Psychology of Decision Making in Economics, Business and Finance

Klaus P. Hofmann
Editor

NOVA

PSYCHOLOGY OF DECISION MAKING IN ECONOMICS, BUSINESS AND FINANCE

PSYCHOLOGY OF DECISION MAKING IN ECONOMICS, BUSINESS AND FINANCE

KLAUS P. HOFMANN
EDITOR

Nova Science Publishers, Inc.
New York

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Psychology of decision making in economics, business and finance / KlausP. Hofmann, editor.
p.cm.
Includes bibliographical references and index.
ISBN- 13; 978-1-60692-568-3
1. Decision making. 2. Economics--Psychological aspects. I. Hofmann, Klaus P.
HD30.23.P79 2007
330.01'9--dc22

2007029215

Published by Nova Science Publishers, Inc. ✣ New York

CONTENTS

PREFACE

In a fast-moving world, the necessity of making decisions, and preferably good ones, has become even more difficult. One reason is the variety and number of choices perhaps available which often arenot presented or understood. Alternatives are often unclear and complex paths to them confusing and misleading. Thus the process of decision making itself requires analysis on an ongoing basis. Decision making is often made based on cultural factors whereas the best alternative might be quite different. The subject touches ethics aspects as well as psychological considerations. This new book presents important research on the psychology of decision making related to economics, business and finance.

Expert Commentary - Sustainable management of natural resources necessitates conscious and organised activity. When planning the use of a certain spatially explicit management unit, competing goals and interests call for multi-objective decision support. The environmental decision making also involves uncertainty concerning available information and future predictions. Traditionally, a typical procedure has included the use of preference elicitation tools, statistical models, and optimisation methods, to find an acceptable solution for the decision problem at hand.

When improving the decision support portfolio to work more efficiently, there is a need to consider the overall structure of decision problems. To demonstrate that, we have conceptualised the variety of forestry decision problems in three-dimensional decision-problem space, which comprises geographical, temporal, and social dimensions. The crucial issue in the problem structuring and, thus, in the selection of decision support method, is scope awareness: recognising the fundamentals of the decision problem at hand. With the aid of that, the unwanted decision-model-based or support-methodology-based psychological biases can be avoided or smoothened.

By the results of the research proposed in this chapter, both the efficiency of single phases of decision making and the quality of the whole decision process would be improved. The results would also enhance learning about the characteristics of decision problems and decision makers' preference structures.

Chapter 1 - We report a study in which methodologies from psychophysics are adapted to investigate context effects on financial decision making related to retirement savings and risky investment. The aim was to determine how the range of the options offered as possible saving rates and levels of investment risk influences decisions about these variables. The respondents were presented with either a full range of choice options or a limited subset of the feasible options. The study was conducted on a sample of working people, and we controlled

whether the participants can financially afford in their real life the decisions taken in the test. The results showed that choices of saving and risk are affected by the position of each option in the range of presented options. Various measures of risk aversion did not account for the risk taken in each condition. Only the simplest and most direct risk preference measure was a significant predictor of the responses within all contexts (conditions), although the actual choices were still very much influenced by each context. Thus, the results reported here suggest that judgments and choices are relative, rather than absolute, which corroborates, in a more applied and realistic setting, previous related work with abstract gambles and hypothetical risky investments.

Chapter 2 - Self-evaluations of performance are important in theory and practice. In contexts with multiple persons performing the same task, the evaluation of one's own performance is expected to be a process involving judgments about the performance of others, and comparisons between one's own and others' performance. We conducted a longitudinal study tracking 79 participants' evaluations of their own and others' performance on five repetitions of a task over a four-month period. Three temporal factors that Radhakrishnan, Arrow, and Sniezek (1996) identified as influences on self evaluations of performance were examined: Temporal Perspective, Time Horizon, and Experience. In the present study, we investigated in more detail, the role of these factors, on judgments evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance as well as on social comparisons. Participants made evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance and social comparisons both had a positively leniency bias. This bias in self evaluations and social comparisons decreased when estimates were made (a) after performance than before; (b) closer to the performance event than farther away from it; and (c) with increasing experience. However, evaluations of only one's own performance were more variable with changes in the temporal factors. Further, the increase in bias with longer time horizons was reduced considerably with increasing experience. Changes in inter- and intra-individual validity followed those for bias. Interestingly, changes in solo evaluations over time were similar to those for social comparisons.

Chapter 3 - The aim of this paper is to draw attention to what is arguably a very general and pervasive feature of human cognition that may have important implications for our understanding of human decision making and also for some aspects of economics. The major claim, defended here, is that when people judge the attributes of choice options (like utilities, payoffs, and probabilities), they are not able to represent the absolute magnitudes of these attributes; instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. Also, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other sampled magnitudes. Such sampling of knowledge from memory and transferring it to the current situation produces certain biases in judgment because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there may be no ability to represent cardinal scales, for any magnitude and judgments involving such magnitudes are determined by the context. The core evidence for this claim comes from recent research in psychophysics on the perception of the intensity of basic psychophysical magnitudes such as the brightness of a

light or the loudness of a sound, and also from research on the effects of context on decision making under risk and uncertainty.

Chapter 4 - The vast majority of the public learns about new risks to health and society predominantly from the media, including the press media directly or indirectly. However, little is known about the role and mechanisms through which the press media influences attitudes and risk perceptions. Some approaches stress the ides that risks are partly created while other state that the media plays a neutral role, however empirical evidence is hard to retrieve and still is scarce. This paper empirically examines both the role of press media coverage and reporting of new genetically modified (GM) foods between 1999-2004. We draw upon a combination of qualitative and quantitative evidence. First, evidence of content analysis of key press media in two countries - Spain and the United Kingdom (UK) - is examined to illustrate preliminary evidence and subsequently, quantitative evidence of survey data (Eurobarometer surveys) is examined to scrutinise for the existence of some media biases, inter-country differences in public perceptions as well as specific media effects connected to role of journalism in the country. Results point towards the existence of significant differences in media reporting and respect for journalism between the two countries, which correlate with public perceptions, although a similar lack of trust was identified. Furthermore, we find evidence suggesting some specific media biases depending on the press media readership.

Chapter 5 - When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints.

Risky and inter-temporal decisions have been captured theoretically primarily by Ecology and Microeconomics but findings from Behavioral Economics, Psychology and Neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive models has emerged to explain the findings. A subset of these models has stressed the similarities between risky and inter-temporal decisions. In this chapter we review both core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

Chapter 6 - While remaining within the traditional micro-economic framework of rational utility maximization, we enrich the standard and random parameters logit choice models with perceptions data. From the estimated models we derive a value of time and we also make a tentative attempt to derive a value of safety. Because we estimate the values simultaneously, we are able to explore whether values estimated in conjunction differ from values estimated in isolation. Survey data is used to measure the individual’s perceptions of five modal attributes (time, cost, safety/risk, environmental friendliness and flexibility) and show how these perceptions affect the modal choice for work trips. The respondents’ perceptions are elicited by a novel approach in which the names of two modes (car and bus) are used as attribute levels instead of objective levels. A difference between our survey and traditional ones is that we do not attempt to educate the respondents about, for example, the risks of travelling. Instead, we *record* the respondent’s perceptions about the risk and the other the modal attributes.

Chapter 7 - Economics has always focused on how individuals make decisions. Traditionally, the discipline has viewed individuals as rational agents maximizing their own

utility. However, economists have recently begun to incorporate research from the field of psychology in creating a richer view of decision making. This push is the result of challenges to the neoclassical model made by theoretical advances such as Kahneman's Nobel-winning prospect theory model and from the field of experimental economics. This growing field has revealed many aspects of human behavior that cannot be explained by traditional economic models. Some of these aspects of behavior include loss aversion (as explored by Kahneman); relative deprivation (the theory that individuals consider their relative position as compared to others when making decisions), motivations of altruism, fairness, and reciprocity; and the endowment effect (individuals tend to value goods more highly if they are already in possession of them). These innovations have impacted economists' views of issues such as consumption, worker-firm relations, labor supply, equities and real estate.

This paper reviews the impact of psychology on economic models of decision making. The major trends will be discussed, along with implications that these changes have for both economics and public policy.

Chapter 8 - Today more than ever, people try to anticipate financial needs and to plan wisely for a lifetime of financial security. Information about financial options is plentiful, and financing for health and long-term care (LTC) is no exception. With all of the information and advice that is available, under what circumstances would a person decide that his/her decision was no longer the best option?

We address this question by looking at the market for LTC insurance, and estimate logistic regressions to model consumer decisions to drop or renew an existing LTC insurance policy. We explore events that occurred after the policy was last purchased and before the current policy was dropped or renewed. The price and benefit design of each policy is not directly observable so several proxy measures of the price of a policy are explored.

Data is obtained from the publicly available Health and Retirement Survey (HRS). Data from 2002 is used to identify those who have a LTC policy and to establish baseline financial circumstances. Data from 2004 is used to determine whether the policy was renewed, and to identify potentially influential events that occurred since 2002.

The study sample includes 1,375 individuals who reported an existing, private LTC insurance policy in 2002, and were therefore eligible to renew the existing policy before 2004. Proxy prices were calculated and assigned using publicly available price schedules.

Preliminary findings suggest that price was an influential factor in the decision to drop an existing policy, even though the price of the policy did not increase as a result of age. Those with newer policies were less likely to allow a policy to lapse. Those with low levels of assets (less than $200,000) were more likely to allow a policy to lapse, as were those with more than $1.5 million in assets.

