where 1 was an extremely negative evaluation and 5 a strongly positive evaluation. In the ACM condition twenty-one adjectives were provided that might be used to describe the pens in a global, emotional way, the approximate English equivalents being *eye-catching, boring, pretty, exciting, practical, nice, mundane, chic, functional, different, amusing, cheap, attractive, novel, "me," comfortable, unusual, "not me," quality, ugly,* and *ordinary*. The participants were instructed to select as many of these as they thought were appropriate to describe each pen (with a minimum of one adjective for each pen). In both conditions the participants were asked finally to select which of the two pens they preferred.

Each of the four experimental groups was a different class of a first-year course in marketing. The classes were composed in a random way at the beginning of the year. Each class was verbally informed as a whole that they were being asked to participate in a study of consumer choice and that this would involve them making evaluations of two different brands of pen. They were also urged that they should attempt to make these evaluations on an individual basis. Next the products were distributed along with the processing manipulation: everyone within a group received the same processing manipulation (i.e., either IPM or ACM), and each group was randomly split into acquisition and forfeiture subgroups. After everyone had made his or her evaluation and indicated a preference, either all the products were collected (WTP condition) or the nonchosen product was collected (WTA condition). As a manipulation check, participants were next asked to complete a questionnaire designed to measure the amount of ACM and IPM. After everyone had returned the questionnaire they were asked to state either the sum of money they would be willing to accept in return for giving up their chosen pen (WTA) or the amount of money that they would need to receive such that it would be preferred to receiving their chosen pen (WTP). This question and the random price mechanism (see below) that was to be used in order to elicit true valuations were explained to them both verbally and in writing on the response sheet. After everyone had indicated a price, one of the participants was invited to draw a chip out of a group upon which were written prices from €0.25 to €2.50 in 25-cent increments. Finally money or pens were awarded to the participants on the basis of the result of the draw: those stating WTA prices equal to or lower than the drawn price got the drawn amount of money, otherwise they kept the pen, whereas those stating WTP prices *higher* than the drawn price received the drawn amount, otherwise they kept the pen.

Unfortunately, the hypothesis that the endowment effect would be greater for those evaluating products under ACM than IPM was not borne out. However, there was a significant main effect of processing ($F(1,141) = 5.918$, $p = .016$) such that people on average indicated that they would pay €0.19 more for the pens in the ACM condition than the IPM condition. It seems likely that this difference in valuation of the products due to processing mode led to a ceiling effect under ACM, which meant that an endowment effect was not observed for this type of processing (i.e., participants could not value these products any more in the WTA condition than WTP since their WTP amounts were already at a maximum price for these products).

Another experiment with classes of Ph.D. and undergraduate students produced similar results. Here processing mode was manipulated by letting the participants evaluate a product either on scales consisting of affective adjectives or on scales concerning attributes of the product. It was intended that affective scales would elicit ACM, whereas the rating scales would elicit IPM. The product evaluated was a candle lamp. After evaluating the candle lamp, students were required to state their WTP. Then one student was selected at random and for this student the candle lamp was auctioned by using the random price mechanism. The average WTP under ACM was €1.69, whereas under IPM processing it was only €0.95 ($p < .05$).

The effect of processing mode may be subtle. The manipulation we used can easily fail if the participants have time to evaluate the product in a different way after completing the questionnaire. In a large class of Danish students, WTP did not differ across conditions, possibly because we waited until every student had completed the questionnaire. By that time, the students might have been thinking about the product in different ways, thus destroying the experimental manipulation. To avoid different ways of thinking after completing the questionnaire, either we walked around in class to present the students with WTP questions immediately after they completed the questionnaire or the students were given the WTP question in an envelope that was opened immediately after completing the questionnaire.

In some more recent experiments we have abandoned the use of the discrepancy between WTP and WTA as a measure of the endowment effect in favor of the swapping paradigm used by Knetsch and Sinden (1984), mentioned above. In one experiment, 102 high school students (ages fifteen to seventeen) and 66 undergraduate students took part. Again a two-by-two design of processing by choice was employed, which resulted in four groups of participants: ACM-Retain, ACM-Switch, IPM-Retain, IPM-Switch. There were eighty-five participants in the IPM group and eighty-three in the ACM: whether participants were in the Retain or Switch group was their own decision and was, in fact, our dependent variable. The products used in the experiment were two different kinds of confectionary: a bag of Autodrop licorice and a bag of Chupa Chups lollipops. These two products cost about the same amount, €1.42 and €1.19, respectively. Moreover, according to a supermarket manager, they were equally popular among the teenagers in the sample.

The processing manipulation consisted of a list of ten product attributes/features of either Autodrop or Chupa Chups to be evaluated on five-point bipolar scales. In the IPM condition, participants were asked to rate functional attributes of each product separately, for example, size, weight, energy, and shelf life. In the ACM condition the products were evaluated on hedonic attributes, for example, taste, brand quality, attractiveness, and ability to satisfy. Each of the four groups at each of the two locations was verbally informed as a whole that they were being asked to participate in a study on consumer behavior. This would involve evaluating the bags of Autodrop and Chupa Chups. The participants were urged to make these evaluations on an individual basis. Next either the Autodrop or the Chupa Chups were distributed along with the processing manipulation—either ACM or IPM—and an envelope to be opened directly after finishing the questionnaire. The envelope asked participants whether they wanted to keep the product they had been given or switch to the other product. The participants could keep or acquire their preferred product. After having made their choice to retain or to switch, participants were asked to estimate the prices of the two products.

The results were in line with the hypothesis that the endowment effect would be stronger for ACM than IPM. Of the 46 participants in the ACM group endowed with Autodrop, only two (4 percent) switched to Chupa Chups. In contrast, in the IPM group, 15 out of 40 (37.5 percent) traded in the licorice for the lollipops. Of the 37 participants in the ACM group endowed with Chupa Chups, just five (13.5 percent) participants switched to Autodrop, whereas in the IPM group 13 out of 45 (29 percent) made the switch. A probit analysis shows that the endowment effect was statistically significant (one-tailed $p < .01$). The expected interaction of processing mode by choice was also significant (one-tailed $p < .05$). There was therefore an endowment effect in both conditions, and it was greater for ACM, as predicted.

