EXPERIMENTAL BUSINESS RESEARCH

Experimental Business Research

Marketing, Accounting and Cognitive Perspectives

VOLUME III

Edited by

RAMI ZWICK

Hong Kong University of Science and Technology, China

and

AMNON RAPOPORT

University of Arizona, Tucson, U.S.A.
and
Hong Kong University of Science and Technology, China

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 0-387-24215-5 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-10 0-387-24244-9 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-0-387-24215-6 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-0-387-24244-6 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands.

Contents

PREFACE

Rami Zwick
Hong Kong University of Science and Technology

Amnon Rapoport
University of Arizona
And
Hong Kong University of Science and Technology

This volume (and volume II) includes papers that were presented at the Second Asian Conference on Experimental Business Research held at the Hong Kong University of Science and Technology (HKUST) on December 16–19, 2003. The conference was a follow up to the first conference that was held on December 7–10, 1999, the papers of which were published in the first volume (Zwick, Rami and Amnon Rapoport (Eds.), (2002) *Experimental Business Research.* Kluwer Academic Publishers: Norwell, MA and Dordrecht, The Netherlands). The conference was organized by the Center for Experimental Business Research (cEBR) at HKUST and was chaired by Amnon Rapoport and Rami Zwick. The program committee members were Paul Brewer, Kenneth Shunyuen Chan, Soo Hong Chew, Sudipto Dasgupta, Richard Fielding, James R. Frederickson, Gilles Hilary, Ching-Chyi Lee, Siu Fai Leung, Ling Li, Francis T Lui, Sarah M Mcghee, Fang Fang Tang, Winton Au Wing Tung and Raymond Yeung. The papers presented at the conference and a few others that were solicited especially for this volume contain original research on individual and interactive decision behavior in various branches of business research including, but not limited to, economics, marketing, management, finance, and accounting.

The following introduction to the field of Experimental Business Research and to our center at HKUST replicates the introduction from Volume II. Readers familiar with the introduction to Volume II are advised to skip Sections 1 and 2 below.

1. THE CENTER FOR EXPERIMENTAL BUSINESS RESEARCH

The Center for Experimental Business Research (cEBR) at HKUST was established to serve the needs of a rapidly growing number of academicians and business leaders in Hong Kong and the region with common interests in experimental business research. Professor Vernon Smith, the 2002 Nobel laureate in Economics and a

current member of cEBR's External Advisory Board, inaugurated the Center on September 25, 1998, and since than the Center has been recognized as the driving force behind experimental business research conducted in the Asia-Pacific region. The mission of cEBR is to promote the use of experimental methods in business research, expand experimental methodologies through research and teaching, and apply these methodologies to solve practical problems faced by firms, corporations, and governmental agencies. The Center accomplishes this mission through three agendas: research, education, and networking and outreach programs.

2. WHAT IS EXPERIMENTAL BUSINESS RESEARCH?

Experimental Business Research adopts laboratory based experimental economics methods to study an array of business and policy issues spanning the entire business domain including accounting, economics, finance, information systems, marketing and management and policy. "Experimental economics" is an established term that refers to the use of controlled laboratory-based procedures to test the implications of economic hypotheses and models and discover replicable patterns of economic behavior. We coined the term "Experimental Business Research" in order to broaden the scope of "experimental economics" to encompass experimental finance, experimental accounting, and more generally the use of laboratory-based procedures to test hypotheses and models arising from research in other business related areas, including information systems, marketing, and management and policy.

Behavioral and experimental economics has had an enormous impact on the economics profession over the past two decades. The 2002 Nobel Prize in Economics (Vernon Smith and Danny Kahneman) and the 2001 John Bates Clark Medal (Matthew Rabin) have both gone to behavioral and experimental economists. In recent years, behavioral and experimental research seminars, behavioral and experimental faculty appointments, and behavioral and experimental PhD dissertations have become common at leading US and European universities.

Experimental methods have played a critical role in the natural sciences. The last fifteen years or so have seen a growing penetration of these methods into other established academic disciplines including economics, marketing, management, accounting and finance, as well as numerous applications of these methods in both the private and public sectors. cEBR is active in introducing these methodologies to Hong Kong and the entire Pacific Basin. We briefly describe several reasons for conducting such experiments.

First and most important is the used of experiments to design institutions (i.e., markets) and for evaluating policy proposals. For example, early experiments that studied the one-price sealed bid auction for Treasury securities in the USA helped motivate the USA Treasury Department in the early 1970 to offer some long-term bond issues. Examples for evaluating policy proposals can be found in the area of voting systems, where different voting systems have been evaluated experimentally in terms of the proportion of misrepresentation of a voter's preferences (so- called "sophisticated voting"). In the past decade, both private industry and governmental

agencies in the USA have funded studies on the incentives for off-floor trading in continuous double auction markets, alternative institutions for auctioning emissions permits, and market mechanisms for allocating airport slots and the FCC spectrum auction. More recently, Hewlett-Packard has used experimental methods to evaluate contract policy in areas from minimum advertised price to market development funds before rolling them out to its resellers, and Sears used experimental methods to develop a market for logistics.

Second, experiments are used to test a theory or determine the most useful competing theories. This is accomplished by comparing the behavioral regularities to the theory's predictions. Examples can be found in the auction and portfolio selection domains. Similarly, business experiments have been conducted to explore the causes of a theory's failure. Examples are to be found in the fields of bargaining, accounting, and the provision of public goods.

Third, because well-formulated theories in most sciences tend to be preceded by systematically collected observations, business experiments are used to establish empirical regularities as a basis for the construction of a new theory. These empirical regularities may vary considerably from one population of agents to another, depending on a variety of independent variables including culture, socio-economic status, previous experience and expertise of the agents, and gender.

Finally, experiments are used to compare environments, using the same institution, or comparing institutions, while holding the environment constant.

3. CONTENT

Whereas Volume II contains papers under the general umbrella of economic and managerial perspectives, the present volume includes papers from the fields of Marketing, Accounting, and Cognitive Psychology. Volume III includes 14 chapters. The 33 contributors come from many of the disciplines that are represented in a modern business school.

Chapter 1 by Zhao, Meyer, and Han explores consumers' ability to optimally anticipate the value they will draw from new product features that are introduced to enhance the performance of existing technologies. The research is motivated by the common observation that consumers frequently purchase more technology than they can realistically make use of. Central to their work is the idea that a general over-buying bias may, in fact, have a strong theoretical basis. Drawing on prior work in affective forecasting, they hypothesize that when buying new technologies consumers will usually have a difficult time anticipating how they will utilize a product after it is purchased, and will be prone to believe that the benefits of attribute innovations that are perceived now will project in a simple fashion into the future. Implicit to this over-forecast is a tendency to underestimate the impact of factors that may likely serve to diminish usage in the future, such as frustration during learning and satiation. Consequently, there is a tendency for consumers to systematically evaluate product innovations through rose-colored glasses, imagining that they will have a larger and more positive impact on the future lives than they

most often will likely end up having. This general hypothesis is tested in the context of a computer simulation in which subjects are trained to play one of three different forms of an arcade game where icons are moved over a screen by different forms of tactile controls. Respondents are then given the option to play a series of games for money with either their incumbent game platform or pay to play with an alternative version that offers an expanded set of controls. As hypothesized, subjects displayed an upwardly-biased valuation for the new sets of controls; adopters underutilized them and displayed a level of game performance that was not better than those who never upgraded. A follow-up study designed to understand the process underlying the bias indicated that while adopters over-forecasted the degree to which they would make use of the new control, they did not over-forecast performance gains. Hence, the key driver of adoption decisions appeared to be an exaggerated belief in the hedonic pleasure that would be derived from owning and utilizing the new control as opposed to any objective value it might provide.