Our results suggest that financial considerations are important, and a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC.

Chapter 9 - The acceptance of risks associated with new technologies is a key issue that is likely to limit the extent of innovation in a 'risk society'. However, given the limited comprehensible information available to the public of new technologies, it is likely that risk information provision will have a heterogeneous effect on public perceptions. In order to examine this issue, we empirically examine the determinants of risk perceptions, benefit perceptions and risks acceptance of new technology developments in Spain. Our findings indicate that risk and benefits perceptions are not independent but affected by common

information sources. Furthermore, by taking into account this effect individual's knowledge of science heterogeneously increases both risks and benefits perceptions.

Chapter 10 - The present theory proposes that investors not only think of future monetary benefits, but also value the choices' implications regarding their self-esteem in decision making. Self-esteem is one's subjective evaluation of the self. Most people want to maintain a positive self-image. When they decide to invest in a project, people expect to receive financial rewards, and they also hope to enhance their self-esteem through the success of the project. Thus, when their initial investment produces negative economic return, they not only suffer financial loss, but also encounter challenges to their self-esteem. They can withdraw from the project to minimize their monetary loss, or they may keep throwing additional money into the project to demonstrate that their initial decision was correct. It is painful to admit a mistake because it poses negatively to the investors' self-concept. As a result, investors may be entrapped within a losing project and suffer accumulated financial loss. The present theory suggests that when investors encounter conflicts between money and self-esteem in decision making, they may choose to give up money in order to defend their self-esteem.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 1-6

Expert Commentary

PSYCHOLOGICAL ASPECTS OF DECISION MAKING AS NEW RESEARCH TOPICS IN NATURAL RESOURCES MANAGEMENT

T. Hujala[1], T. Kainulainen[2] and P. Leskinen[2]
[1]Department of Forest Resource Management,
P.O.Box 27, FI-00014 University of Helsinki, Finland.
[2]Finnish Forest Research Institute Metla, Joensuu Research Unit,
P.O. Box 68, FI-80101 Joensuu, Finland.

ABSTRACT

Sustainable management of natural resources necessitates conscious and organised activity. When planning the use of a certain spatially explicit management unit, competing goals and interests call for multi-objective decision support. The environmental decision making also involves uncertainty concerning available information and future predictions. Traditionally, a typical procedure has included the use of preference elicitation tools, statistical models, and optimisation methods, to find an acceptable solution for the decision problem at hand.

When improving the decision support portfolio to work more efficiently, there is a need to consider the overall structure of decision problems. To demonstrate that, we have conceptualised the variety of forestry decision problems in three-dimensional decision-problem space, which comprises geographical, temporal, and social dimensions. The crucial issue in the problem structuring and, thus, in the selection of decision support method, is scope awareness: recognising the fundamentals of the decision problem at hand. With the aid of that, the unwanted decision-model-based or support-methodology-based psychological biases can be avoided or smoothened.

By the results of the research proposed in this chapter, both the efficiency of single phases of decision making and the quality of the whole decision process would be improved. The results would also enhance learning about the characteristics of decision problems and decision makers' preference structures.

Keywords: *adaptive decision analysis, decision support systems, measurement scales, mental patterns, preference elicitation, psychological biases, uncertainty.*

BACKGROUND

In natural resources management, the intertwining of global and local, as well as public and private perspectives evokes the need for more sophisticated decision support methods. These methods would contribute e.g. to collaborative planning (e.g. Leskinen 2006), conflict management (e.g. Niemelä et al. 2005, Losa and Belton 2006), biodiversity maintenance (e.g. Kurttila et al. 2006), and legitimacy of environmental and forest policies (e.g. Rantala 2004). All of these aspects imply that a decision-makers' subjective views should be taken into account thoroughly, with the aid of cognitive and social psychology, as well as game theoretical aspects (e.g. Bender and Martin 2003). By the term decision-maker we do not refer merely to national, regional or local politicians or business managers but also to landowners, representatives of interest groups, and ordinary citizens, who participate planning or comparable decision making processes in some way or another.

The properties of decision problems as well as the understanding, motives, and commitment of decision-makers vary largely in different cases. Therefore, a traditional decision support system with fixed parameters and properties is susceptible to fail in fulfilling its expectations. In order to respond to this challenge we have enhanced the grounds for *adaptive decision analysis*, which adapts flexibly to the requirements of the output data and availability of the input data, taking into account the social system of decision considerations. With the adaptive approach, it is possible to select and apply appropriate decision models for different types of decision-problems and decision-makers.

A useful decision support system (Sprague 1993) helps the decision-maker to make better decisions or to find acceptable solutions more comfortably. If the system is transparent and easy to understand, the usage of such system results in learning about the problem (e.g. Buchanan 1994) and in improved commitment to the decisions. When the decision problem is straightforward and has a simple structure, our previous research (e.g. Hujala et al. 2007) shows that merely an illustration and discussion about the decision alternatives may be sufficient decision support to the decision-maker. Otherwise, e.g. when the decision-problem is about a large geographical area with a long planning period, cardinal, i.e. interval (e.g. Keeney and Raiffa 1993, Kainulainen et al. 2007) or ratio scale preference, elicitation techniques (e.g. Saaty 1980) are advisable.

The advantage of the cardinal techniques is that they offer a more sophisticated view on the problem, including the analysis of trade-offs between the conflicting decision criteria (c.f. Leskinen 2001). In addition, cardinal methods enable versatile possibilities for uncertainty analysis (e.g. Alho et al. 2001), which is an essential feature in natural resource management with complex structures. However, when reaching efficiency and acceptability of the solutions, profound computational methods will require also qualitative conversational support, especially in structuring the decision problems and assuring the decision-makers' understanding of the results.

DISCRETION REQUIRED IN USING ELICITED PREFERENCES

The mental aspects of preference elicitation can have significant impacts in decision making. Psychological biases in decision making are well-documented in the behavioral decision making literature (c.f. Hammond et al. 1998, Pöyhönen et al. 2001). Also, the performance of multi-criteria decision making (MCDM) methodology has been examined from the behavioral perspective. As for the results, *anchoring effects* have turned out to be a common type of psychological bias in interactive MCDM methods. For example, Buchanan and Corner (1997) identified that the starting point of decision process affects the weight assessment in the interactive Zionts and Wallenius (1983) method. In cardinal MCDM methods, a specific alternative or a certain interval is used as a reference for the preference elicitation process. One interesting issue for future research would be to examine whether the changing of the reference alternative impacts the outcome of the analysis. If such effect can be observed and verified in experiments and sensitivity analyses, some anchoring-based scaling recommendations for preference elicitation could be derived.

Another important topic for future research would be to conduct a systematic preference elicitation test, in which both interval and ratio scale models would be tested through direct weighting and pairwise comparisons techniques. It would then be interesting to examine whether the different methods lead to systematic differences in the estimated preferences (c.f. Belton 1986, Hujala and Leskinen 2006) and uncertainty measures. In addition, the rationale of individual decision-makers' thinking could be explored through qualitative interviews and direct observations. Implications of the empirical findings could then be drawn back to the design of preference enquiries.

INTRODUCING DECISION-PROBLEM SPACE

In order to respond to the requirements of taking into account decision-makers' perspectives in natural resource management, we hereby introduce a concept of *decision-problem space* (Figure 1). The space illustrates, in the context of forestry decision making, the potential diversity of decision-problems. In geographical dimension, one may focus on a single tree, forest compartment, an estate, region etc. The time horizon may be only one season, a couple of years, a decade or two, sometimes a century, or even more. The people who participate in the process form the social dimension of decision making. We measure the social dimension by the depth of participation and by adapting Arnstein (1969).

The scales in different dimensions are dependent, i.e. on the applicability of a decision support method in one certain dimension that is simultaneously affected by the case-specific scales in the other two dimensions. Thus, we can not study decision aid needs in one dimension at a time, independently of the other two dimensions. On the other hand, because decision support systems are tools to be used by humans, we consider that it is especially critical to study the methodological problems related to the social dimension.

In addition to comparing single points in the decision space, it might be beneficial to study decision scales hierarchically so that decisions made at one level are dependent upon decisions or information at other levels.

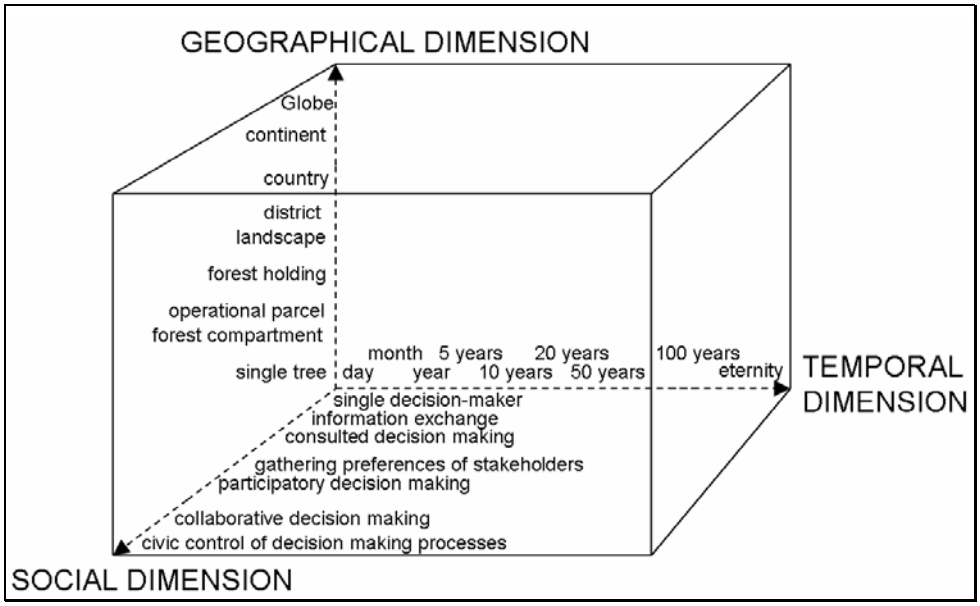


Figure 1. Illustration of three-dimensional decision-problem space (case forestry decisions).