SUBJECTIVE DISCOUNTING

Discounting refers to valuing present outcomes higher than equal future outcomes (Fishburn and Rubinstein 1982). Usually in economics exponential discounting is assumed, implying an equal discounting rate in each future period. For example, if someone values $100 today as equal to $110 in one year (discount rate of 10 percent), then according to these assumptions $121 in two years will also be valued equally. However, the standard assumption is not realistic in the area of consumer behavior. It appears that consumers frequently use higher discount rates in the near future and lower discount rates in the distant future (e.g., Thaler 1981). An alternative discounting function has therefore been proposed (Loewenstein and Prelec 1992; Ahlbrecht and Weber 1995) that reflects the idea of changing discount rates over time. This is called hyperbolic discounting.

It is quite easy to demonstrate hyperbolic discounting in class, and we have reported several experiments elsewhere (Antonides and Wunderink 2001). For example, one may ask students for future amounts they are willing to accept in order to forgo $1.50 payable on the same day. Future dates may vary between one week and one year. Hyperbolic discounting will be evident from the data by decreasing amounts per time period for periods of one week ($3), two weeks ($4.50, or $2.25 per week), ten weeks ($8, or $0.80 per week), and fifty weeks ($30, or $0.60 per week).

Hyperbolic discounting may lead to preference reversals. For example, viewed from today, an amount of $1,000 in two years may be preferred to an amount of $800 in one year because both outcomes occur in the future. However, after one year, the situation is receiving $800 the same day or receiving $1,000 in one year. At that moment, cashing in the $800 may be more likely because it has become a present outcome. Likewise, a pregnant woman who is asked six months before the event may prefer delivering the baby naturally to delivery under anesthesia, because natural delivery has larger long-term benefits than the short-term benefits of anesthesia. However, when the labor starts, she may prefer the immediate, smaller benefits of anesthesia (Christensen-Szalanski 1984).

Also, different types of good may be associated with different discount functions. For example, for healthful items such as fruit the benefits may be perceived as higher in the long run than for less healthful snacks. Hence, one may prefer an apple to a less healthful snack to be consumed in one week (Read and Van Leeuwen 1998). However, after one week consumption can take place immediately and many people change their preference in favor of the less healthful snack.

Another distinction related to time preference is between hedonic and utilitarian goods. Gattig (2002) showed that time preference is higher for hedonic items (e.g., CD or television set) than for functional items (e.g., computer diskette or washing machine). Hence, the participants in his studies preferred advancing the delivery of hedonic goods to advancing the delivery of utilitarian goods. However, when monetary compensation was given to postpone the delivery of the goods, no significant differences were found between advancement choices for hedonic and utilitarian goods. It seems that adding monetary aspects to the choices made people decide more rationally.

EXPERIMENTS ON THE EFFECT OF SITUATION ON CONSUMER BEHAVIOR

The effect of situation in consumer judgment has become of interest to marketers because volatile consumer behavior can only partly be explained on the basis of personal characteristics, income, attitudes, and social norms. Situational effects can be demonstrated easily with a questionnaire asking for preferences for goods in different situations. For example, it can easily be shown that an ice cream is preferred to an apple on a hot beach, whereas the reverse is true after lunch. Likewise, the probability of consumption differs across social situations (Belk 1974; Lutz and Kakkar 1975).

Also, the situation may affect preferences for the same good. Thaler (1980) asked his students for their willingness to pay for a beer under two different conditions: when the beer was purchased from a fancy resort hotel or when the beer was purchased at a run-down grocery store. In both cases, the beer was to be consumed at the beach. WTP appeared to be higher when the beer

Table 19.5

Classified Reactions to Receiving Each of the Social Resources

Student's reaction	Money	Product	Service	Love	Status	Information
Other's gift						
Money	11	4	20	4	3	0
Product	1	19	7	0	1	0
Service	0	8	22	0	0	5
Love	0	7	13	18	1	0
Status	0	0	0	1	15	0
Information	0	14	6	1	3	2

was to be purchased at the hotel than at the grocery store. The different WTP could only be due to the nature of the point of sale. Thaler (1980) assumed the existence of two different kinds of utilities: acquisition and transaction utility. Acquisition utility is derived from the product itself, whereas transaction utility is derived from the purchase environment. Although acquisition utility was the same for the beer from the hotel and the beer from the grocery store, the transaction utilities differed across the two points of sale.

Framing is just another instance of a situational effect. Framing refers to a particular description of a good that may be considered as an information situation. For example, Levin and Gaeth (1988) found that preference for a steak that was "50 percent fat free" was higher than for a steak that "contained 50 percent fat."

The type of item appears to influence its suitability in mutual exchange for another item, which is considered another situational effect. Foa (1971) developed a theory explaining the likelihood of exchange for different "resources," including goods, services, money, information, status, and love. In the original experiment participants were asked which of a pair of resources was most appropriate in return for a particular resource given to another person. For example, participants were asked: "What is the proper compensation you wish to receive in exchange for giving information to a person? Money or a good?" Since there are fifteen possible combinations of the six resources, participants were presented with fifteen pairs of choices. Information was most likely to be exchanged for status and money and less likely to be exchanged for love and services. This procedure was repeated six times for each resource, amounting to ninety pairs of choices. It appeared that personalized resources such as love, status, and services generally were not preferred in exchange for general resources such as cash, information, and goods. Also, abstract resources, such as status and information, were not preferred in exchange for concrete resources such as goods and services.