What is notable about their results is that the evidence for the optimism bias was derived from a context designed to facilitate rational assessments of innovation value. Specifically, subjects were given a clearly-stated metric by which the objective value of the innovation could have been assessed, there was a direct monetary penalty for overstating value (the game innovation was paid for by a point deduction), and the innovation itself was purely functional rather than aesthetic (a new control added to the same graphic game platform). Yet, subjects still succumbed to the same biases.

Chapter 2 by Kim and Waller reports on a behavioral accounting experiment on strategic interaction in a tax compliance game. The experiment employed a three-step approach. First, subjects were assigned to the opposing roles of auditor and strategic taxpayer. This step addressed a past criticism of behavioral accounting research: economic mechanisms such as the interaction of players with conflicting preferences potentially eliminate the decision biases found in individual settings. Second, the experiment operationalized a game-theoretic model of the tax compliance problem by Graetz, Reinganum, and Wilde. In the model, the taxpayer chooses a strategy $\{\alpha, 1 - \alpha\}$ when true income is high, whereby he under-reports income with probability α and honestly reports income with probability $1 - \alpha$. The auditor chooses a strategy $\{\beta, 1 - \beta\}$ when reported income is low, whereby she conducts a costly audit with probability β and does not audit with probability $1 - \beta$. The model assumes two types of taxpayer: proportion ρ of strategic taxpayers who maximize expected wealth, and proportion $1 - \rho$ of ethical taxpayers who adhere to an internalized norm for honesty. The auditor maximizes expected net revenue, i.e., tax plus fine minus audit cost. Before conducting an audit, the auditor cannot distinguish between the taxpayer types. When the auditor conducts an audit and detects under-reporting, the taxpayer must pay a fine plus the tax for high true income. The model implies that the optimal audit rate β^* is insensitive to an exogenous change in ρ, as long as ρ exceeds a threshold. The strategic taxpayer fully absorbs the change in ρ by adjusting the optimal rate of under-reporting income α^*. Third, the experiment manipulated two variables that are considered irrelevant by the game-theoretic

model, i.e., the level of ρ and uncertainty about ρ, in order to test hypotheses about auditors' choice of the audit rate, β.

Contrary to the model, Kim and Waller hypothesized that an auditor with limited rationality will use ρ as a cue for adjusting β. The hypotheses assume a simple additive process: $\beta = \beta' + \beta''$, where β' depends on ρ, and β'' depends on a belief about the taxpayer's strategy. The results show positive associations between ρ and β', and between auditors' uncertainty about ρ and β'. The auditors formed incorrect beliefs about the taxpayers' responses, which affected β''. The auditors incorrectly believed that the taxpayers increased the rate of under-reporting income as ρ increased, and that the taxpayers expected a higher audit rate when the auditors faced uncertainty about ρ. The taxpayers correctly believed that β increased as ρ increased, and responded by decreasing the rate of under-reporting income.

Chapter 3 by Bodoff, Levevq, and Zhang explores the beliefs that underline policies such as the SEC's Fair Disclosure Rule, and technologies such as SEC EDGAR, that aim to disseminate corporate disclosures to a wider audience.

Rational expectations models have been successful in predicting equilibrium prices in experimental markets of risky assets. In previous work, the authors explored whether such models are also useful in their other predictions regarding welfare in the sense of *ex ante* expected utility. They previously found that they are not, i.e. that subjects did not prefer the predicted market condition. In particular, when subjects could select the environment in which to trade, and the environment was characterized by the proportion of informed traders, subjects' preference for the fraction of informed traders was "Half > None > All", i.e. investors most favored a situation where a random half of investors are informed. Analytical predictions based on theories of non-revealing and full-revealing prices would predict a different preference order: "None > All > Half". In this chapter, the authors explore the tension between the correct predictions of the equilibrium solution and the incorrect predictions of subjects' preferences. In analytical models, predictions of EU follow by definition from the equilibrium prices, so it would be expected that if a theory properly characterizes the equilibrium, then it will properly predict *ex ante* EU. But this is apparently not the case, which suggests an anomaly. If market equilibriums were perfectly accurate, then the anomaly would be total. Because the predictions of market equilibrium are not perfect, the authors explored the possibility that perhaps subjects' preferences were consistent with the expected utility of the actual market equilibriums, if not with the analytically predicted market equilibrium. They found that they still were not. Ultimately, the authors adopt another approach, and propose that subjects have different attitudes toward different sources of risk, a phenomenon which traditional analytical models do not consider.

In Chapter 4, Amaldoss and Rapoport report the results of an experiment designed to investigate the effects of idiosyncartic investments in collaborative networks. The research is motivated by a desire to better understand the emerging phenomenon of networks, rather than individual firms, developing new products. In contrast to the common belief of alliance managers, the authors have shown that in theory the joint investment of network partners does not decrease as a network

grows in size. Specifically, if the investments are recoverable, the joint investment should increase as the network size increases. But if they are not, then joint investment should not change with network size. On extending the theoretical model to investigate competition among a large number of networks ($N > 2$), the authors found that the effect of number of competing networks on joint investment depends on whether the investments are recoverable. If they are, it exerts a positive effect, but if they are not, it has a negative impact. In this chapter they describe an experimental test of these predictions in a laboratory setting. The experimental results support the qualitative predictions of the model. That is, they report that the joint investment increases as network size increases when investment is recoverable. But joint investment does not change significantly with increase in network size when investments are nonrecoverable. Amaldoss and Rapoport also detected a trend toward equilibrium behavior over multiple iterations of the stage game, and found that an adaptive learning model (EWA) accounts for the investment patterns of the subjects over time.

Chapter 5 by Hertwig and Ortmann discusses the methodological insights that experimental economists may derive from the debate in psychology about the reality of cognitive illusions. The authors have argued elsewhere that psychologists can learn from the experimental practices of economists. In this chapter, the proposed directional cross fertilization is reversed.

Hertwig and Ortmann discuss the heuristics-and-biases program launched by Kahneman and Tversky in the early 1970s. This program stresses that people have only limited "reasoning power" at their disposal and hence must rely on cognitive heuristics to make judgments and choices. Although these heuristics are highly economical and usually effective, they can lead to systematical and predictable errors that are variously referred to as biases, fallacies, or *cognitive illusions*. The heuristics-and-biases program has attracted the attention of numerous social scientists, including economists and legal scholars. In fact, much of today's work in behavioral economics and behavioral finance draws inspiration and concepts from the heuristics-and-biases program. This attention is warranted because systematic biases may have important implications for economic behavior.

As the heuristics-and-biases program has gained acceptance outside psychology, it has also drawn criticism within psychology. Some critics have suggested that the heuristics-and-biases research strategy has a built-in bias to find cognitive illusions, and others have claimed that some cognitive illusions are themselves illusory. Perhaps the most influential objections were voiced by Gigerenzer, who has argued that the heuristics to which cognitive illusions are attributed are not precise process models; that the heuristics-and-biases program relies on a narrow definition of rationality; and that cognitive illusions can be reduced or made to disappear by representing statistical information differently than it typically had been in the heuristics-and-biases experiments.