The aim in hierarchical planning is to find formulations that preserve consistency between the decision levels. Production of information that supports finding solutions for problems arising from violated constraints, inefficient solutions, or infeasible problem formulations between or inside the levels of hierarchy are important (e.g. Davis and Liu 1991, Weintraub and Cholaky 1991, Davis and Martell 1993). The hierarchical analysis is important e.g. when projecting operational actions from strategic goals, implementing national policy in regional level, or compiling aggregate goals from the preferences of several individuals. The decision hierarchy may be both top-down- or bottom-up-oriented. It will be crucial that the available decision support methods both enable and support the hierarchical decision analysis.

CONCLUSION

We emphasise that it is important to attain and distribute new knowledge of decision-makers' perceptions of various decision support techniques and processes, especially the psychological biases that emerge within decision-making processes that need anticipation and smoothening. The research should open-mindedly take advantage of mixing quantitative and qualitative approaches. Developers and facilitators of decision support systems as well as decision-makers themselves need state-of-the-art recommendations for more convenient and transparent decision making. These guidelines would help decision-makers in natural resource management to meet the contemporary challenges such as multi-objectivity, interactivity, collaboration, and risk management. They would also improve learning possibilities in the context of decision making about different scales and facilitate the legitimacy of policies related to all levels and disciplines of natural resources management.

REFERENCES

Alho, J.M., Kolehmainen, O., Leskinen, P. (2001). Regression methods for pairwise comparisons data. In: Schmoldt, D.L., Kangas, J., Mendoza, G.A., Pesonen, M. (eds.). *The Analytic Hierarchy Process in natural resource and environmental decision making*. Kluwer Academic Publishers, Dortrecht. p. 235–251.

Arnstein, S., (1969). A ladder of citizen participation. *Journal of American Institute of Planners* 35: 216–224.

Belton, V. (1986). A comparison of the Analytic Hierarchy Process and a Simple Multi-attribute Value Function. *European Journal of Operational Research* 26: 7–21.

Bender, H.W., Martin, W.E. (2003). Modelling conflict using spatial voting games: an application to USDA Forest Service lands. *International J. of Global Environmental Issues* 3:149–167.

Buchanan, J.T. (1994). An experimental evaluation of interactive MCDM methods and the decision making process. Journal of the Operational Research Society 45: 1050–1059.

Buchanan, J.T. and Corner, J.L. (1997). The effects of anchoring in interactive MCDM solution methods. *Computers and Operational Research*, 24: 907–918.

Davis, L.S., Liu, G., (1991). Integrated forest planning across multiple ownerships and decision makers. *Forest Science* 37: 200–226.

Davis, R.G., Martell, D.L. (1993). A decision support system that links short-term silvicultural operating plans with long-term forest-level strategic plans. *Canadian Journal of Forest Research* 23: 1078–1095.

Hammond, J. S., Keeney, R. L. and Raiffa, R. (1998). The hidden traps in decision making. *Harward Business Review*. Sept.-Oct.: 47–58.

Hujala, T. and Leskinen, P. (2006). Preferences in interval and ratio scales - An experiment with non-industrial private forest owners. Oral presentation at EUROXII European Conference on Operational Research 2-5 July, 2006. Reykjavik, Iceland. Abstract published in: Plà, L.M. (ed.) (2006). Proceedings of the Second Meeting of the EURO Working group on Operational Research (OR) in Agriculture and Forest Management. *Journal of Agricultural Science*, 144, 467–472.

Hujala, T., Pykäläinen, J., Tikkanen, J. (2007). Decision-making among Finnish non-industrial private forest owners – the role of professional opinion and desire to learn. To appear in *Scandinavian Journal of Forest Research*.

Kainulainen, T., Leskinen P., Korhonen, P., Haara, A. (2007). *A Statistical approach to assessing interval scale preferences in discrete choice problems* (submitted).

Keeney, R. L., Raiffa, H. (1993). Decisions with Multiple Objectives. *Preferences and Value Tradeoffs*. Cambridge University Press, Cambridge.

Kurttila, M., Pykäläinen, J., Leskinen, P. (2006). Defining the forest landowner's utility loss compensative subsidy level for a biodiversity object. *European Journal of Forest Research* 125: 67–78.

Leskinen, L.A. (2006). Purposes and challenges of collaborative forest planning. *Working Papers of the Finnish Forest Research Institute* 38: 66–69.

Leskinen, P. (2001). Statistical methods for measuring preferences. University of Joensuu. *Publications in Social Sciences* 48. D.Sc. thesis. 111p.

Losa, F.B., Belton, V. (2006). Combining MCDA and conflict analysis: an exploratory application of an integrated approach. *Journal of the Operational Research Society* 57: 510–525.

Niemelä, J., Young, J., Alard, D., Askasibar, M., Henle, K., Johnson, R., Kurttila, M., Larsson, T.-B., Matouch, S., Nowicki, P., Paiva, R., Portoghesi, L., Smulders, R., Stevenson, A., Tartes, U., Watt, A. (2005). Identifying, managing and monitoring conflicts between forest biodiversity conservation and other human interests in Europe. *Forest Policy and Economics* 7: 877–890.

Pöyhönen, M., Vrolijk, H. and Hämäläinen, R.P. (2001). Behavioral and procedural consequences of structural variation in value trees. *European Journal of Operational Research* 134: 216–227.

Rantala, T. (2004). Conceptions of democracy of key informal interest groups in Finnish forest policy. In: Andersson, F., Birot, Y., Päivinen, R. (eds.). (2004). Towards the Sustainable Use of Europe's Forests – Forest Ecosystem and Landscape Research: *Scientific Challenges and Opportunities. EFI Proceedings* No. 49.

Saaty, T.L. (1980). *The Analytic Hierarchy Process*, McGraw-Hill, New York, NY.

Sprague, R.H. (1993). A framework for the development of decision support systems. In: R.H. Sprague, Jr. and H.J. Watson (eds.). (1993). *Decision Support Systems* (3rd ed.): putting theory into practice; pp.3–28. Prentice-Hall, NJ, USA.

Weintraub, A., Cholaky A. (1991). A hierarchical approach to forest planning. *Forest Science* 37: 439–460.

Zionts, S. and Wallenius, J. (1983). An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. *Management Science* 29 (5): 519–529.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 7-36

Chapter 1

RELATIVITY OF FINANCIAL PREFERENCES: HOW CHOICE OPTIONS INFLUENCE INVESTMENT DECISION MAKING

Ivo Vlaev [*] ***and Nick Chater***
Department of Psychology
University College London
26 Bedford Way, London, WC1H 0AP
United Kingdom

ABSTRACT

We report a study in which methodologies from psychophysics are adapted to investigate context effects on financial decision making related to retirement savings and risky investment. The aim was to determine how the range of the options offered as possible saving rates and levels of investment risk influences decisions about these variables. The respondents were presented with either a full range of choice options or a limited subset of the feasible options. The study was conducted on a sample of working people, and we controlled whether the participants can financially afford in their real life the decisions taken in the test. The results showed that choices of saving and risk are affected by the position of each option in the range of presented options. Various measures of risk aversion did not account for the risk taken in each condition. Only the simplest and most direct risk preference measure was a significant predictor of the responses within all contexts (conditions), although the actual choices were still very much influenced by each context. Thus, the results reported here suggest that judgments and choices are relative, rather than absolute, which corroborates, in a more applied and realistic setting, previous related work with abstract gambles and hypothetical risky investments.

Keywords: *context effects, decision making, judgment, investment risk, saving*.

[*] Correspondence concerning this article should be addressed to Ivo Vlaev, Department of Psychology, University College London, London, WC1H 0AP, United Kingdom. E- mail: i.vlaev@ucl.ac.uk.

This article presents a study that investigated how far, simple, and practically relevant, modifications in the decision making context can affect the way people make financial decisions related to retirement savings and investment. Specifically, we aim to see to what extent a new and powerful context effect, prospect relativity (Stewart, Chater, Stott, and Reimers, 2003), is likely to be important in practical settings. The applied objective of the experiment presented here is to test whether contextual factors can be used to stimulate financial consumers to save more for retirement and, also, to encourage them to take higher investment risks in relation to their retirement savings products in order to obtain the long-term benefits from higher expected returns in the stock market. The relevance of this objective is particularly significant given the major demographic shift that is occurring across the industrialized and post-industrialized world. Rising life-expectancy implies that growing, aging population will live longer after retirement and they would need substantial income to support their lifestyles. Without growing younger working population, our aging society requires working people to save more for retirement and also to opt for bigger returns on their investments by utilising the opportunities offered by the capital markets. The interests of individuals, the financial services industry, and governments, are all likely to be served both by a higher overall level of saving and by increased uptake of long-term investments in the capital markets (rather than risk-free fixed-interest products). Therefore, such an applied objective is in the consumer's interest and governments around the world are also encouraging such an initiative, because previous research suggests that people save too little and do not take enough financial risk (e.g., for analysis of this issue in UK see Oliver, Wyman and Company, 2001).