Foa's idea can be replicated rather easily in class. First we gave students an overview of the kind of resources that one may use in social exchange. Then we asked students about the most appropriate reaction to each of six situations: (1) someone who gave you money (500) when you needed it, (2) someone who gave you a product to be used in your room, (3) someone who helped you clean your room for one day, (4) someone who gave you emotional support when you had a difficult time, (5) someone who praised you about your good exam results in the presence of other people, and (6) someone who gave you information about a job vacancy (you got the job). Students formulated their answers themselves (in contrast with Foa's original research), which were then coded into the resource categories. An overview of the answers is shown in Table 19.5. The results are by and large in agreement with Foa's theory: most reactions fell within the same category as the resource that was given (numbers on the diagonal of the matrix in Table 19.5).

Also common were reactions including resources that in Foa's theory were close to the resource that was given (numbers around the diagonal). Other reactions were less common (numbers away from the diagonal). Services appear as a quite popular resource given in exchange for another person's gift.

METHODOLOGICAL CONSIDERATIONS

Incentives

One important way in which the experimental methods of psychologists and economists differ is in the use of incentives. Many economists strongly believe (see, e.g., Binmore 1987; Hertwig and Ortmann 2001) that experimental participants must be given large external incentives that are performance-related if they are to be adequately motivated to give responses with external validity (i.e., that will be generalizable to situations outside the laboratory). Meanwhile, psychologists tend to regard participants as being motivated by many factors other than external financial reward, which renders financial incentives either at best unnecessary or at worst counterproductive (see, e.g., Loewenstein 1999; Rakow 2001). For example, Roth (2001) points out that the endowment effect never would have been observed if only the valuation of monetary amounts had been investigated. This debate remains to be resolved empirically, although at least one metastudy suggests that providing financial incentives does not have any significant effect on the reliability of data, which is an indication that the effort made by participants is not necessarily contingent on external reward (Camerer and Hogarth 1999). Whatever the eventual outcome of this debate, by the nature of the field of inquiry, there will still be many classroom experiments in behavioral economics that require some money or goods to pass from the experimenter(s) to the participant(s): for instance, participants have to be endowed with a good in endowment-effect experiments and should receive a payout commensurate with their performance in a prisoner's dilemma game.

In an ideal world there would be no issue: classroom experimenters would be able to ensure external validity by giving sizable incentives to participants in all cases where it was deemed advantageous to do so. Further, they would always use products or sums of money that ensured high involvement (i.e., the degree to which a participant feels it is important to give a correct or truthful response; see the discussion of involvement above in relation to dual processing for more details). Unfortunately, in the real world, classroom experimenters will usually find themselves funding their experiments out of very restricted budgets (or, indeed, their own pockets; the same arguments apply to projects run by students, but to an even greater degree). Although a certain amount of expenditure of this sort might be considered worthwhile on the basis that it both provides potentially useful pilot data and is a valuable teaching tool, ways of minimizing one's expenditure as an experimenter are, we are sure, to be welcomed. We will therefore now briefly provide some suggestions regarding this aspect of classroom experimentation.

In some cases it may be possible to get good results without any cost at all to the experimenter. For example, we have managed to obtain significant effects of framing, mental accounting, time discounting, and satiation, as well as sizable (and statistically significant) reversals of preference, money illusions, overconfidence, sunk costs, and certainty effects, using purely hypothetical situations. Although the use or otherwise of hypothetical rather than real situations is a hotly debated issue in its own right (e.g., Roth 1995) that we do not wish to get involved in here, we suggest that as far as the classroom experiment goes, the use of hypothetical situations is perfectly justifiable if it is solely for demonstration purposes. However, one would be advised to stick to phenomena with relatively large and frequently replicated effects, such as those listed above.

In other instances, products used or the amount of financial incentive given can be fairly small and still produce significant effects. For example, strong endowment effects have been obtained with inexpensive products such as chocolate bars and coffee mugs, while small sums of money can be sufficient to produce the expected results in experimental games such as the ultimatum bargaining game. Where larger inducements are required—for instance, where effects might be quite small—then alternative procedures exist. A common technique is to use some random allocation of a subset of the participants to the prizes: this may be done by giving participants raffle tickets as payment, or by selecting one or more winners of a prize by drawing from a hat (a variation on this latter procedure is that these winners are then rewarded on the basis of performance on the experimental task). To give a couple of specific examples: one of the authors has had a student endow participants with ten raffle tickets each to win a (relatively) expensive product A, and later give them the chance to swap some or all of their tickets for raffle tickets for a chance to win an equally expensive product B. At the end of the experiment, all tickets are put into a hat and two tickets are drawn, one for each product. The winners are notified by e-mail. The number of tickets retained for product A (or B; whichever is selected is arbitrary) is a measure of the endowment effect and, as a measure, has the advantage over the frequencies obtained by the usual swapping method (see above) in terms of the range and power of statistical analyses that can be applied. It also has an advantage over the random price mechanism (see below) in terms of transparency to the participant (it also seems to work, but it should be noted that there may be some disadvantages, for instance, in the interpretation of the phenomenon that is being measured by this procedure, i.e., is it really the endowment effect?). Another example is that of the ultimatum bargaining game. Here one of us has randomly allocated students to roles, paired them up, then asked them to make their allocations with the understanding that one of the pairs would be selected, also at random, to make the transaction for real.

An alternative approach that can be used in experiments with multiple trials—such as an iterated prisoner's dilemma or market entry game—is to pick one or more trials at random and allocate the rewards according to performance on that particular trial. Thus in an iterated ultimatum bargaining game with six trials, one of the trials can be chosen to be played for real using dice. If in the selected trial the proposer had offered only 10 percent of the stake and the receiver had refused this proposal, then neither player would receive any money (even if proposals had been accepted in all the other five trials).