Hertwig and Ortmann's focus in this chapter is neither the controversy about cognitive illusions nor its implications for rationality. Instead, it is what they see as the important methodological insights that have emerged from the controversy, which

can inform the choices that all behavioral experimenters wittingly or unwittingly make when they sample and represent stimuli for their experiments. In particular, Hertwig and Ortmann discuss the issues of stimulus sampling and the way these stimuli are presented to subjects, and then show that both factors matter in experiments with economical context.

For example, the question whether and how to sample from the environment has not been of much concern to experimental economists. Little attention has been paid to how representative these environments are of their real-world counterparts and the neglect of representative design has been amplified by the practice of using abstract tasks. However, there is now ample evidence that stripping away content and context prevents participants from applying the strategies that they use in their usual habitats.

Similarly, the authors argue that stimulus representation is an important factor in experimental economics and demonstrate how representing the stimuli in different formats (e.g., graphical) can dramatically reduce inconsistent behavior in an Allais type task even if boundary gambles are used.

Chapter 6 by Kramer and Budescu explores the role of vagueness (ambiguity) in choice. Ellsberg's paradox (1961) involves an inconsistent set of choices amongst two urns, each filled with red or blue marbles, but whose composition is known with different levels of precisions. In the "classic paradox" the DMs' choices indicate that the more certain urn is more likely to produce the desired marble *for each color*, implying that Pr(red) + Pr(blue) >1. Several empirical studies have investigated variations of this paradigm, but none have demonstrated conclusively the presence of Ellsberg's paradox in situations where the composition of neither urn is known precisely. In the present study the authors investigate this Vague-Vague (V-V) case, where neither of the urns' color probabilities are specified precisely, but one urn's probabilities are always more precise than the other. They show that people prefer precisely specified gambles and succumb to Ellsberg's paradox in these "dual vagueness" situations. The tendency to avoid the more vague urn and the prevalence of the classic paradox (and all the other two-choice patterns) is similar, but not identical, in the standard P-V (Precise-Vague) and the V-V situations. When conditioning on the midpoint (the middle of the probability range[s]), there is a reversal in vagueness avoidance between P-V and V-V cases. Otherwise, their results indicate that P-V and V-V cases are not qualitatively different, and it is more appropriate to think of them as defining a continuum of "degree of vagueness." The P-V case is just one, admittedly critical and intriguing, endpoint of this continuum. In both P-V and V-V cases, the prevalence of the paradoxical pattern of choices depends primarily on the ranges of the two gambles (i.e., the relative precision and minimal imprecision of the pair) and, to a lesser degree, on the pair's common midpoint.

In Chapter 7, Levy and Levy experimentally test the overweighing of recent return observations in an investment experiment with business school students and financial practitioners. They find that it is mainly the most recent observation that is overweighed, and that this overweighing is very strong. They estimate the decision weight attached to the most recent observation as approximately twice the

objective probability. In this framework, probabilities are subjectively distorted on the basis of the temporal sequence of the observations, unlike the distortion that takes place in single-shot lottery type decisions (as in Prospect Theory, Cumulative Prospect Theory, or Rank Dependent Expected Utility models). This framework is applicable to circumstances where individuals are given observations as time series, as they are in financial markets, rather than a "given" set of outcomes and probabilities, as in many decision-making experimental setups. The case of the temporal probability distortion seems more relevant to actual economic decisions because in practice investors observe time series data regarding corporate earnings, mutual fund returns, etc., and their decisions are based on these time series. The findings of this paper suggest a simple explanation to several important economic phenomena like momentum (the positive short run autocorrelation of stock returns) and the relationship between recent fund performance and the flow of money to the fund. The results also have strong implications to asset allocation, pricing, and the risk-return relationship.

Chapter 8 by Blume, DeJong, and Maier concerns cognitive processes in common-interest spatial dispersion games in which the agents' common goal is to choose distinct locations. The games are characterized by multiple, non-strict equilibria. It is an open question whether players can select and attain equilibrium in such games and if equilibrium can be achieved, how long will it take and what are its characteristics. A further question is whether the insights from matching games extend to dispersion games. The authors report on an experiment designed to answer these questions. In their setup, cognition matters because agents may be differentially aware of the dispersion opportunities that are created by the history of the game. Their main finding is that strategic interaction magnifies the role of cognitive constraints. Specifically, with cognitive constraints, pairs of agents fail to solve a dispersion problem that poses little or no problem for individual agents playing against themselves. When they remove the cognitive constraints, pairs of agents solve the same problem just as well as individuals do. In addition, they report that when playing against themselves agents do not change the mode by which they solve the dispersion problem when their design removes the cognitive constraints.

In chapter 9, Chong, Camerer, and Ho further develop their cognitive hierarchy (CH) model. Strategic thinking, best-response, and mutual consistency (equilibrium) are three key modeling principles in non-cooperative game theory. In a previous paper, the authors relaxed mutual consistency to predict how players are likely to behave in one-shot games before they can learn to equilibrate. They introduced a one-parameter cognitive hierarchy (CH) model to predict behavior in one-shot games. The CH approach assumes that players use k steps of reasoning with frequency $f(k)$. In their previous paper they assumed $f(k)$ to be a one-parameter Poisson distribution. This chapter investigates and lends support to the generality and precision of this Poisson CH model in three ways: 1. An unconstrained general distribution CH model is found to offer only marginal improvement in fit over its Poisson cousin and hence this suggests that the Poisson approximation is reasonable. 2. The steps of thinking players use in games are found to positively correlate with response time

and schools they attend which suggests that cognitive hierarchy captures realistically a reasoning mechanism that goes on in the brain of these players. 3. Several classes of interesting economic problems, including asset pricing and business entry, can be explained by the iterated reasoning of the Poisson CH model. When compared to the Quantal Response Equilibrium model, which relaxes the best-response assumption of equilibrium theory, the better fit of Poisson CH model seems to suggest that mutual consistency is a more plausible assumption to relax in explaining deviation from equilibrium theory.

Chapter 10 by Fox, Bardolet, and Lieb explores a wide range of judgment and decision tasks in which people are called upon to allocate a scarce resource (e.g., money, choices, belief) over a fixed set of possibilities (e.g., investment opportunities, consumption options, events). The authors observe that in these situations people tend to invoke maximum entropy heuristics in which they are biased toward even allocation. Moreover, they argue that before applying these heuristics, decision makers subjectively partition the set of options into groups over which they apply even allocation. As a result, allocations vary systematically with the particular partition that people happen to invoke, a phenomenon called *partition dependence*. The authors review evidence for maximum entropy heuristics and partition dependence in the following domains: (1) decision analysis in which the degree of belief and importance weights must be distributed among possible events and attributes, respectively; (2) managerial decision making in which money and other organizational resources are allocated among risky projects, divisions, and organizational stakeholders; and (3) consumer choice in which individuals select among various consumption goods and consumption time periods.

In Chapter 11, Gneezy investigates the influence of prior gains and losses on the risk attitude of people. Empirical findings suggest that in decisions under uncertainty people evaluate outcomes relative to a reference level: they are risk-seeking in the domain of losses and risk-averse in the domain of gains. The finance literature uses this finding to predict/explain the "disposition effect," which is the tendency of investors to sell assets that have gained value ("winners") too early and ride assets that have lost value ("losers") too long. The purpose of the experiment reported in this chapter is to investigate the influence of prior gains and losses on the risk attitude of people. Unlike the case of real market data, the stylized experimental setup allows the author to gain insight into the decision-making process of individuals. Furthermore, using a stylized decision problem makes the benchmark prediction very clear and testable. One of the main goals was to find evidence on how prior gains and losses influence the risk behavior of people, by shifting the reference level. The results show that prior gains and losses do influence the risk attitude, and in a different way from that predicted by the rational theory (expected utility). The disposition effect prediction that people will be reluctant to sell losing assets found strong empirical support with the traditional assumption that the reference level is the initial purchase price of the stock. This finding supports the empirical research done on real market data. The use of a stylized process also allows for more refined tests about the way reference levels are formed. In particular, it is

possible to learn about how it depends on the history of gains and losses. This is important because, for example, prospect theory is useless as a descriptive theory without a "good" assumption about the reference levels. It was found that when the peak of the process was used as a reference level, the descriptive power of the theory increased dramatically.