In this article, we present a study that investigates the effects of the framing and presentation of financial information when asking people to express their preferences in relation to different retirement savings and investment scenarios. The experimental design and method are based on the recent discovery of a substantial dependence of human preferences on the set of options they are presented with. This phenomenon was termed *prospect relativity* and indicates a lack of stable underlying preference function (Stewart, Chater, Stott, and Reimers, 2003). This finding, which we describe in more detail below, is a striking illustration of the view that preferences are constructed, on the fly, rather being a stable basis for decision making (Slovic, 1995).

A theoretically important question is how far these effects transfer from abstract low-stakes gambles, to major financial decisions that could significantly affect long-term well-being. If we observe such transfer to realistic financial decisions, then this would have crucial practical significance for marketing, sales, and provision of advice in the financial services industry. Our goal was to investigate how far individual variation in risk preferences is stable across different realistic decision contexts.

THEORETICAL BACKGROUND

Various forms of context dependence of human decision making have been studied extensively in the past. These effects are inconsistent with the normative theory of choice in economics, and therefore, they can be viewed as irrational (although see Sher and McKenzie, 2006, for discussion of a sense in which many context effects may have a rational basis).

Traditionally, psychologists and economists have tried to empirically assess the assumptions of the standard theory of choice in economics, revealing mounting evidence that human behaviour diverges from the predictions of the theory, especially in the context of decision making under risk (e.g., Kagel and Roth, 1995; Kahneman and Tversky, 2000; and also Camerer, 1995, for a review). Thus, a large literature in the psychology of decision making has demonstrated that expected utility theory does not fit with observed behaviour in a large variety of ways (e.g., Kahneman, Slovic, and Tversky, 1982). Some of these disparities between observed behaviour and the normative strictures may be due to lack of familiarity, practice or understanding of the problem in hand; however, behaviour that appears inconsistent with rational choice theory is very difficult to eliminate (e.g., Shafir and LeBoeuf, 2002). Thus, there has been considerable interest, in both psychology and economics, in building models that provide more accurate descriptions of observed decision making behaviour. These psychological theories have typically involved variations of the expected utility, to bring its predictions in line with people's observed behavior (e.g., prospect theory [Kahneman and Tversky, 1979], regret theory [Loomes and Sugden, 1982], rank-dependent utility theory [Quiggin, 1982]). The result involves building a bridge from expected utility theory (the normative theory) to psychology (the descriptive theory), where the core elements of the normative approach are maintained.

Rather than starting from a normative economic theory, and attempting to make modifications that render it descriptively acceptable, an alternative research strategy is to start from assumptions about elementary cognitive processes, and attempt to construct an account that can address the economic problem of choice. That is, our attempt is to bridge from psychology to economics (instead from economics to psychology as it used to be the case) with the ultimate aim of showing that the resulting model predicts decision-making behaviour (see, e.g., Gigerenzer and Todd, 1999 and Payne, Bettman and Johnson, 1992, for examples of this strategy).

In search of such realistic psychological foundations for descriptive decision theory, Stewart et al. (2003) considered whether context effects observed in psychophysics might transfer to risky decision making. Specifically, Stewart et al. (2003) found that the set of options from which an option was selected almost completely determined the choice. They demonstrated this effect in selection of a risky prospect and in a certainty equivalent estimation task (the amount of money for certain that is worth the same to the person as a single chance to play the prospect). Similarly, the selection of a preferred option from a set of prospects was strongly influenced by the prospects available. More recently, Stewart, Chater, and Brown (2006) developed a model of risky choice, *decision by sampling*, which assumes relative judgments and provides one explanation for the results in Stewart et al. (2003).

The cognitive claim that Stewart et al. (2003) argue for, is that people are unable to represent absolute magnitudes, whether psychophysical or abstract (including a choice option's attributes like utilities, payoffs, and probabilities). That is, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other magnitudes retrieved from memory or observed in the environment. If people cannot represent the absolute value of magnitudes on any cardinal scale, and the subjectively judged utility of an option is determined by its relationship to comparison options, then judgments will be strongly affected, or even determined, by the context.

The core evidence for this claim comes from the study of the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a

sound. Much traditional research in psychophysics has assumed the existence of some cardinal internal scale of intensities, onto which physical stimulation must somehow be mapped; and there has been consequent debate concerning the nature of this mapping (e.g., whether it is logarithmic, as argued by Fechner, 1966; or a power law, as argued by Stevens, 1957). But more recent theory (reviewed and analysed in Laming, 1997) suggests a different point of view---that the very idea of an internal scale is incoherent. In particular, Laming (1997) has shown that empirical data in line with Stevens' power law (relating psychophysical variables and free numerical judgments) can arise without assuming any representation of absolute information. In addition, Stewart, Brown, and Chater (2005) demonstrated that a wide range of data from the psychophysical task of absolute magnitude identification can be captured by a model which has no absolute scales, and relies entirely on local comparisons between recent stimuli.

A particularly telling example supporting this viewpoint is an elegant experiment conducted by Garner (1954), who asked participants to judge whether tones were more or less than half as loud as a 90 dB reference loudness. Participants' judgments were entirely determined by the range of tones played to them. There were three groups of participants who received tones in three different ranges respectively. Participants who heard tones in the range 55-65 dB had a half-loudness point (i.e., where their judgments were "more than half as loud" 50% of the time and "less than half as loud" 50% of the time), of about 60 dB. Another group, who received tones in the range 65-75 dB had a half-loudness point of about 70 dB. A final group, who heard tones in the range 75-85 dB, had a half-loudness point of about 80 dB. Garner's experiment indicates, therefore, that people have no idea of the absolute intensity of the sound or what it means for one sound to be half as intense as another. Instead, it seems that people adjust their responses depending on the presented sound intensities from which they are asked to choose.

Other examples of similar context effects abound in psychophysics. Thus, empirical investigations in absolute identification (e.g., Garner, 1953; Holland and Lockhead, 1968; Lockhead, 1984; Luce, Nosofsky, Green, and Smith, 1982; Ward and Lockhead, 1970), magnitude estimation (e.g., Jesteadt, Luce, and Green, 1977), relative intensity judgment (Lockhead and King, 1983), and matching tasks (Stevens, 1975), have shown that perceptual judgments of stimuli varying along a single psychological continuum are strongly influenced by the preceding material. In summary, context effects, like those found by Garner (1954), are consistent with participants making perceptual judgments on the basis of relative magnitude information, rather than absolute magnitude information (see also Laming, 1984, 1997).

Here we have digressed briefly into psychophysics, in order to make it explicit how we apply the resulting conclusion to a decision-making context. Note, though, that the parallel between the two domains, psychophysics and economic decision making, is relatively close. After all, just as perceptual theorists traditionally assumed that people have internal scales for the representation of loudness and brightness, so a traditional psychological or economic picture of an agent assumes that the agent must have internal scales for the representation of the utility of various outcomes; for representing, perhaps distorted (e.g., Kahneman and Tvesky, 1979) of the probability that they will occur; and so on. Without some type of scale for utility or probability, the model of the economic actor (or any decision-making agent in general) would look very different.

One prediction based on such model is that the attributes of the previously or currently seen risky prospects influence the decisions in the current prospect. Stewart et al. (2003)

argued for the existence of what they call "prospect relativity": that the perceived value of a risky prospect (e.g., "*p* chance of *x*") is relative to other prospects with which it is presented. Note that Stewart et al. studied peoples' perception of utilities in individual decision making tasks in gambling situations. The prediction, based on the psychophysical studies described above, is that the option set (i.e., the context) will affect peoples' choices because there is no fixed internal scale according to which people make their judgements of the values of certain options. Recall that the results demonstrated a powerful context effect in judging the value of different risky prospects - the set of options offered as potential certainty equivalents for simple prospects was shown to have a large effect on the certainty equivalents selected. To illustrate this result, for example, when during judging the value of a 50% chance of winning £200 people have options of £40, £50, £60, and £70, the most popular choice is £60 and then second choice is £50. When people have options of £90, £100, £110, £120 pounds, the most popular choice is £100, and then second choice is £110. So the set of alternatives affected valuation by a factor of (almost) 2. This effect was replicated despite monetary incentives designed to encourage participants to deliver accurate and truthful certainty equivalents. In another experiment, the set from which a simple prospect was selected was also shown to have a large effect on the prospect that was chosen. Vlaev and Chater (2006) discovered similar results in a very different context, where people play the strategic games based on Prisoner's Dilemma, indicating the generality of this effect. Finally, Vlaev, Chater, and Stewart (2007) report three studies, in which methodologies from psychophysics (which were similar to the methods used by Stewart et al., 2003) were adapted to investigate context effects on individual financial decision making under risk. The aim was to determine how the range and the rank of the options offered as saving amounts and levels of investment risk influence people's decisions about these variables. In the range manipulation, participants were presented with either a full range of choice options or a limited subset, while in the rank manipulation they were presented with a skewed set of feasible options. The results showed that choices are affected by the position of each option in the range and the rank of presented options, which suggests that judgments and choices are relative.