Both these procedures are variations on what is known as the random lottery incentive scheme. Although there are some who argue against the use of such schemes (e.g., Holt 1986), these criticisms can be largely ignored if one is simply running an experiment for pedagogical reasons. There are also good counterarguments to the criticisms (see Cubitt, Starmer, and Sugden 1998), and since this procedure is rather widespread in the literature, a paper almost certainly would not be barred from publication for employing it.

A further way of minimizing costs is by having a small sample size. In classroom experiments one usually has little control over sample size, and often the size is not optimal (i.e., either too big or too small) for one's purposes, an issue we will come to in a moment. In the United Kingdom at least, classroom experiments in economic psychology or behavioral economics will most commonly be conducted with undergraduates in their final year or with postgraduates. In either case, the classes will be rather small, and it will be not expense so much as experimental power that will be the greatest and most common problem. When sample sizes are small, power can be increased by using repeated-measures designs. In the extreme, taking a large number of measurements (or fewer but very rich measurements) from participants can allow the sample size to be reduced to one, as in psychophysical experiments (or in case studies). It follows, then, that where one does have control over sample size but one's budget is tight, expense may be kept down by using a small number of participants in a repeated-measures design.

Repeated-measures designs cannot, however, always be used because of learning and other carryover effects, or contamination, from one condition to another. As a specific example, let us consider an experiment to test whether choosing the product one wishes to be endowed with leads to a stronger endowment effect than when one is given no choice, but that this only works when the products are evaluated under IPM, not ACM (note that this is a hypothesis constructed for illustrative purposes only). This proposal could be operationalized by presenting participants with two products, getting judges to evaluate these products under either IPM or ACM (see above), then either giving them one of the products at random or allowing them to choose which one to keep. It is obviously going to be difficult to manipulate choice versus no choice within subjects, as the expectations from the first trial are going to be carried over to the second trial (i.e., the expectation being that if one chooses a product the first time one will choose again the second time, and so on; if one attempts to dispel these expectations through instructions, then the participants may be disgruntled), and these expectations may interfere with the evaluation (processing) of the second pair of products. In contrast, there is no particular reason to believe that the type of processing evoked in the first trial will be carried over to the second trial (as we have already seen, the effects of the processing manipulation appear rather short-lived), so this could potentially be manipulated within subjects. However, even if one believes that there should not be any carryover effects in one's experiment, it is necessary to counterbalance the order of presentation of the within-subjects conditions (in the above example, half the participants should be asked to evaluate the products analytically first and holistically second, the other half holistically first and analytically second). One can then check that there is no difference between the two orderings to ensure no (or negligible) carryover effects.

One problem with providing incentives in classroom experiments that one probably would not anticipate is that student participants can be reluctant to accept the prizes or sums of money offered, or to take them seriously. This can occur for at least three reasons. First, incentives that are very small might be rejected purely on the basis that they are too trivial. Second, and more commonly, incentives are rejected because they are perceived as both too trivial to receive individually and rather costly to the experimenter in the aggregate (there can also be an element of embarrassment about receiving "gifts" from the teacher). Third, if the prizes or monetary amounts are rather large, then there may be disbelief that they will actually be awarded. Obviously there is a problem if incentives are not wanted in some way or are disbelieved, because then they are not really acting as incentives. If one is intending to run a series of experiments with the same class, then one can "train" students in the acceptance of incentives. The first class where incentives are used is therefore something of a loss leader: the students get used to the idea of incentives being awarded, with the data collected being of little use. Henceforth there should be no significant problems so long as the experimenter is conscientious about giving out the rewards as stated, even if they are protested against by the recipients. Conscientiousness on the part of the experimenter is very important in order to build trust and positive reputational effects, an issue we will return to later. It should additionally be noted that some rewards are more acceptable to students than others, with chocolate bars being a generally safe bet in terms of acceptability (although there will always be someone who is not that interested in chocolate; a savory snack such as potato chips can be used as a complementary alternative, and the two together will generally cater to everyone's tastes).

A last brief comment on monetary incentives is that there may also be religious or cultural objections to their use, so some care should be taken if one is conducting experiments in a country of which one is not a native, or with culturally heterogeneous participants. For example, many Muslims find gambling unacceptable, so using monetary payoffs in experimental setups that may be interpreted as gambling situations (as is the case with many tests of economic axioms) may not be possible in Islamic countries or where there are a number of Muslim students in the class unless these tasks are heavily disguised.

Random Price Mechanism

An important methodological issue is how to elicit true preferences, or prices, from participants. This is related to the issue of incentives in that highly motivated and involved participants are likely to try to respond as accurately as possible; however, there are some other factors that also influence the reliability and validity of participants' responses. There is not the space here to discuss all the arguments regarding the obstacles to eliciting true preferences or all the means that have been devised to try to overcome these obstacles (see Bateman et al. 1997 for a discussion and comparison of several different elicitation techniques). Rather, we shall focus our attention on one procedure—the random price mechanism (Becker, DeGroot, and Marschak 1964)—that we have used frequently in our classroom experiments (and is therefore mentioned in several places above). When attempting to elicit prices—for instance, WTP and WTA—there may be a number of reasons participants may not give their true prices. The most serious of these reasons would be that the participants do not actually have a single true price to state, but, assuming for pragmatic reasons that they do, they may reasonably wish to reduce their WTP to a minimum that they think they can get away with, while similarly attempting to maximize their WTA. The random price mechanism essentially seeks to punish people who do not state their best price by potentially making them pay more than their true WTP or receive less than their true WTA. This is usually done by informing participants that a price (within a set range) will be drawn at random. If this random price is smaller than participants' stated WTP or greater than their WTA, then they pay (receive) the random price; otherwise they pay (receive) their stated price. If we take WTP as an example, imagine the quoted price range of a good is between 1 and 11. If one's true WTP is 6 but one states the minimum price of 1, then there is a high likelihood that the random price will be greater than one's stated price (thus one will have to pay it) and an even chance that it will be greater than one's true WTP. With the random price mechanism, there is thus a disincentive to state prices that are different from one's true price, for if one does, one stands a chance of paying more than one would ideally like to obtain a product or receiving less than one ideally wants in order to part with a product. A variation of the random price mechanism used for the endowment effect involves swapping an endowed product for money or trading potential ownership of a product for a sum of money. In this case, if one does not state one's true price, then one may end up parting with the product for too little money or "buying" it for too much.