Chapter 12 by Dufwenberg and Gneezy investigates the relationship between gender and coordination. Groups of six females or six males played the minimal effort coordination game for ten periods. Little difference was found between the groups of men and women with regard to their ability to avoid the least efficient equilibrium. The results show some differences in the initial stages of the game, but these differences quickly disappear and no difference is found in later stages. In addition to reporting this result, the authors raise a methodological issue: Is there a bias in the research community against reporting or publishing results that document the absence of a gender effect? The results reported in this chapter are not "positive," in the sense that no difference in behavior between females and males was found. The authors believe that in order to truly understand the differences in behavior between genders, one should not only report or publish experiments and results that show positive differences because such practice would bias perceptions about the magnitude and the limits of the differences.

The last two chapters discuss the use of laboratory- and class-based experiments intended to enhance teaching and learning. Successful attempts to teach business related courses through experiments and projects conducted in computerized laboratories (e.g., the Economic Science Laboratory at the University of Arizona, the Laboratory for Economic and Political Research at the California Institute of Technology) all testify to the benefit of integrating this new methodology in the teaching of business related courses at the undergraduate, graduate, and MBA levels. There is by now ample evidence that "hands-on" learning through experimentation, in which different economic scenarios are created under controlled laboratory conditions, is a very effective way of acquiring new concepts and procedures, gaining insight into business practices, and learning how to make better decisions.

The basic idea that underlines this new teaching methodology is that actual experience in carefully designed experiments, whether they are selected to test basic theoretical concepts or mirror business problems that appear in practice, is critical for effective teaching of business. The experience takes two forms: a personal experience of participation and a supply of data produced by the participants. The personal experience is invaluable for maintaining the student's attention and motivating his/her understanding of the material, but it is the data produced by the group that truly make clear the power of economic principles in understanding markets, bargaining, and other business decision environments.

Chapter 13 by Croson, Donohue, Katok, and Sterman describes an experiment that illustrates the challenges of supply chain management. Supply chain management involves the management of orders and shipments of goods through a supply chain; for example, shipping beer from the manufacturer to the distributor to the

wholesaler and then to the retailer for sale to customers, and transmitting the orders for beer back up the supply line. A large body of research investigates these issues theoretically. However, in addition to the theoretical operational challenges, there are also cognitive limitations that managers face which prevent them from optimally managing their supply chains. Chapter 13 describes an in-class experimental game that can be used to illustrate a number of these challenges, operational and cognitive, that managers face in supply chain management. The experiment is well-suited for undergraduate, MBA, or executive teaching, and has been used in all those forums. Exactly which treatments to choose, and how deep the debriefing should be, will depend on the sophistication of the audience as well as the manner in which the teacher chooses to implement the experiment (physical or computer).

The last chapterby Erev and Livne-Tarandach describes an innovative approach to the use of experimentally derived findings in experiment-based exams in the social sciences. The authors have analyzed GRE exams and highlighted an important difference between the natural and the behavioral sciences. Most questions in Physics ask the examinee to predict the results of particular experiments. On the other hand, nearly all questions in Psychology deal with abstract terms. The analysis in Chapter 14 clarifies this difference, and proposes two related steps that can lessen the gap.

The first step addresses the difficulty of developing experiment-based questions in the behavioral sciences. The authors assert that the main stumbling block, from the developer's point of view, lies in identifying questions with unambiguous correct answers. The solution proposed here is technical. It requires focusing each question on a particular experiment that has been run. With this focus in mind, the correct answer is crystal clear: It is the observed experimental result. Their analysis suggests that the discriminative power of experiment-based questions based on this technical solution is at par with the discriminative power of more typical abstract questions. The second step requires some changes in the information collected by researchers and presented to students. The authors assert that the discriminative power of experiment-based questions can be improved through the standardization of descriptive models and experimental procedures. The standardization of descriptive models as suggested, for example, by Erev, Roth, Slonim, and Barron is expected to have three benefits: It would allow unbiased selection of experimental tasks; it would clarify the boundaries of descriptive models; and it would provide guidance where models conflict with intuition, introspection, and or personal experience. The standardization of experimental procedures is expected to be beneficial in that it would facilitate clear and parsimonious presentations of experiment-based questions.

Erev and Livne-Tarandach believe that the use of experiment-based questions to evaluate students in behavioral science courses is likely to have many attractive outcomes. In addition to making behavioral science exams more similar to those in the natural sciences, this effort will advance the behavioral sciences in substantial ways. A focus on predictions in exams is likely to have a similar effect on courses, textbooks, and mainstream research.

ACKNOWLEDGEMENTS

We owe thanks to many for the successful completion of this volume. Most importantly, we express our gratitude to the contributors who attended the conference and participated in insightful discussions. The conference was supported financially by a grant from the Hong Kong University Grant Commission to cEBR (Project No. HKUST-3, Experiential based teaching for networked economy), and by an RGC Direct Allocation Grant (Project No. DAG02/03.BM78) to Rami Zwick and Soo Hong Chew. Additional financial support was given by HKUST. Special thanks are due to Professor K C Chan the Dean of the HKUST Business. We wish to thank Maya Rosenblatt and Maggie Chan, the conference secretaries, without their help the conference would have been a total chaos, and Chi Hang Chark for the splendid and dedicated work in preparing and formatting all the chapters for publication. We also thank Kluwer for supporting this project.

Chapter 1

THE RATIONALITY OF CONSUMER DECISIONS TO ADOPT AND UTILIZE PRODUCT-ATTRIBUTE ENHANCEMENTS: WHY ARE WE LURED BY PRODUCT FEATURES WE NEVER USE?

Shenghui Zhao
University of Pennsylvania

Robert J. Meyer
University of Pennsylvania

Jin Han
Singapore Management University

Abstract

The ability of consumers to optimally anticipate the value they will draw from new product features that are introduced to enhance the performance of existing technologies is explored. The work tests a hypothesis that when consumers are given the opportunity to buy a new generation of a products that offers enhanced features consumer will overvalue them, a bias the accrues to a tendency to overestimate both the extent that they will utilize these new features and the impact they will have on utility. This general hypothesis is tested in the context of a computer simulation in which subjects are trained to play one three different forms of an arcade game where icons are moved over a screen by different forms of tactile controls. Respondents are then given the option to play a series of games for money with either with their incumbent game platform or pay to play with an alternative version that offered an expanded set of controls. As hypothesized, subjects displayed an upwardly-biased valuation for the new sets of controls; adopters underutilized them and displayed a level of game performance that was not better than those who never upgraded. A follow-up study designed to resolve the process underlying the bias indicated that while adopters indeed over-forecast the degree to which they would make use of the new control, they did not over-forecast performance gains. Hence, the key driver of adoption decisions appeared to be an exaggerated belief of the hedonic pleasure that would be derived from owning and utilizing the new control as opposed to any objective value it might provide.