Effects of the type presented above suggest that people's expressed (or revealed) risk preferences are not absolute, but are, to some degree at least, relative to the range of available options (see Stewart, Chater, and Brown, 2006, for model of risky choice that assumes relative judgments only). A plausible account of the context effects caused by the range of options is the *range-frequency theory* proposed by Parducci (1965, 1995). Parducci found that the neutral point of the judgment scale did not correspond to the mean of the contextual events, contrary to popular at that time adaptation level theory (Helson, 1964), but rather to a compromise between the midpoint (defined by the range) and median (depending on the skew) of the distribution of contextual events. For example, satisfaction judgements are different between two distributions of life events, which have different skew and identical means. The *range principle* reflects tendency to judge an event relative to its position within the range of stimuli on the specified dimension of judgment, while the *frequency principle* reflects a tendency to judge an event relative to its rank within the immediate context. The subjective value given to an attribute is a function of its position within the overall range of attributes, and its rank. Thus, this model implies that attributes are judged purely in relation to one another and their subjective value is independent of their absolute value. Range-frequency theory has already been used to account for context effects in decision making

under risk. Birnbaum (1992), Stewart et al. (2003), Vlaev and Chater (2006), and Vlaev, Chater, and Stewart (2007) found their data to be consistent with the theory.

In this article, we present a study that aimed to test the practical relevance of the prospect relativity principle to more realistic financial decision making scenarios. The results we present here suggest that, when people make financial decisions, the attractiveness of the choice options significantly depends on the other available options.

PROSPECT RELATIVITY PRINCIPLE AND REALISTIC FINANCIAL DECISION SCENARIOS

The goal of this experimental study was to make a provisional estimate of the degree to which the kinds of effects that are revealed by Stewart et al. (2003) could also be applicable to real financial decisions. Our study builds on, and is similar to, the study by Vlaev, Chater, and Stewart (2007). The domain used here, and also by Vlaev, Chater, and Stewart, was saving and investment for retirement, because it is an issue having serious social relevance at the moment. We currently live in a financial environment, in which aging population and younger consumers are increasingly expected to take command of their own pension and investment decisions. Therefore, the following two key issues arise: How does the range of options people choose between affect the level of pension investment they choose? How does the range of options from which people choose affect the level of risk they accept with that investment?

This experiment followed logic similar to the decision experiments reported by Stewart et al. (2003) and Vlaev, Chater, and Stewart (2007), which were in turn inspired by Garner's (1954) loudness judgment experiment. In various questions, the participants in our study were asked to select among a predefined set of values related to five variables: (a) the desired percentage of the annual income that will be saved for retirement, (b) the investment risk expressed as the percentage of the saving that will be invested in risky assets, (c) retirement age, (d) expected retirement income, and (e) possible variability of the retirement income.

There was a control condition called a *free choice* condition, in which the participants had to freely decide the value of each one of these variables selecting from the full range of options. In two *context* conditions, participants were asked to select these values from sub-ranges of the set of options offered by the experimenter in the free choice condition. Thus, there were three between-participant conditions in the experiment presented here, i.e., with separate groups for the free choice, low range, and high range conditions. In the free choice condition, all options were presented. In the two other conditions, the choice of prospects was limited to either the first or second half of the prospects available in the free choice condition, so that the participant in the high range condition were presented with a range of values the lowest of which coincides with the highest option in the low context condition. In the free choice condition for saving, the options were presented in monetary terms and varied from 2% to 22% of the hypothetical salary (£25,000) increasing with 2% between the options; so there were eleven options to choose among, while the low range condition spanned from 2% to 12% and the high range condition was from 12% to 22%. The same design was applied for the other risk variable. The choice option values in the free choice condition for investment risk varied from 0% to 100% and were increasing with 10% between the options. For

retirement age the values varied from 48 to 68 increasing with 2 years. Table 1 presents the values for savings, risk, and retirement age, in the three conditions. For the retirement income and its variability, the values were different for every question depending on the combination of saved amount, investment risk, and retirement age.

Table 1. Figures for saved amount (£), investment risk (%), and retirement age in the three conditions of the experiment

Free Choice			Low Range			High Range		
Save	Risk	Retire	Save	Risk	Retire	Save	Risk	Retire
500	0	48	500	0	48			
1,000	10	50	1,000	10	50			
1,500	20	52	1,500	20	52			
2,000	30	54	2,000	30	54			
2,500	40	56	2,500	40	56			
3,000	50	58	3,000	50	58	3,000	50	58
3,500	60	60				3,500	60	60
4,000	70	62				4,000	70	62
4,500	80	64				4,500	80	64
5,000	90	66				5,000	90	66
5,500	100	68				5,500	100	68

Note that for risk, it is natural to assume that people anchor the possible percentage of their retirement savings that can be invested in risky assets to be between on 0% and 100% (i.e., people cannot have negative savings, or invest more than 100% of their savings). For the range of possible savings values, it is equally easy to imagine that the lower bound is fixed at 0%, while the upper bound depends on many factors like for example legal requirements, cost of living, etc. We fixed the savings range across all conditions at values 0% and 22% (the range for the full range condition) because 22% approximates the upper bound for the retirement savings rate in UK due to legal and tax restrictions. The team of professional actuaries who monitored our research project suggested 22% as the practically relevant upper bound.

Our prediction was that if a participant is not influenced by the set of options, then his or her choice of each value in the high and low range conditions should be independent of the other values in the set and the chosen values should be the nearest to his or her free choice. The first prediction was that if people are influenced by the context (i.e., the other available options), then the mean saving and risk selected in the both high range and low range conditions should be different from the free choice condition. This is because we assume that if people's true preferences are represented by the results in the free choice condition, and their choices are not influenced by the context in the high or low range condition, then at least one of the sub-range conditions should not be different from the full range condition. Thus, for example, if peoples' preferences are naturally amongst the lower options in the full range, then there should not be any significant difference between the results in the low range and the full range; the only exception should be the highest value in the low range, which ought, if there are no context effects, to include all the people who would choose that option, or all higher options, in the full range condition. Conversely, if all participants that truly prefer

options in between the higher options in the full range, then there should be no significant difference between the average results in the high range and the full range; the only exception should be the lowest value of the high range, which should, in the absence of context effects, include all those people who would choose that options or a lower one, in the full range. However, if *both* the low range and the high range conditions are significantly different from the full range condition, in terms of the low range being lower than the full range *and* the high range being higher than the full range, then we can conclude that the context had significantly influenced the choices in the sub-range conditions (i.e., this result should be due to the effects of the choice set in the high range and low range conditions). In other words, the choice set is affecting the responses and inducing people to select higher/lower options than they would have selected if they were in the full range condition.

We also conducted the following statistical comparison between the high range and low range conditions in order to test furthermore whether the context (range of options) had a significant effect on the choices. We compared the lowest option in the high range, against the sum of the all the items in the low range, except the highest item (i.e., every option in the low range, which was missing in the high range). In other words, we compared whether the lowest option in the high range condition was significantly lower than the proportion of times the options below it were selected in the low range condition. Conversely, we compared the highest option in the low range, against the sum of the options in the high range, except the lowest option. In either direction (and we tested both directions), if the latter is bigger, we have prospect relativity effect (or at least, a rational choice model will fail to predict this result) and we can conclude that this result should be due to the effects of the choice set in the high range and low range condition respectively. The logic behind this analysis is that people who do not select the highest item in the low range condition should definitely select the lowest item in the high range condition. So the former ought to be less numerous than the latter. This is because their true preferences should be within the options lower than the highest item in the low range condition. In other words, the reason we are missing the highest item in the low/high range, is that we do not know what proportion of the people selecting this option want to select options above/below it, but they do not have the opportunity in the low/high range. Alternatively, if participants' responses are solely determined by the set of options presented to them, then the distribution of responses across options should be identical for both the low and the high range conditions.

In addition, the design of the experiment presented in this article had four new features relative to the work reported by Vlaev, Chater, and Stewart (2007). We designed these new characteristics in order to increase our study’s relevance to real-world financial advice. These four new design features were:

1) Representative Sample

The study was conducted on a sample of working people, rather than university students. In other words, we used sample of participants who are more realistic to be consumers of financial advice (e.g., people who are already working, have a family, and need to save for retirement pension provision) than the student population used in Vlaev, Chater, and Stewart (2007).

2) Realistic Financial Assumptions

The future financial outcomes (e.g., expected annuity values) were calculated using very plausible financial assumptions (like inflation, risk free rate, risk premium rate, etc.). We undertook this work on consumer understanding of risk, both from a mathematical and from a psychological standpoint, with the help of the Actuarial Profession's Personal Financial Planning Committee in United Kingdom (who actively participated in creating the test materials and the descriptions of the risky assets). The age of the participants was also taken into account in calculating the time horizon of future returns.

3) Financial Affordability Questionnaire

In the light of the important discussion raised at our meetings with professional actuaries and personal financial advisers, we took account of financial affordability, as a constraint on people's choice of pension. We created a financial affordability test, which categorised people according to their individual financial circumstances. The financial affordability test was designed to help and encourage the participants to think through the practical viability of the financial options that they choose. An additional purpose of the financial affordability test was to make the experimental situation appear as a very realistic example of a financial advisory process. This was achieved by asking the participants concrete questions about their real life financial circumstances and problems. Thus, by explicitly focusing respondents' attention on their real life struggles at the beginning of the experimental session, we expected them to provide more adequate and valid responses to our saving and investment questions.