This is all very well, but the random price mechanism can be difficult to administer in practice, particularly within the constraints of the classroom experiment. One problem is that by giving a range of values one provides an anchor for an estimation of the "objective" value of the product (i.e., the price one could obtain it for in the store). Thus in the above example, participants might be drawn toward the price of 6, the midpoint of the range provided. Another problem that some of us have experienced is that logic of the procedure can be difficult to explain to participants, and if they do not understand it fully, then the procedure is unlikely to achieve its desired effects. This problem can be alleviated by careful wording of the instructions, examples, and practice trials, but all this can be rather time-consuming: in general, it is a good idea to pilot any instructions and other materials with a similar group of participants, if at all practicable, in order to ensure as smooth a running of the classroom experiment as possible. A rather more mundane problem with the random price mechanism is that if one has to do the draw individually for many participants, it can also be very time-consuming: a solution to this is to select one student to draw the random price and apply it to everyone. A third, even more mundane problem, but real nonetheless, is that with the random price mechanism the experimenter is never sure how much he or she will have to pay out. This means one has to prepare for the worst in terms of the amount of money to hand out (and goods, if appropriate). This also applies to many other random lottery incentive schemes. The obvious solution to this problem would be to rig the random lottery, but we strongly advise against doing that (see the section on use of deception below).

Sample Size

Returning to the issue of sample size, if sample sizes are small, then there will be a number of experiments that simply are not possible. For example, some effects are fairly small and will only reliably show up in large samples (e.g., some of the effects of dual processes; see above). In other cases, data might be such that one needs lots of observations to reveal even moderate effects, as is the case, for example, with frequencies and much categorical (or nominal) data (i.e., one does not get much out of each participant, and the methods of analysis available are not very sensitive). Further, as already mentioned, complex experimental designs with several different conditions will obviously not be possible with small sample sizes, so it would not be possible to answer certain research questions.

Although big sample sizes are desirable from the viewpoint of experimental power, they also can be problematic to the classroom experimenter in logistic terms. In particular, a large class may be difficult to manage without assistance. For example, if the class has to be split up into two or more experimental groups requiring separate instructions, then this will be difficult to do unless one has help (although not necessarily impossible). Separate rooms might also be required if there are experimental manipulations that cannot be conducted on paper (e.g., a mood manipulation using film or music).

If one is attempting to run a study single-handed, then a questionnaire might be the best bet for a large group, although this is not totally without difficulties either. For example, if one has a class of more than a hundred, then it can take quite a few minutes to distribute a questionnaire. Our experience therefore suggests that one should not permit students to start answering the questionnaire as soon as they receive it because the first students to receive the survey may well have finished before the last students to receive it have even started. As a general rule, it is not good to have large numbers of students idle while others are busy, as it will be difficult to keep noise levels down and/or participants become bored and unmotivated. It can also be difficult to regain control of the class after the questionnaire is completed if many people are talking. One strategy to deal with this problem is to allow students to leave as soon as they have finished (and possibly reconvene later), although this too creates a disturbance and tends to reward those who are least diligent in their responses. Another strategy is simply to make sure that one provides enough tasks to keep everyone occupied during the available time, but with the least important tasks toward the end.

Another common problem with classroom experiments, and one that is exacerbated by large sample sizes, is the provision of feedback about the results. For pedagogical reasons it is desirable to get the results back as soon as possible; ideally the feedback should be given within the session where the data is collected. Three ways of dealing with this are to have an assistant calculate some preliminary results while one is doing something else with the students, calculate the results oneself during a break, and get the students themselves to calculate the results. In each case the calculations should be fairly simple (i.e., not require sophisticated analyses), for example, mean prices or counts of number of exchanges. Getting the students themselves to do some analysis can be a good ploy, as in addition to speeding things up and reducing one's workload it can help the students gain insight into the experiment. An alternative to doing the analysis on the spot is to do the classroom experiment some time in advance of the lecture when one wishes to make use of the results. This is obviously a less desirable option, since it reduces immediacy and continuity, but it may be necessary if some complex analysis is required; if the delay between data collection and feedback is fairly short (no more than a week), then any undesirable effects should not be great. An additional strategy for providing feedback is to do so via a Web page. This way detailed information about the rationale for the classroom experiment, the procedure, the results, and their interpretation can be provided. We recommend, where possible, providing rough results on the spot and then more detailed feedback on the Web as soon as possible thereafter.

One way of running complex classroom experiments single-handedly is to use computers. Networks of computers that can be used for teaching purposes are common today, and software for running experiments is widely available (e.g., ELSE G4 software for running experimental games [Tomlinson 2002]; the downloadable software package z-Tree from the Institute for Empirical Research in Economics, University of Zürich; the experiments available at http://veconlab.econ.virginia.edu/admin.htm). Computers can be made to do the hard work of coordinating the activities of many students simultaneously, collecting their responses, analyzing them, and providing rapid feedback. For example, computers are ideal for running experiments with repeated trials of varying types such as a market entry game with different market capacities varied randomly over trials. Computers are not, however, a panacea. First, it can take a great deal of effort to program the experiment: even with good software a number of iterations of development will be required, and there is also the time required to learn the software. Second, in our experience, computers and their software have an uncanny knack for failing at crucial moments, so a good amount of piloting is recommended. Third, there will be a limited number of machines in a classroom, which rules out most large classes unless students double up (which creates its own problems). Fourth, computer labs may not be the most conducive places for doing whatever else one wants to do with students, such as give a lecture or do group work.