R. Zwick and A. Rapoport (eds.), Experimental Business Research, Vol. III, 1–33.

As consumers we have always had something of a love-hate relationship with new generations of products. On one hand, innovations that hold the promise of being the latest and best in a class of technologies often hold an allure that seems to go beyond the objective incremental benefits they provide. Manufacturers of new gaming systems seem never to produce enough units to meet initial demand, we brag about the multitude of features that endow our new cell phones (even if we never use them), and the wealthy compete to see who can fill their homes with the most advanced technological gadgets. Even those who lack the wealth to acquire technological enhancements are no less subject to their appeal; society surrounds us with images of innovations in magazines, television ads, and billboards.

Yet, it is equally transparent that whatever appeal consumers may see in acquiring new technologies, it is an appeal that has real limits. As attracted as we may be to the idea of acquiring that which is new and innovative, we are also often averse to incurring the switching costs that are often associated with adopting innovations – an effect cognitive scientists term *lock-in* (e.g., Norman 1998; Johnson, Bell, and Lhose 2003; Zauberman 2003). Hence the origin of Klemperer's (1987) paradox of the early innovator: individuals who are the first to adopt new technologies often turn into laggards, inhibited from keeping up with the pace of innovation by the need to constantly incur switching costs.

How do consumers balance these instincts when forming assessments of their willingness to adopt product innovations? The answer to this question is uncertain. On one hand, there is ample anecdotal evidence that would seem to support the often-heard claim that consumers over-estimate the degree that they will make use of enhanced features carried by new technologies. For example a 2003 *Harris Poll* revealed that 45% of cell phones owners never use voice mail features, and 50% have never exercised the option of setting their phones to silent or vibrate[1]. But the mere fact that consumers make limited used of the advanced features of new products, of course, does not necessarily imply that a forecasting error had been made at the time of purchase, or that they would be happier if they put them to greater use. An un-used feature may have been acquired simply because it was part of a sales bundle, or the feature may have been purchased for its option value. That is, only by acquiring the feature could the consumer learn whether they would be useful or not, or gain access to it at an uncertain later point in time. Finally, it should be noted that, by definition, the reciprocal error of *under*-forecasting is difficult to document; while it is easy to observe attributes that are purchased but never used, we never observe attributes that would have been used had they been purchased.

The purpose of this paper is to take a step toward resolving this research uncertainty by systematically investigating the quality of consumer decisions to adopt and then subsequently utilize innovative features in new products. We undertake our investigation in a controlled laboratory setting where subjects are trained to play a new arcade game for a monetary incentive where game tokens are moved using a certain set of computer controls. Subjects are then given the opportunity to purchase alternative versions of the platform that offer expanded sets of controls. The objective of the paradigm is to identify biases in consumers' willingness-to-pay for product attribute enhancements as well as how these attributes are subsequently utilized. In

addition, we also examine biases that arise in the reverse case where the product innovation offers a design *simplification.*

The core finding of the work is strong support for what might be termed an *enhancement bias* in new-product adoption decisions. When given the opportunity to purchase an enhanced game platform subjects reveal levels of willingness-to-pay that are greatly in excess of that which can be explained based on either their own best forecasts of score improvement or a simplified options-value analysis of the adoption decision. In essence, subjects act as if mere access to the new set of controls – regardless of their functional value – provides a source of prospective utility worth paying for. Yet, once this ability is in place few seem to utilize it; players who acquire the enhanced platform withdraw use of the new controls after overly-short periods of experimentation, and do not realize higher levels of performance compared to those who never had the chance to upgrade.

We organize our presentation of our research in three phases. We first develop a more complete background for the research by reviewing the normative basis for consumer new product-adoption decisions and exploring prior behavioral research that suggests how actual decisions may depart from this benchmark. We then test these hypotheses using data drawn from two laboratory experiments. We conclude with a general discussion of the implications of the work for both basic research in consumer response to product technologies as well applied work in new-product design.

1. THE PSYCHOLOGY OF NEW-PRODUCT ADOPTION DECISIONS

In this work we consider how consumers solve a class of new-product adoption problems that have the following structure. A consumer currently owns a durable good that conveys utility through the utilization of a set of features (such as options in software or capabilities of a home entertainment device). A manufacturer offers the consumer the opportunity to purchase an enhanced version of the good that retains the features of the old but also offers a new set of discrete attributes of uncertain value. The existence of these new attributes does not affect the functionality or utility derived of the older attributes, however they do compete for usage time. That is, the new attributes cannot be used simultaneously with the old. Hence, analogies might be software packages that provide users with the option to utilize either older or newer interfaces (similar to *Windows XP*), or digital cameras that give users the option to operate it with basic or advanced settings.

We can formally model the consumer's problem as follows. Assume that the utility that the consumer realizes from consuming an incumbent good with attribute α at any point in time t is scaled to be 0. Let $d_t \in \{0,1\}$ denote the consumer's decision whether or not to utilize some new feature δ given its ownership at time t, let $x_t = u(\delta) - c(\delta)_t$ be the net utility that is realized given a decision to utilize δ at t, and z_t denote the consumer's beliefs about the probability distribution associated with x_t^2. In addition, let $\pi_T = d_0, \ldots d_T$ be a sequence of attribute-usage decisions defined over a T-period ownership horizon, and let $V_0(\pi_t)$ be the total discounted expected utility implied by this sequence, defined as follows:

$$V_0(\pi_T) = E_0 \sum_{t=0}^{T} \beta^t v(x_t, z_t, d_t) \quad (1)$$

The decision maker's goal would then be to find that sequential decision policy π_T^* that maximizes expression (1), yielding an optimal ownership valuation ($V^* = V_0(\pi_T | \pi_T^*)$). The consumer would then be prescribed to buy the new product if $V^* > C$; that is, if the lifetime expected value of the new product that follows from assuming optimal utilization of the innovative feature δ exceeds good's purchase price.

It goes without saying that making a new-product adoption decision in this manner would be a formidable feat of cognition. One would need to possess good skills not only in intuitive dynamic programming (to derive the optimal ownership policy π_T^*), but also hedonic forecasting – accurately anticipating the various possible states of long-term pleasure one might come to associate with a new technology (the distribution over net asymptotic values of $u(\delta) - c(\delta)$) as well as how this pleasure may change over time in the course of ownership.

How potentially damaging would failure of these assumptions prove? On one hand, the literature is replete with examples of intuitive decisions that closely correspond with those prescribed by highly complex normative models (e.g., Hogarth 1981; Meyer and Hutchinson, 2001; Rust 1992). Yet, there is growing evidence that this same robustness may *not* extend to tasks – like the current – where decision makers are required to forecast their future preferences. Specifically, as skilled intuitive decision makers we may be in many domains, predicting how we will feel and act in the future does not appear to be one of them (see, e.g., Loewenstein and Schkade 1999; Wilson and Gilbert 2003). A core hypothesis of this research is that when making product-adoption decisions biases in hedonic forecasts will yield systematic inefficiencies in both the quality of initial decisions to buy new goods and their subsequent utilization after purchase.

We will briefly review lines of evidence that suggest systematic biases that may arise when consumers attempt to develop two kinds of forecasts that would be central to the normative solution to expression (1): forecasts of the mean potential value of an innovative attribute (beliefs about x and z); and forecasts of the dynamic utilization of the new attribute (beliefs about the decision policy π).