The financial affordability questionnaire is presented in Appendix A and it has the following main features:

a) Question 2 asks people to estimate to what extent their current income (question 1) is sufficient to cover their expenditures, and then to judge in percentage term whether, and by how much, this income is sufficient or insufficient to cover these expenses (e.g., "I would be happy to earn around 20% more than my current salary"). In addition, the participants were asked to indicate how much they are able to save at the moment (question 3).
b) In question 4, the participants were provided with a list of various types of spending and they had to answer how much of their current income is spend on each of these expenditures. In general, there were two types of expenditure examples – discretionary (e.g., leisure activities) and essential ones (e.g., food and rent) and we asked the respondents to give estimates of their expenditure across these categories.
c) Question 6 asks people whether they can give up some of their discretionary spending in order to increase their savings. Here the focus again was on the amount and type of current saving and discretionary spending; and hence the degree to which people can readily reallocate money towards a pension. Here we also aimed to test how important and essential some of these discretionary expenditures are (e.g., some people might be unwilling to give up some types of social life, hobbies, sport activities, etc.). Participants were also informed that at the and of the experimental session if their average preferred savings rate is above their current savings as

indicated in question 3, then they had to readjust some of their expenditures in question 4 so that to be able to provide the additional capital that is the lacking difference between their real savings and the saving levels selected in the second part of the experiment (in addition, question 5 separately asked what is the maximum amount that they would consider saving each year).

4) Risk Preference Tests

We collected information about participants' risk attitudes in order to investigate to what extent their decisions were influenced by their general risk preferences. Thus, we expected to test whether people can be manipulated to take more risk than their genuine risk preferences are, or whether they intuitively know how much investment risk to take. Such results could also inform us whether it is worth trying to stimulate people to invest in a way that best matches their risk and time preferences (for example, to increase their investment risk exposure if they are particularly risk seeking), or, if people's decisions are easily manipulated by the context, whether to offer them financial products only on the bases of their individual goals and social and financial circumstances (e.g., to offer them relatively risky investments that will accomplish the desired retirement income, independently of their risk preferences). We used five different measures of risk aversion (presented in Appendix C) in order to measure whether the choices of investment risk in the high and low range conditions are due to natural risk preferences instead of context effects. These measures represented typical self-report hypothetical measures (as used in the literature) in the form of simple direct questions and hypothetical gambles

Questions 1-4 (Direct Risk, Direct Concern, Relative Risk, Relative Concern). These questions are rather simple and direct measures. We used these both as a base-line, and also because of existing results showing that simple self-report measures of risk preferences could be more powerful predictors of portfolio allocation than sophisticated measures based on economic theory (Kapteyn and Teppa, 2002). Two of these questions measured risk attitudes with the basic questions "How much risk are you prepared to take?" (Direct Risk) or "How much are you concerned about your financial future?" (Direct Concern) and the participants had to answer on a scale from 1 (not at all) to 5 (very much) to what extend they agree with these statements. There were also two questions about how people perceive their level of risk aversion in relation to other people – "Are you more or less willing to take risks than the average person?" (Relative Risk) or "Are you more or less concerned about your financial future than the average person?" (Relative Concern) and the participants had to answer on the following scale: 1 - much less, 2 – less, 3 - the same as the average, 4 – more, and 5 - much more.

Question 5 (Income Gamble). Question 5 is a well-known test by Barsky, Juster, Kimball, and Shapiro (1997), who constructed a measure of risk aversion by asking respondents about their willingness to gamble on lifetime income. By contrast, experiments in the existing literature ask people to gamble over spending or consumption and typically involve stakes that have little impact on lifetime resources. However, a gamble whose outcome is too small to be meaningfully related to consumption should not require a risk premium, on normative grounds, and therefore such gamble is not a good measure of economic risk preference. So the principal requirement for a question aimed at measuring risk aversion according to Barsky

et al. is that it must involve gambles over lifetime income. In addition, after pre-testing, Barsky et al. concluded that survey respondents would better understand income than consumption lotteries. The three questions in this test, in the first paragraph and then in (a) and (b), separate the respondents into four distinct risk preference categories, depending on the combinations of their answers (see Question 5 in Appendix C): (1) reject the risk to cut the (family) income by one-third in the first question and also reject the risk in (b) to cut the income by one-fifth (20%); (2) reject the risk for one-third income cut in the first question but accept the possibility for one-fifth cut in (b); (3) accept the possibility for one-third income cut in the first question but reject the one-half cut risk in (a); and (4) accept both possibilities for one-third income cut in the first question and one-half cut in (a). These four categories can be ranked by the level of risk-seeking without having to assume a particular functional form for the utility function. Barsky et al. (1997) provide four numerical indices of relative increasing risk-seeking corresponding to each category respectively: 0.11, 0.36, 0.68, and 1.61. In the original study by Barsky et al., their measure was significantly correlated with various demographic factors, and it was positively related to risky behaviors, including smoking, drinking, failing to have insurance, and holding stocks rather than treasury bills.

METHOD

Participants. We sent the materials (the financial affordability questionnaire and the savings and investment questionnaire) by post to a population of working individuals, as this population is typically not able to attend laboratory sessions. We sent out the questionnaire to 600 people, and received 64 completed questionnaires. These respondents were typical of the demographics in the geographical area and selected from a big subject pool of people who expressed desire to participate. Participants who completed their questionnaires were paid £10 for their participation (received as a check after they have returned the answer sheet). There were 24 men with average age 36.5 and 40 women with average age 37. The Low Range Condition had 20 participants: 7 men (av. age 37) and 13 women (av. age 38); the Free Choice Condition had 21 participants: 9 men (av. age 33) and 12 women (av. age 36); and the High Range Condition had 23 participants: 8 men (av. age 40) and 15 women (av. age 36).

Design. The questions in the prospect relativity test were formulated as long-term saving/investment decision tasks related to retirement income provision. The participants had to make decisions about five key variables. These variables were the saved proportion of the current income, the risk of the investment expressed as the proportion invested in risky assets,[1] the retirement age, the desired income after retirement, and the preferred variability of this income (we explained that such variability is due to favourable and respectively unfavourable economics conditions).

The experimental materials were designed as 10 independent hypothetical questions, in which we varied each of the five key variables. However five of the questions focused only

[1] There are, of course, various types of risky assets, including a wide variety of bonds and equities; but in reality these various investment vehicles differ mainly in their risk-return characteristics. Therefore, we simply described the characteristics of these two assets – the High Risk Asset and the Low Risk Asset, rather than labelling them explicitly as bonds and equities, although while setting the basis so that it is not out of line with typical assumptions made about actual assets. This aimed to avoid some of the potential challenges that might otherwise result and which could draw attention away from the results.

on savings while the other five questions focused on risk, and some questions showed how changing savings or risk would affect another variable or set of variables. For example, how changing the investment risk can affect the projected retirement income and its variability - with higher risk offering higher expected income on average, but also wider spread of the possible values.[2]

As an example, Figure 1 presents a question in the free choice condition, in which the participants were asked to choose their preferred level of investment risk by selecting one of the rows in the table (note that in this format the key choice variable is in the first column of the table while the other columns are showing the effects on the other variables like the minimum, average, and maximum retirement income[3]). This future distribution (risk) of the investment in risky assets is calculated as follows. Assuming a variable annual interest rate with mean μ and standard deviation σ, the expected return on an *n*-year investment is also log normally distributed with mean μ^n and standard deviation $\sigma = \mu^{2n}(((\sigma^2/\mu^2)+1)^{n-1})$. We also assumed that an annuity that provides 1/14th of the lump sum saved each year is purchased, which is a typical figure used in the UK financial services industry.

In Appendix B, there is a detailed description of each question and its purpose (the questions are grouped by the key variable that participants are asked to select – savings or investment risk). The ten questions were presented in different order in the various conditions. We also counterbalanced the order of saving and risk questions by dividing the participants into two groups: one that first answered the saving questions and then the risk questions, and second group that answered the risk questions before the saving ones.

Note that since we used mature population of participants who vary in age, this might create problems with using the same test materials for all participants, because older people would have so save for fewer years (compared with younger people) in order to get the same retirement income. Thus for example if the materials (saving and risk choice options) are created for people with average age of around 25 yrs, and if we give the same materials to somebody who is 50, then of course the older respondent would be willing to save the highest possible amounts for the remaining 10-15 years until retirement, while a younger person has to commit to this higher saving (and lower consumption) rate for 25-30 years. In order to avoid this problem, we decided to create test materials for three different age groups, namely, 30, 40, and 50. We sent the identical test materials to respondents who are plus or minus 5

[2] In order to derive plausible figures for the various economic variables we implemented a simple econometric model into a spreadsheets Monte Carlo simulator that calculates the likely impact of changes in each variable on the other four variables. For example, this model can derive what retirement income can be expected from certain savings, investment risk, and retirement age, or what are the possible potential investment options that could lead to the preferred retirement income. The sort of basis the professional actuaries suggested was 2.5% for Inflation, 1.5% real return on Low Risk asset, 4.5% real return on High Risk asset, and 15% annual volatility. Note also that all figures are in pounds and the participants knew this. It is important to stress that all figures shown were in today's money terms (i.e. after taking out the effects of inflation). This is important when comparing figures for different retirement ages.

[3] Most of the questions showed the expected retirement income and its variability like in the example above. The possible variability of the retirement income was explained by referring to the 95% and respectively 5% confidence intervals of the income variability, i.e. maximum and minimum possible values of the income, for which there is 5% chance to be more than the higher or less than the lower value respectively. On each row of the table these two values were placed on the both sides of the average expected retirement income. The confidence intervals were expressed also in verbal terms using the words very likely. For example, the participants were informed that it is very likely (95% chance) that their income will be below the higher value and above the lower value, and that these two values change depending on the proportion of the investment in equities.

years around each age group (e.g., the financial options calculated for somebody who is 30, were also sent to all respondents between 25 and 35 yrs old). Thus the projected retirement income was calculated for three time horizons: after 35, 25, and 15 years of investment respectively.