Deception

We started this discussion of methodological issues in classroom experiments by examining one way in which psychologists and economists differ with regard to experimentation: the use of incentives. To conclude this section we want to briefly discuss another major methodological difference between the experiments of psychologists and economists, the use of deception. Economists are very strongly against the use of deception (see, e.g., Ortmann and Hertwig 1997), whereas psychologists, especially social psychologists, regard deception as an essential tool for the investigation of certain research questions (e.g., Baron 2001; Davis and Durham 2001; Goodie 2001; Hilton 2001). The economists argue that the use of deception leads to a breakdown in trust between experimenters and participants, which produces undesirable reputational effects for researchers. In other words, once participants learn that they are likely to be deceived by researchers, they no longer trust them: this results in, at best, attempts to divine the "true purpose" of the experiment, which can lead to an increase in error variance, a deliberate lack of cooperation, or even sabotage, which can destroy the validity of experimental findings. Psychologists' response to this is that deception of some sort is often necessary in order to conceal the true nature of the experiment and remove compliance and demand effects; they argue that the negative reputational effects can be removed by a thorough debriefing. As is the case with the use of incentives in experiments, the "truth" about the effects of the use of deception is far from clear. For one thing, there are several degrees of deception: is deception by omission to be regarded as negatively as deception by commission? If so, we would always have to inform our participants of our experimental hypotheses, which we doubt is what most economists have in mind (and see, e.g., McDaniel and Starmer 1989; Hey 1989). Also, the effects of deception are likely to be different for different people. Psychology students, for instance, tend to become poor participants as a result of cynicism arising out of overexposure to psychologists' methods, including deception. Other groups who may act as participants rather infrequently may well never learn to mistrust researchers. Returning to classroom experiments, if one makes experimentation a regular feature of one's class, then one's student participants are going to be very susceptible to reputational effects, so it is particularly important that one does not use deception. In addition, one should ensure that one's reputation as an experimenter is spotless in other respects too, for instance, promptly providing promised incentives and/or feedback.

CONCLUSION

We have presented our experiments as behavioral economic experiments. This implies that the behavior of the participants may be, and frequently is, different from what standard economic theory would predict. This comes as no surprise to psychologists and sociologists, as well as people from marketing and many other disciplines. However, to a number of economists the results of our experiments may be unacceptable for several reasons.

Since we considered classroom experiments, the behavior of participants was less under control than in economic laboratories. Hence the results may be influenced by random error, type of participants, social influences, and systematic error due to classroom settings, logistic difficulties, and lack of attention, among other things. Rather than viewing the lack of control as a disadvantage, we believe that the results indicate the robustness of the phenomena studied. Stated differently, when standard economic theory is considered the null hypothesis in our experiments, we believe that we make a correct decision by rejecting it.

Other reasons for rejecting results from economic classroom experiments are similar to those that have been mentioned in relation to behavioral economic research in general (Thaler 1986):

- The use of small incentives in experiments (however, see our discussion on incentives above).
- Negligibility of heterogeneous preferences in aggregate predictions (Musgrave 1981). However, many aggregate predictions, seem to be false and heterogeneous preferences may be systematically related to personal characteristics and contextual circumstances.
- Negligibility of unstable preferences in long-run predictions. However, unstable preferences may also be systematic (e.g., in the case of hyperbolic discounting).
- In practice, learning may lead to more rational behavior than in one-shot experiments. However, experiments on melioration (Herrnstein and Prelec 1991) and overconfidence (Fischhoff, Slovic, and Lichtenstein 1977; Barber and Odean 2000) show that learning may not prevent anomalous behavior. Furthermore, in many practical situations learning opportunities are absent.
- In practice, irrational behavior is weeded out because of arbitrage and competition. However, markets do not always eliminate error (see, for example, Odean 1998 on loss aversion in financial markets).

We believe that the standard economic model should not be abandoned but needs to be adapted by including insights from behavioral experiments. In this respect we agree with Thaler, who offered two false statements: "1. Rational models are useless. 2. All behavior is rational" (1986, S283).

NOTES

1. Such effects may also occur when the same class is divided into groups. One of the authors has run a classroom experiment on donations to charity organizations, several of which were introduced briefly in class. The class was divided into those sitting in the front and those sitting in the back. The two groups were given different anchors for their donations. The front group was given a high anchor by being asked: "Would you donate more or less than 10 euros?" The back group was given a low anchor (1 euro). Then all students donated money. Although we hypothesized that the high-anchored students would give more than the low-anchored students, the reverse result was obtained. Why? In the discussion afterward, it turned out that the front group consisted mostly of foreign students who were not familiar with the charity organizations mentioned in the introduction. For this reason, they donated less.
2. This variation was suggested by Daniel Read, who was also involved in carrying out the experiment.
3. Together with Alessandra Arcuri.
4. Formally, this is not a test of the endowment effect since nothing had to be given up in order to acquire something.
5. We thank Tjeerd van den Berg for computing the results of this experiment.

REFERENCES

Ahlbrecht, M., and M. Weber. 1995. "Hyperbolic Discounting Models in Prescriptive Theory of Intertemporal Choice." *Zeitschrift für Wirtschafts- und Sozialwissenschaften* 115: 535–68.

Anderhub, V., R. Müller, and C. Schmidt. 2001. "Design and Evaluation of an Economic Experiment via the Internet." Merit-Infonomics Research Memorandum Series, Maastricht, the Netherlands.

Antonides, G., and S.R. Wunderink. 2001. "Time Preference and Willingness to Pay for an Energy-Saving Durable Good." *Zeitschrift für Sozialpsychologie* 32, 3: 133–41.

Barber, B.M., and T. Odean. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *Journal of Finance* 55: 773–806.

Baron, J. 2001. "Purposes and Methods." *Behavioral and Brain Sciences* 24: 403.