1.1. Intuitive forecasts of new attribute values

Assessing what one should be willing to pay for new product features is not an easy task. Such assessments should rationally reflect not just the pleasure one anticipates drawing from the feature over the expected future of ownership, but also the costs that will be incurred learning to use the feature, and, most critically, the long-term utility of *not* acquiring it; keeping the current device and spending the money on something else. How skilled will consumers be in making these kinds of assessments? While no work has examined this question directly, research that has examined the quality of human hedonic forecasts would not seem encouraging (e.g., Kahneman 1999; Loewenstein and Schkade 1999; Wilson and Gilbert 2003). Prior evidence suggests that not only will consumer assessments of the likely future value of attribute

innovations often depart from normative benchmarks, but that these departures will have a distinct bias: toward overvaluation.

The core argument is as follows. One could view the above normative framework as requiring consumers to hold three kinds of expectations when valuing product enhancements: an initial short-term expectation of the relative value offered by the innovation, an expectation of how these beliefs will evolve over time through ownership, and a belief about the option value of the attributes – the utility of being able to decide in the future *not* to use the feature if its value turns out to be limited. We argue that consumers will commonly systematically overvalue new-product features because of the cascading effect of congruent distortions in each of these judgments: a tendency to hold overly optimistic priors about value, under-assess the likelihood that pleasure may diminish in the future, and over-assess future option values.

Consider, first, the direction of affect consumers will first associate with a product innovation. There are strong normative and psychological arguments that predict that these assessments will routinely be positively biased, with consumers feeling a lure to acquire the new good that is not based in any objective knowledge of value. Common experience, of course, offers numerous anecdotes that would seem to support this idea: we are attracted to new rides at amusement parks and new flavors of ice cream, and are anxious to read about the latest innovations in computer technology. In many cases these kinds of reactions have a sound rational basis in information economics: one *should* be tempted to try new options that appear in markets because it is only through the experience of trial will we know which options will give us the highest utility in the future.

There is also evidence, however, that the lure consumers feel toward product innovations is triggered by more than curiosity: novel products also often are evoke heuristic expectations of heightened quality. To illustrate, Miller and Kahn (2003) offer data showing that merely affixing novel names to the color or flavor of an otherwise familiar product enhances its perceived quality among consumers. They suggest that the effect arises not as a result of a rational desire for information but rather by a simpler effect of conversational norms (Grice 1975). Given a communication that is seen as potentially ambiguous (in their case, a color or flavor name), consumers implicitly assume that it holds relevance to the purpose of the communication (conveying something about the nature of the product), and its valence is inferred from the presumed intended consequence of the communication (that the consumer would be more inclined to buy the good). In the case of innovative product attributes conversational norms would predict a similar result; even if consumers were not lured by curiosity, most would believe that if a firm took the time and effort to add new features to a good it was with the intention of enhancing its value.

This same effect is likely to be compounded by yet another documented bias in decision making: the tendency of individuals to overvalue options that allow for flexibility (e.g., Lowenstein and Alder 1996; Simonson 1990). Translated to product design, such a preference would reinforce a "more is better" heuristic in evaluating new product attributes: even if one suspects that that an expanded set of a feature offered by a product innovation carry little immediate value (e.g., an imbedded

camera in a cell phone), one might nevertheless desire having it as a hedge against future changes in preference or usage norms.

Of course, such assessments *per se* are far from wrong; recall that in a normative analysis the prospective value of an innovative product feature depends not just on the utility that one expects to receive from it given its use ($u(\delta)$), but also the option value of *not* using it. The problem comes from the fact that individuals routinely overvalue the merits of such flexibility.

For example, Simonson (1990) and Loewenstein and Alder (1996) report data showing that when consumers are asked to make a one-time choice of a basket of product flavors that will be consumed in the future they tend to choose a wider assortment than is actually consumed when these choices are made individually over time. Likewise, Shin and Ariely (2003) report a sequential search task where people are willing to pay to keep search routes open even when the odds that they will be utilized is small. Finally, Gilbert and Ebert (2002) and Wilson and Gilbert (2003) offer evidence that consumers tend to strongly prefer transactions that allow for revocability (e.g., liberal exchange policies), even when they are unlikely to be exercised. Hence, while there is indeed a rational basis for desiring products that offer a flexible assortment of features, the value that consumers place on this capability may be excessive.

Of course, one might argue that these kinds of visceral assessments of product value might fade once consumers begin thoughtful analyses of the real *net* utility they would draw from an innovation given its purchase price. Consumers might (and should) come to recognize, for example, that with these new features comes the cost of having to learn how to use them, and recall times in the past when they were lured to buy new goods in the belief that they would dramatically enhance pleasure, only to find that the enhancement was modest at most. Yet, the weight of evidence is that consumers will under-attend to these considerations, perpetuating a positive assessment bias.

Supporting this idea is empirical evidence that affective forecasts are often subject to what that Loewenstein, O'Donoghue, Matthew Rabin term a *projection biases*, a tendency to presume that one will feel in the future much as how one feels today. What seems to drive this bias is an effect that Wilson and Gilbert (2003) call *focalism*: when a decision maker is in one affective state it is difficult to imagine being in another, or project the preferences one will have at future points in time (see also Kahneman and Snell 1992). Gilbert, Gill, and Wilson (2002) and Read and van Leeuwen (1998) illustrate thus effect by showing there is real truth to the old adage that one should never shop on an empty stomach; shoppers who are hungry systematically buy more than those who are full, presumably due to inability to anticipate how they will feel in the future when they begin to consume the goods they are purchasing. Likewise, DellaVigna and Malmendier (2002) and Gourville and Soman (1998) offer evidence from health-club attendance patterns that people systematically underweight future costs in the form of effort. Specifically, subscribers pay large up-front fees to join a gym (implying high expectations of usage), but then underutilize it after joining, implying an under-forecast of the effort required to attend. The implication here is that while learning costs may ultimately play a major

role in influencing how new-product attributes are actually used, they will tend to be undervalued at the time product-adoption decisions are made.

Taken together, these streams of work suggest a straightforward hypothesis about how consumers will prospectively value new attributes carried by product innovations:

H1: The Innovation Bias. *When given the opportunity to purchase a new product that possesses an expanded set of attributes relative to an incumbent, consumers will display an overvaluation bias, revealing rates of adoption and levels of willingness-to-pay in excess of those would be justified by both actual subsequent utilization patterns and a rational a priori options valuation.*

The logic that underlies **H1** rests, on an assumption that the most salient initial reaction that consumers will have when exposed to a product innovation will always be that of optimism about its value. The degree to which this would hold in natural settings, of course, would be expected to vary from consumer to consumer. For example, a consumer who has recently incurred extremely high learning costs when adopting an innovation might have far more tempered – or even negative – visceral reactions to a product that offers yet another new set of features. In the same way that focalism predicts that optimistic consumers will be prone to underweighting future learning costs when valuing products, pessimistic consumers may be prone to underweighting future pleasure. Given this, we might expect that individual differences in difficulties encountered when learning to use new product features in the recent past could serve to moderate the general prediction in H1. Formally,

H1a: The moderating effect of past learning costs: *the mean tendency of consumers to overvalue product innovations will be moderated by past learning costs, with the bias being tempered among decision makers who have experienced associate steep learning curves with innovations.*

Now that they've bought it, will they use it?