Assume that you will retire at 65 and decided to save 11 percent of your current salary (£2750) in order to provide for your retirement income. The following options offer different ranges of retirement income (in pounds) depending on the percentage of your savings allocated to shares (in the stock market) and you can see the effects on the expected average retirement income and its variability (minimum and maximum). Note that you are very likely (have 95% chance) to be between the minimum and maximum figures indicated in the table below. Please select one of the following options.

Invest	Minimum	Average	Maximum
0 %	16,000	16,000	16,000
10 %	17,000	19,000	22,000
20 %	17,000	21,000	23,000
30 %	17,000	23,000	29,000
40 %	16,000	26,000	35,000
50 %	15,000	29,000	42,000
60 %	14,000	33,000	51,000
70 %	11,000	37,000	62,000
80 %	7,000	41,000	76,000
90 %	2,000	47,000	92,000
100 %	0	53,000	112,000

Figure 1. A question in the free choice condition, in which the participants were asked to choose their preferred level of investment risk by selecting one of the rows in the table bellow. In this format the key variable is in the first column of the table bellow while the other columns are showing the effects on the other variables like the minimum, average, and maximum retirement income.

The high range condition was derived by deleting the lower five rows of the table for each question in the control condition and the low range condition was derived by deleting the higher five rows in the tables in the free choice condition (i.e., the same was done for each question). Therefore, in the free choice condition, the participants had to choose among eleven possible answer options for each questions while in the high and low range conditions there were only six available answer options. Note that in this design the participant had to choose among predetermined option values in all conditions. This design was similar to the design used in Experiment 4 reported by Stewart et al. (2003), where in the free choice condition the participants had to choose among predefined set of risky prospects (gambles), while in the two context conditions they were asked to choose among predetermined choice options that were either the higher halve or the lower half of the list of options offered in the free choice condition. Vlaev, Chater, and Stewart (2007) also used very similar design.

Procedure. Participants were sent a booklet containing the financial affordability questionnaire, the ten saving and risk questions, and the five questions measuring risk aversion. They received written instruction explaining that the purpose of the experiment is to answer series of questions about savings and investment related to retirement income provision, and that there were no right and wrong answers and that they are free to choose

whatever most suits their preferences. It was explained that the choice options are predetermined because these are the outcomes that can be realistically accomplished according to a standard economic model and that the task is to choose the option that is nearest to the participant's preferences. The participants were also informed that if they find them unsatisfactory then they can indicate values outside these ranges.

The questions and the answer options were presented as in the example question in Figure 1, which showed the projected retirement income for 30 years old age group (i.e., after 35 years of investment). The participants chose one of the figures in the first column of the table (which were either savings or investment risk values) and they were provided with a separate answer sheet on which to write their answers. Participants were informed that their answers do not need to be consistent between the questions, and that they can freely change their preferences on each question and choose different savings and risk values.

Another issue that we had to deal with was how to account for people's existing savings because we wanted to make our session as realistic as possible. If we give our questions to somebody who already has got a good pension scheme, then she might choose very low saving amounts and investment risk just because she does not need to save much more. On other side, if we tell them to imagine that our scheme is offering them to start anew, then our calculation will have to include also their accumulated savings up to date. This would also require some sophisticated software to be used online with every individual (and which is probably used by the real financial advisors). Our solution to this problem was to write in the instruction that most people in UK are underprovided and that we research what kind of pension top-up product people might find attractive (in addition to the social security scheme), and therefore this is an extra to what they already have.

RESULTS

Financial Affordability Questionnaire. Table 2 presents the results from the Financial Affordability Questionnaire. We checked whether the participants were in a position to afford the saving levels selected in the experiment. None of the participants had selected inappropriate saving rates in relation to their income and expenditures. Only one person decided to give up half of her essential spending and also discretionary spending in order to provide the additional capital that is required to cover the difference between her real savings rate and the maximum savings rate that she had indicated in the main test. Otherwise, 54% of the participants indicated that they can give up some discretionary spending in order to increase their current savings rate if it is bellow their preferred maximum amount (indicated in question 5). This result indicates that all respondents took their task seriously and carefully selected the saving options in the main test so that their choices reflected their real financial circumstances. We also hoped that asking people to provide this financial information would encourage them to consider their saving decisions carefully and as a result to give answers that are reasonably close to what they would choose for real (recall that the purpose of this questionnaire was mainly to prompt the respondents to give realistic answers to the prospect relativity test).

Table 2. Results from the Financial Affordability Questionnaire

Question		Mean	Std Deviation	Category
Annual income		£19,235.5	15,492.6	General
Spend less than you earn by…		£3,601.1	3,032.7	
Spend exactly the amount you earn		£18,658.8	8,729.8	
Spend more than you earn by		£1,593.8	1,136.4	
Current Saving		£1,829.3	2,363.2	
Food		£2,147.3	1,905.0	Essential expenditure
Rent / Mortgage		£3,177.0	1,990.6	
Utilities (electricity, gas, water, etc.)		£659.9	832.4	
Car		£1,301.5	1,616.8	
Other transport (train, busses)		£343.6	686.3	
Debt repayment		£1,021.2	1,234.8	
Communications (telephone, etc.)		£424.1	300.5	
Childcare and Schooling		£283.5	747.0	
Health		£76.3	109.7	
Repairs and Maintenance		£439.6	542.6	
Other (like health and life insurance, etc.)		£425.0	479.6	
Holiday		£327.0	392.3	Discretionary expenditure
Entertainment (e.g., cinema)		£234.0	356.5	
Sport		£177.5	222.1	
Hobbies		£554.7	516.4	
Meals and Drinks		£663.7	1,351.6	
Other		£1,1249.0	5,816.9	
Total Expenditure		£3,177.0	1,990.6	Demographics
Desired Saving		£3,606.9	3,589.4	
Household Income		£30,328.4	£34,202.2	
Give up discretionary spending to save	Yes	54.2%		
	No	45.8%		
Employment	Part-time	31.7%		
	Full-time	68.3%		
Education	School	5.08%		
	College	28.8%		
	University	66.1%		
Time spend managing finances	Not at all	13.3%		
	Occasionally	28.3%		
	Regularly	35.0%		
	Often	15.0%		
	Very often	8.3%		

Prospect Relativity Test. Note that although the questions related to saving and risk asked the participants to trade-off different variables (e.g., savings versus retirement income in one question, and savings versus risk in another question), we used the weighted average of the answers of each participant across all five questions related to saving and all five questions related to risk, in order to derive the mean values for saving and risk in each condition; and

these averaged results are presented here. This was done because the results showed qualitatively the same pattern and there were no significant differences across the five questions for saving and risk respectively.

There are two sets of responses to consider. First, we present the results on the savings rate. The proportion of times each saving option was chosen in the free choice, low range, and high range conditions is plotted in Figure 2. The presented results were averaged over all participants (which was also done for all statistical tests presented here). The error bars represent the standard error of the mean, which is also presented in all other figures in this article. The mean savings in the high range condition (£3,540) was significantly higher than the full range condition (£2,285), $t(38) = 4.94$, $p < .0001$, while the mean savings in the low range condition (£1,660) was significantly lower than the full range condition (£2,343), $t(38) = 2.05$, $p = .0469$. Thus, since both the low range and the high range conditions are significantly different from the full range condition in the predicted directions, we can conclude that the range of offered saving options has strongly affected the mean of the selected values in each group.

Now we consider the direct test of whether these data are compatible with stable absolute preferences, by comparing the restricted (high and low range) conditions (using the logic outlined earlier). The proportion of times the lowest saving option in the high range condition (the £3000 option) was selected was .42 and was significantly lower than .77 which is the proportion of times the options below it were selected in the low range condition, $t(38) = 3.34$, $p = .0019$. This result indicates that the context has affected choices in the high range condition.

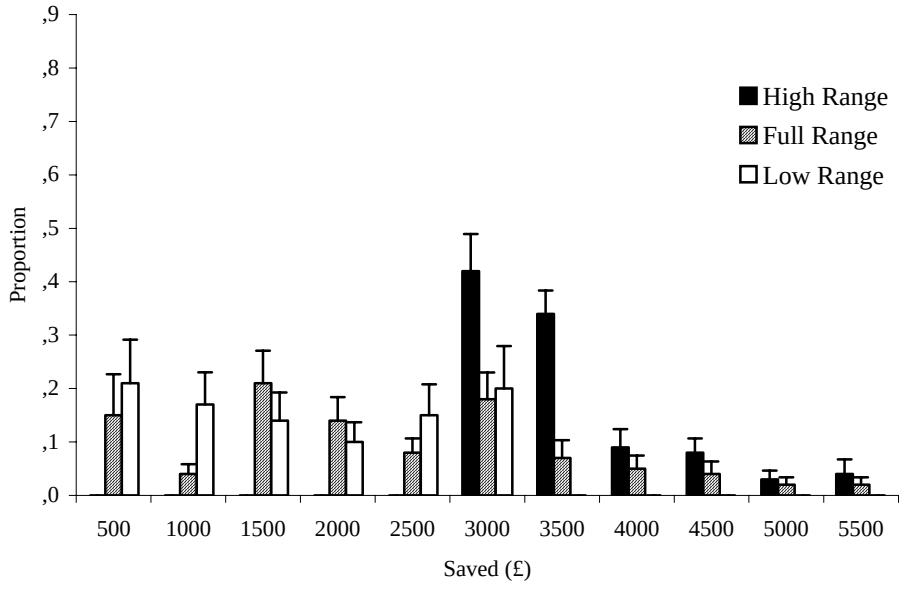


Figure 2. Proportion of times each saving option was chosen in the low range, full range and high range conditions. (Error bars are standard error of the mean).