Bateman, I., A. Monroe, B. Rhodes, C. Starmer, and R. Sugden. 1997. "A Test of the Theory of Reference-Dependent Preferences." *Quarterly Journal of Economics* 112: 479–505.

Bazerman, M. 1998. *Judgment in Managerial Decision Making.* New York: John Wiley.

Becker, G.M., M.H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9: 226–32.

Belk, R.W. 1974. "An Exploratory Assessment of Situational Effects in Buyer Behavior." *Journal of Marketing Research* 11: 156–63.

Bergstrom, T.C., and J.H. Miller. 1999. *Experiments with Economic Principles.* New York: McGraw-Hill.

Binmore, K. 1987. "Experimental Economics." *European Economic Review* 31: 257–64.

Camerer, C.F., and R.M. Hogarth. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty* 19: 7–42.

Chaiken, S., and Y. Trope. 1999. *Dual Process Theories in Social Psychology.* New York: Guilford Press.

Chamberlin, E.H. 1948. "An Experimental Imperfect Market." *Journal of Political Economy* 56, 2: 95–108.

Chapman, G.B. 1998. "Similarity and Reluctance to Trade." *Journal of Behavioral Decision Making* 11: 47–58.

Charness, G., and M. Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117: 817–69.

Christensen-Szalanski, J.J.J. 1984. "Discount Functions and the Measurement of Patients' Values: Women's Decisions During Childbirth." *Medical Decision Making* 4: 48–57.

Coase, R.H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1–44.

Cooper, J., and R.H. Fazio. 1984. "A New Look at Dissonance Theory." In L. Berkowitz, ed., *Advances in Experimental Social Psychology* 17:229–66. New York: Academic Press.

Cubitt, R.P., C. Starmer, and R. Sugden. 1998. “On the Validity of the Random Lottery Incentive System.” *Experimental Economics* 1: 115–31.
Davis, H.P., and R.L. Durham. 2001. “Economic and Psychological Experimental Methodology: Separating the Wheat from the Chaff.” *Behavioral and Brain Sciences* 24: 405–6.
Dawes, R.M. 1980. “Social Dilemmas.” *Annual Review of Psychology* 31: 169–93.
DeGroot, I.M. 2003. “Product Trials: The Effects of Direct Experience on Product Evaluation.” Ph.D. thesis, Tilburg University.
Denes-Raj, V., and S. Epstein. 1994. “Conflict Between Intuitive and Rational Processing: When People Behave Against Their Better Judgment.” *Journal of Personality and Social Psychology* 66: 819–29.
DeYoung, R. 1993. “Market Experiments: The Laboratory Versus the Classroom.” *Journal of Economic Education* 24: 335–51.
Dhar, R., and K. Wertenbroch. 2000. “Consumer Choice Between Hedonic and Utilitarian Goods.” *Journal of Marketing Research* 37: 60–71.
Epstein, S. 1973. “The Self-Concept Revisited, or a Theory of a Theory.” *American Psychologist* 28: 404–16.
Fels, R. 1993. “This Is What I Do, and I Like It.” *Journal of Economic Education* 24: 365–70.
Festinger, L. 1957. *A Theory of Cognitive Dissonance.* Evanston, IL: Row, Peterson.
Fischhoff, B., P. Slovic, and S. Lichtenstein. 1977. “Knowing with Certainty: The Appropriateness of Extreme Confidence.” *Journal of Experimental Psychology: Human Perception and Performance* 3: 552–64.
Fishburn, P.C., and A. Rubinstein. 1982. “Time Preference.” *International Economic Review* 23: 677–94.
Foa, U.G. 1971. “Interpersonal and Economic Resources.” *Science* 171: 345–51.
Frank, R.H., T. Gilovich, and D.T. Regan. 1993. “Does Studying Economics Inhibit Cooperation?” *Journal of Economic Perspectives* 7: 159–71.
Gattig, A. 2002. “Intertemporal Decision Making: Studies on the Working of Myopia.” Ph.D. thesis, University of Groningen, the Netherlands.
Goodie, A.S. 2001. “Are Scripts or Deception Necessary When Repeated Trials Are Used? On the Social Context of Psychological Experiments.” *Behavioral and Brain Sciences* 24: 412.
Hanemann, W.M. 1991. “Willingness to Pay and Willingness to Accept: How Much Can They Differ?” *American Economic Review* 81: 635–47.
Herrnstein, R.J., and D. Prelec. 1999. “Melioration: A Theory of Distributed Choice.” *Journal of Economic Perspectives* 5: 137–56.
Hertwig, R., and A. Ortmann. 2001. “Experimental Practices in Economics: A Methodological Challenge for Psychologists.” *Behavioral and Brain Sciences* 24: 383–451.
Hey, J. D. 1998. “Experimental Economics and Deception: A Comment.” *Journal of Economic Psychology* 19: 397–401.
Hilton, D.J. 2001. “Is the Challenge for Psychologists to Return to Behaviorism?” *Behavioral and Brain Sciences* 24: 415–16.
Hirschman, E.C., and M.B. Holbrook. 1982. “Hedonic Consumption: Emerging Concepts, Methods and Propositions.” *Journal of Marketing* 46: 92–101.
Hofstaedter, D. 1983. “Metamagical Themas.” *Scientific American* 248: 14–28.
Holt, C.A. 1986. “Preference Reversals and the Independence Axiom.” *American Economic Review* 76: 508–15.
———. 1996. “Trading in a Pit Market.” *Journal of Economic Perspectives* 10: 193–203.
Hsee, C.K. 1996. “The Evaluability Hypothesis: An Explanation for Preference Reversals Between Joint and Separate Evaluations of Alternatives.” *Organizational Behavior and Human Decision Processes* 67: 247–57.
Johnson, E.J., J. Hershey, J. Meszaros, and H. Kunreuther. 1993. “Framing, Probability Distortions, and Insurance Decisions.” *Journal of Risk and Uncertainty* 7: 35–51.
Kagel, J.H., and A.E. Roth. 1995. *Handbook of Experimental Economics.* Princeton, NJ: Princeton University Press.
Kahneman, Daniel. 2003a. “A Psychological Perspective on Economics.” *American Economic Review* 93: 162–68.
———. 2003b. “Maps of Bounded Rationality: Psychology for Behavioral Economics.” *American Economic Review* 93: 1449–75.
Kahneman, D., J.L. Knetsch, and R.H. Thaler. 1990. “Experimental Tests of the Endowment Effect and the Coase Theorem.” *Journal of Political Economy* 98: 1325–48.