Central to the work on affective forecasting that forms much of the basis of **H1** and **H1a** is the idea that biased forecasts arise because individuals are poor at anticipating how they will make decisions in a future world where the on-going judgment tasks and inputs are substantially different from those that are faced today. In the case of new-product judgments this disconnect would seem particularly acute. At the time of purchase the consumer's cognitive efforts are focused on solving a rather formidable normatively decision problem: that of whether the option value of acquiring a new generation of a technology is worth the purchase price, given assessments of the likely horizon of ownership, likely utilization over that horizon, and the affect associated with loss of the incumbent good and liquidity. But once an affirmative decision to acquire the innovation is made, cognitive efforts shift to solving a quite different – and seemingly much simpler–task: making moment-to-moment decisions about whether to make use of the innovative attributes of the good now that it is owned. These judgments, in turn, will be influenced by a range

of hedonic factors that were not salient at the time of the initial choice, such as the frustration of learning how to use a new product attribute, and the appeal of momentarily deferring this learning to a future time period during ownership.

The implication is that consideration of these new factors will not only lead to levels of attribute utilization that are below those envisioned at the time of purchase, but also below those that would maximize the absolute long-term utility of ownership. Specifically, when a consumer who has purchased an new product is deciding whether or not to try utilizing one of its new features the decision is not simply one of whether this action might yield long-term benefits, but whether these benefits – which are uncertain – will be higher than those afforded by continuing to use older, more familiar, attributes. A systematic finding of work on technological utilization is that when consumers have well-developed skills in utilizing one technology they often find it difficult to learn new ones, and are frequently averse to learning – an effect called termed cognitive lock-in (e.g., Johnson, Bell, and Lhose 2003; Norman 1998; Zauberman 2003). The usual explanation is that expertise with using one generation of a technology tends to increase as logarithmic function of practice (the power law), implying that the more familiar one becomes with one technology, the higher the short-term relative cost of learning to utilize new technologies (Klemperer 1987; Zauberman 2003).

The fact that new technologies involve switching costs, however, does not by itself imply that consumers will be prone to error in how they initially value technologies or how they utilize them once acquired. As we noted earlier, normative assessments of the value of innovations *should* anticipate such costs (through the consumer's beliefs about how $u(\delta)$ and $c(\delta)$ will evolve over time), and after purchase the observed magnitude of switching remain a normatively-relevant consideration in usage decisions. For limited utilization to be judged an *error*, therefore, the effect of switching costs on usage must be greater than what would be anticipated in a rational analysis.

Prior work on dynamic decision making in other contexts provides strong hints that processing of switching costs may well be biased in just such a manner. First, one of the most pervasive findings in the study of decision making over time is that people frequently undervalue the long-term benefits of learning and experimentation (see, e.g., Meyer and Hutchinson 1994; 2002). For example, in experimental armed-bandit tasks decision makers frequently cease gathering data on unfamiliar options after overly-short periods of experimentation (e.g., Meyer and Shi 1985), and melioration experiments find a similar aversion to making short-term costly investments when the benefits are long-run and distant (e.g., Herenstein and Prelec 1992). Hence, while consumers might well concede the long-term benefits of learning about new technologies at some abstract level, day-to-day decisions about the attribute utilization may be dominated by short term assessments of which product features yield the greatest benefit at the lowest cost – leading to underutilization.

A related influence that may further contribute to underutilization is the fact that in product-adoption settings learning is deferrable. In other words, for most consumers the decision about whether to take up learning about a new product feature is not one of whether it will *ever* be beneficial to learn (for most, the answer would

be, "probably yes"), but rather whether *now* is the best time to start. From a normative perspective, of course the answer to this question will always be "yes"; one should always want to resolve uncertainty as early as possible so as to allow the benefits of information can be realized over the longest-possible time horizon. Yet, this is an instinct that is often lost on real decision makers (e.g., Meyer and Hutchinson 2002).

Specifically, there is extensive evidence showing that when individuals are presented with a choice between a set of uncertain alternatives versus deferral, growing indecision leads to a growing preference for postponement (e.g., Dhar 1997; Tversky and Shafir (1992). Hence, one might speculate that the more consumers are unsure whether the benefits offered by a new attribute are worth the learning costs, the greater will be their urge to delay the onset of experimentation. What is particularly attractive about delay in this context is that it allows consumers to mentally justify the short-term action of utilizing familiar attributes while still retaining the abstract goal of wanting to learn new technologies. By deferring one is not abandoning this long-term commitment, just delaying its onset to an unspecified future time when costs be lower (e.g., "I'll read the manual over the weekend").

A final factor that would contribute to under-utilization errors is if learning and usage costs at the time of initial purchase turn out to be much larger than was anticipated. In some cases this under-forecast will arise due to the inherent difficulty that comes from envisioning future affective states that we discussed above (e.g., DellaVigna and Malmendier 2002 and Gourville and Soman 1998). But an even more acute basis for under-forecasts would be if consumers use an inappropriate analogic-reasoning process to generate expectations about learning costs. That is, assume that knowledge about product usage gained in one domain can be directly transferred to the new one to greater degree than is the case (e.g., Moreau, Lehmann, and Markman 2001). While learning-by-analogy can often greatly reduce learning costs, it can also substantially raise them if the assumed analogies prove inappropriate; for example, assuming that short-cuts useful in one text editor holds for others (e.g., Norman 1988). In such cases consumers have the added burden of not just learning how to use the new technology, but also *un*learning interfering mappings to old ones – mappings for which they may be unaware (e.g., Wood and Lynch 2002). A relevant suggestive illustration of this effect has recently offered by Zauberman (2003), who reports data showing that people tend to over-forecast how productive they will likely be using new web interfaces, implying that the difficulties involved in switching to new formats went largely unanticipated.

Taken together, these discussions lead to the following general hypothesis about post-purchase utilization of new-technology attributes:

H2: The Under-Utilization Bias. *Given a decision to acquire an innovation that possesses a mixture of innovative and familiar attributes, utilization of the new attributes will be- downwardly biased relative to the levels implied by stated wiliness-to-pay for the good, direct forecasts of benefits, and objective benefits that would come from optimal usage.*

An intriguing consequence of the discussion we offered about how projection biases might influence both prior new-product valuations and subsequent utilizations is that it implies a possible paradox in how individual differences in post-purchase attribute utilization might relate to pre-purchase willingness to pay. In **H1a** we proposed that consumers who had more positive experiences when consuming past technologies would produce the most optimistic assessments of the prospective value of a new product that offered an attribute innovation. Yet, because much of this optimism comes from the under-forecasting of learning and switching costs (as above), it is these same consumers who would most likely experience the greatest disappointment when they come to utilize attribute innovation that they paid for. This disappointment, in turn, would lead to more rapid decisions to abandon use of the new attributes relative to those who entered ownership with more modest expectations. We summarize this idea in following hypothesis:

H2b The Paradox of the Technological Optimist: *Consumers who reveal the greatest optimism in their willingness to pay for a technological innovation will also be the most prone to abandon trial usage of attribute innovations given ownership.*

2. EMPIRICAL ANALYSIS

2.1. Overview and Design Consideration

In this section we describe the results of three experiments designed to test the empirical validity of the research hypotheses summarized in **H1**, **H1a**, **H2**, and **H2a**, as well as provide descriptive insights into the process by which consumers make decisions to buy and then subsequently utilize product innovations. These issues were examined by observing how a sample of experimental subjects learned to play an original arcade-like computer game where performance was rewarded by a monetary incentive. After a period of training with one of several basic platform designs subjects were given the opportunity to purchase an enhanced platform that offered a combined set of features that were drawn from the basic platforms. In a third experiment we examine the reciprocal case: subjects trained on the enhanced platform are given the opportunity to exchange it with a reward for a simplified platform containing only them most-used controls.

The game was called "Catch'em" and bore similarities to the popular late 70's, early 80's arcade game *Pac Man*. In the game players viewed a square grid on which, at the start, was superimposed a number of stationary green dots called "cookies". Also on the grid were two larger red and black dots that depicted the staring position of the player and his or her robotic opponent, termed the "Monster". Upon triggering the start of the game both the Monster's and player's icons began moving over the grid. While the Monster moved at a random speed and direction, the player controlled the speed and direction of his or her icon. Each time either the player's icon (or the Monster) moved over a cookie a point was scored for the player

(or the Monster). If all of the cookies were consumed from the board by the player and/or the Monster, the play ended and the player received a point total equal to the number of cookies he or she had captured. If, however, at any point the Monster's icon touched the player's icon, the player's icon was declared "caught" and play also ended, with all points having been earned to that point being forfeited. The basic board layout and instruction are reproduced in Appendix 1.

We chose this – admittedly unusual – stimulus context because it was one that satisfied four ideal design criteria:

1. It provided us with experimental control over the design and familiarity subjects had with a basic generation of a technology;
2. It allowed experimental introduction over the value of enhanced features in a new technology;
3. It provided a natural objective for measuring performance that could be used for providing a monetary incentive to subjects; and
4. The task context – an arcade game – was one that was likely to be seen as highly involving and familiar to the subject pool, primarily undergraduate college students.

The technology in this case was the nature, complexity, and quality of the controls available to subjects for moving their icon. A basic technology was one where subjects had access to only one kind of control at one calibrated level of performance, while the enhanced technology was one where subjects had access to multiple controls – both those with which they were familiar and a "new" set that was derived from one of the other basic models (the existence of which was unknown to subjects).

Our analysis focuses on the results of two experiments conducted within this paradigm. The purpose of Experiment 1 was to conduct a basic test of the four hypotheses in a setting where there was minimal measurement intervention; we observed learning paths, the dynamics of control utilization, and adoption decisions in the absence of direct elicitations of either forecasts of behavior or elicitations of reasons for decisions – interventions that might influence behavior. In Experiment 2 we attempt to more deeply probe the process that underlies the data uncovered in Experiment 1 by gathering such process measures.

3. EXPERIMENT 1

3.1. Design, Subjects, and Procedure

Subjects were 149 business-school undergraduates who volunteered to complete the task for a monetary incentive. Subjects performed the experiment seated in computer cubicles in the school's behavioral research lab. At the outset of the experiment subjects were told that the purpose of the experiment was to learn how consumers such as themselves learned to play gaming devices, and that they would be paid depending on their performance in the game. Subjects were told that there would be

a show-up fee of $5 (US) per subject, and they could earn up to $10 more depending on how well they learned to play the game.

All subjects were told that they would be playing the "Catch'em" game a total of 30 times, with the first 15 being practice rounds that would not count toward their final earnings, and the second 15 being money rounds on which their pay would be based. After reading this basic instruction subjects were randomly assigned to either a control or treatment condition, with which they were also assigned to play one of three different basic game platforms (described below). Subjects in the control condition played with the same platform over all 30 rounds of the experiment. Subjects in treatment condition played the first 15 training rounds with one platform, but were then given the opportunity to pay to play the money rounds with a new platform that offered a broader range of controls. The opportunity to pay to switch to a new platform was offered only once; if a subject declined the purchase he or she played the 15 money rounds with the same game platform that they trained on, the same as those in the control condition.

The game platforms. The three basic game platforms on which subjects trained on were defined by the physical form and reliability of the controls used to move the player's icon. There were three mechanisms:

1. *A Scroll Bar Control (Figure 1a)*: Subjects continuously adjusted the speed and direction of movement of their icon by moving each of two horizontal scroll bars displayed on the computer screen. Use of the directional control was aided by a steering-wheel-like graphic that displayed the current directional heading of the icon.

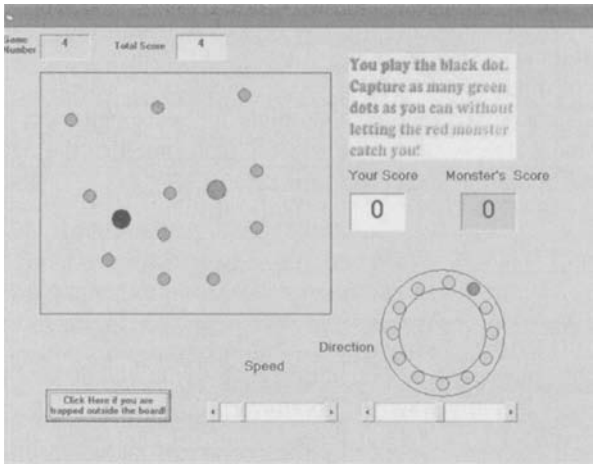

Figure 1. The Three Game Platforms.
1a: Scroll-Bar Control

2. *A Button Control with high reliability (Figure 1b).* Subjects adjusted speed and direction by repeatedly clicking two sets of button controls. One pair of buttons allowed subjects to reverse the current heading of their icon either horizontally or vertically, while the other pair induced discrete increases or decreases in speed. High reliability meant that the icon's movement responded 80% of the times to player actions in the intended manner given activation of any control.

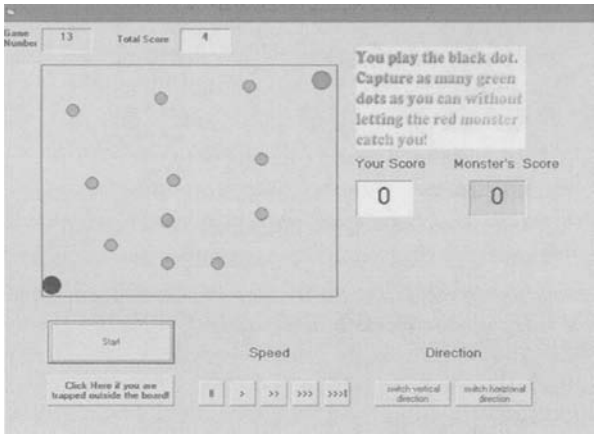

Figure 1b. Button Control.

3. *A Button Control with low reliability (Figure 1b).* The appearance and function of this platform was identical to (2), except that random noise was added to the

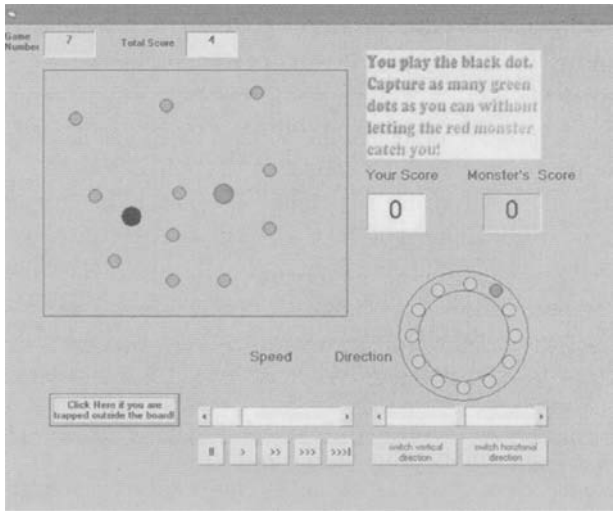

Figure 1c. The Enhanced Platform: Combined Controls.