Kahneman, D., and A. Tversky. 1979. "Prospect Theory: Analysis of Decisions Under Risk." *Econometrica* 47: 263–91.

———. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5: 297–324.

Knetsch, J.L. 1989. "The Endowment Effect and Evidence of Nonreversibility of Indifference Curves." *American Economic Review* 79: 1277–84.

———. 1995. "Asymmetric Valuation of Gains and Losses and Preference Order Assumptions." *Economic Inquiry* 33: 134–41.

Knetsch, J.L., and J.A. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics* 99: 507–21.

Levin, I.P., and G.J. Gaeth. 1988. "How Consumers Are Affected by the Framing of Attribute Information Before and After Consuming the Product." *Journal of Consumer Research* 15: 374–78.

Loewenstein, G. 1999. "Experimental Economics from the Vantage Point of Behavioral Economics." *The Economic Journal* 109: F25–34.

Loewenstein, G., and D. Prelec. 1992. "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *Quarterly Journal of Economics* 107: 573–97.

Lutz, R.J., and P. Kakkar. 1975. "The Psychological Situation as a Determinant of Consumer Behavior." *Advances in Consumer Research* 1: 439–53.

McDaniel, T., and C. Starmer. 1998. "Experimental Economics and Deception: A Comment." *Journal of Economic Psychology* 19: 403–9.

Mitchell, A.A. 1981. "The Dimensions of Advertising Involvement." In K.B. Monroe, ed., *Advances in Consumer Research,* 25–35. Ann Arbor, MI: Association for Consumer Research.

Mittal, B. 1988. "The Role of Affective Choice Mode in the Consumer Purchase of Expressive Products." *Journal of Economic Psychology* 9: 499–524.

———. 1994. "A Study of the Concept of Affective Choice Mode for Consumer Decisions." *Advances in Consumer Research* 21: 256–62.

Musgrave, A. 1981. "Unreal Assumptions in Economic Theory: The F-Twist Untwisted." *Kyklos* 34: 377–87.

Odean, T. 1998. "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance* 53: 1775–98.

Orne, M.T., and K.E. Scheibe. 1964. "The Contribution of Nondeprivation Factors in the Production of Sensory Deprivation Effects: The Psychology of the Panic Button." *Journal of Abnormal and Social Psychology* 68: 3–12.

Ortmann, A., and R. Hertwig. 1997. "Is Deception Acceptable?" *American Psychologist,* July, 746–47.

Park, C. W., and B. Mittal. 1985. "A Theory of Involvement in Consumer Behavior: Problems and Issues." In J.N. Sheth, ed., *Research in Consumer Behavior,* 1:201–31. Greenwich, CT: JAI Press.

Purohit, D. 1995. "Playing the Role of Buyer and Seller: The Mental Accounting of Trade-ins." *Marketing Letters* 6: 101–10.

Rakow, T. 2001. "Theorize It Both Ways?" *Behavioral and Brain Sciences* 24: 425–26.

Read, D., and B. Van Leeuwen. 1998. "Predicting Hunger: The Effects of Appetite and Delay on Choice." *Organizational Behavior and Human Decision Processes* 76: 189–205.

Roth, A.E. 1995. "Introduction to Experimental Economics." In J.H. Kagel and A.E. Roth, eds., *The Handbook of Experimental Economics.* Princeton, NJ: Princeton University Press.

———. 2001. "Form and Function in Experimental Design." *Behavioral and Brain Sciences* 24: 427–28.

Samuelson, W., and R. Zeckhauser. 1988. "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty* 1: 7–59.

Scott, W.A. 1957. "Attitude Change Through Reward of Verbal Behavior." *Journal of Abnormal and Social Psychology* 55: 72–75.

Shefrin, H.M., and M. Statman. 1985. "The Disposition to Sell Winners Too Early and Ride Losers Too Long." *Journal of Finance* 40: 777–90.

Sloman, S.A. 1996. "The Empirical Case for Two Systems of Reasoning." *Psychological Bulletin* 119: 3–22.

Smith, V.L. 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy* 70, 2: 111–37.

Strahilevitz, M., and G. Loewenstein. 1998. "The Effects of Ownership History on the Valuation of Objects." *Journal of Consumer Research* 25: 276–89.

Thaler, R.H. 1980. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization* 1: 39–60.

———. 1981. “Some Empirical Evidence on Dynamic Inconsistency.” *Economics Letters* 8: 201–7.

———. 1986. “The Psychology and Economics Handbook: Comments on Simon, on Einhorn and Hogarth, and on Tversky and Kahneman.” *Journal of Business* 59, 4: S279–84.

Tomlinson, C.D. 2002. “Specification for the 4th Generation of ELSE Experimental Software.” ESRC ELSE Centre Internal Report, University College, London.

Tversky, A., and D. Kahneman. 1991. “Loss Aversion in Riskless Choice: A Reference Dependent Model.” *Quarterly Journal of Economics* 106: 1039–61.

Van Dijk, E., and D. Van Knippenberg. 1996. “Buying and Selling Exchange Goods: Loss Aversion and the Endowment Effect.” *Journal of Economic Psychology* 17: 517–24.

Varian, H.R. 2002. “Observe, Theorize, Measure, Test and Don’t Overlook What Goes Wrong: Nobel Experiments.” *New York Times,* October 24.