The proportion of times the highest option in the low range condition (again the £3000 option) was selected was .20 and this value was significantly lower than 0.58, which was the sum of the options in the high range condition, except the lowest option, $t(38) = 3.60$, $p = .0009$. This result also means that the hypothesis that participants' choices were unaffected by context should be rejected; the pattern of prospect relativity (Stewart et al., 2003) is evident.

Now we turn to the second set of data, concerning the level of risk people were willing to take with their saving. The proportion of times each investment risk option was chosen in the free choice, low range, and high range conditions is plotted in Figure 3. The average investment risk was significantly higher in the high range condition (56.1%) compared to the free choice condition (31.8%), $t(38) = 8.52$, $p < .0001$, and also the average investment risk in the low range condition (22.9%) was significantly lower that in the free choice condition (31.8%), $t(38) = 2.35$, $p = .0239$. This result also indicates significant context effects on the mean risk preferred in each condition because both the low range and the high range conditions are significantly different from the full range condition in the expected directions (higher in the high range condition and lower in the low range condition).

Turning to the comparison of the high and low range conditions, the proportion of times the lowest option in the high range condition (the 50% option) was selected was .58 and this value was significantly lower than the proportion of times the options below it were selected in the low range condition, which was .85, $t(38) = 3.11$, $p = .0035$. The proportion of times the highest option in the low range condition (again the 50% option) was selected was .11 and this result was significantly lower than .41, which was the sum of the options in the high range condition, except the lowest option, $t(38) = 3.60$, $p = .0009$. Again, the results are incompatible with the assumption that people have stable absolute preferences among the choices options, which is a further illustration of prospect relativity.

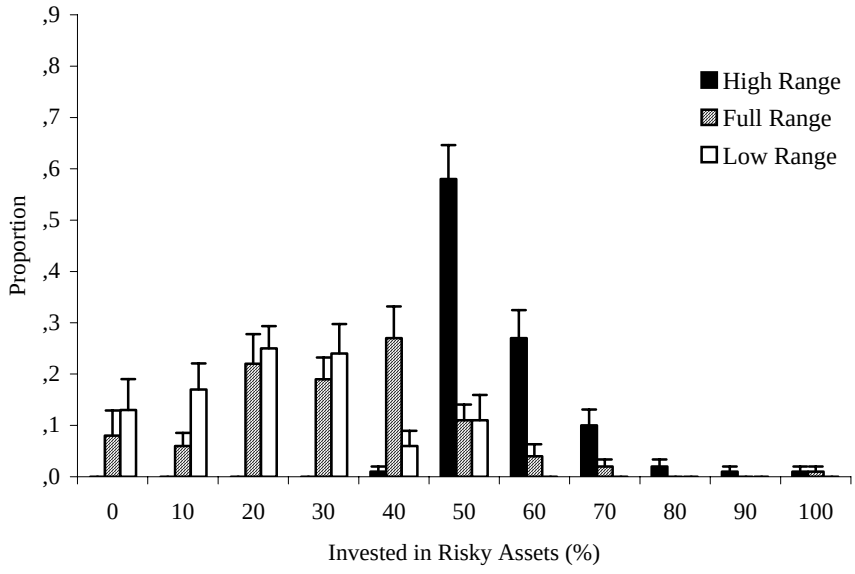


Figure 3. Proportion of times each investment risk option was chosen in the low range, full range, and high range conditions. (Error bars are standard error of the mean).

In summary, the results for saving and risk clearly demonstrate that the choices were strongly influenced by the set of offered choice options.[4] Thus, we replicated our previous findings (Vlaev, Chater, and Stewart, 2007) with realistic financial assumptions and population that does need to make these decisions in real-life. Note, however, that even though the choices were significantly affected by the context in the high range condition, the skewed results in this condition clearly show that there is a tendency towards certain most preferred values for savings and risk. This result suggests that people's preferences are not completely malleable by the context.

Thus, we have demonstrated that prospect relativity arises when people are faced with familiar situations with which they are likely to have some exposure and practice like saving, consumption, pension plans, and investment in the capital markets (at least the media provides enough information on the last issue). It seems also that people might, nonetheless, have developed some more stable preferences or anchors for savings and investment risk, although their responses are still very malleable to context effects. This could be viewed as evidence for some stable absolute scale for assessing money or risk; or, alternatively, as arising because people sample salient comparison 'anchors' from memory (e.g., income, expenditure, current savings---the type of information elicited in the first part of the experiment) against which the choices options are compared (Stewart, Chater, and Brown, 2006).

Risk Preferences. Table 3 presents the results from the five questions measuring respondents' risk preferences and the mean levels of investment risk in each condition. All five risk-aversion measures indicate that the respondents typically perceived themselves to be moderately risk averse. Note that risk-averse preferences are implied by values that are: a) lower than 3.0 for Direct Risk; b) higher than 3.0 for Direct Concern; c) lower than 3.0 for Relative Risk; d) higher than 3.0 for Relative Concern; and e) lower than 0.69 for the Income Gamble, which is the mean between the four indices of relative risk-seeking.

The significance value of the F test in the Analysis of Variance (ANOVA) test shown in Table 3 indicates that the average scores were significantly different between the three conditions (there was a main effect of the factor Condition) only for the investment risk chosen in the main test, $F(2, 61) = 65.98$, $p < .0001$. The investment risk levels, in line with our predictions, were lowest in the low range condition and highest in the high range condition. The average scores for the five risk preference tests were very similar (and for most measures the score in the low range condition was even higher than the score in the high range condition). In summary, these results demonstrate that the self-reported subjective risk preferences did not change as the context changed, while the investment risk choices did change accordingly.

However, a rather interesting result in presented in Table 4, which shows the correlations between the risk preference measures and the investment risk in each condition. We used the Spearman's correlation coefficient because some of the measures were quantitative variables (the Investment Risk and the Income Gamble) and some were variables with ordered categories (the questions 1-4). Table 4 shows that in the Low Range condition, there was a

[4] Additional analysis also established that the context effects are relatively similar for people from different income ranges. In other words, people who can least afford it are not more or less likely to be influenced by the range of options offered to them. We also did not find any gender differences in terms of saving and investment risk preferences, and context malleability on these two dimensions (i.e., women were not more context sensitive than men). All these analyses are not reported here, but these additional results are available on request.

strong correlation between the Direct Risk measure and Investment Risk, r = .64, p = .0022. In the Full Range condition, the Investment Risk correlated significantly again with Direct Risk, r = .68, p = .0015, but also with Relative Risk, r = .48, p = .0345. In the High Range condition, the Investment Risk correlated significantly only with Direct Risk, r = .50, p = .0264.

Table 3. Means risk levels chosen for each risk preference measure and for Investment Risk in each condition. Investment Risk is calculated as the mean proportion (%) of the savings invested in the Risky Asset. p is the significance value of the F test in the ANOVA testing the hypothesis that average scores are equal across conditions

Risk Measure	Condition			ANOVA
	Low Range	Full Range	High Range	*p*
Investment Risk	22.9%	31.8%	56.1%	.0000
	(12.1)	(11.7)	(5.00)	
Direct Risk	2.40	2.47	2.20	.5864
	(1.14)	(0.70)	(0.62)	
Direct Concern	3.30	3.47	3.15	.6863
	(1.38)	(1.12)	(0.93)	
Relative Risk	2.65	2.75	2.47	.6126
	(0.93)	(0.91)	(0.77)	
Relative Concern	2.85	3.05	2.84	.7270
	(1.04)	(0.83)	(0.90)	
Income Gamble	0.70	0.67	0.53	.5921
	(0.62)	(0.53)	(0.50)	

Standard deviations within parentheses.

In summary, only Direct Risk was significantly associated with risky choice in all three conditions. In other words, within each context, people who selected the options with higher risk also indicated that they were more risk seeking, and vice versa. This result suggests that while people's choices are dependent on the context, their subjective risk-aversion is a stable trait. The predictive power of the simple question measuring Direct Risk could be explained if we assume that people are more or less aware about their risk preferences, but their risk perception is determined by the context. The fact that only the simplest Direct Risk measure was a significant predictor, suggests that people use some very crude heuristics (e.g., "How much risk I am prepared to take?") to select choice options, which are perceived as relatively safe or risky only in comparison to the other available options (i.e., in the current context). Thus, the significant predictive power of the Direct Risk measure implies that people define their preferences in relation to the available set of choice options, which again corroborates our claim that judgments are made relative to the available reference points in the current environment.

Table 4. Spearman's rho correlations between the investment risk and the five risk aversion measures in the three conditions of the experiment. Investment Risk is calculated as the mean proportion (%) of the savings invested in the Risky Asset

Condition	Risk Measures	Investment Risk	Direct Risk	Direct Concern	Relative Risk	Relative Concern	Income Gamble
Low Range (N = 20)	Investment Risk	–					
	Direct Risk	.64**	–				
	Direct Concern	.38	.27	–			
	Relative Risk	.17	.29	-.10	–		
	Relative Concern	-.02	-.02	.39	-.37	–	
	Income Gamble	.34	.16	.08	.64**	-.39	–
Full Range (N = 20)	Investment Risk	–					
	Direct Risk	.68**	–				
	Direct Concern	.29	.33	–			
	Relative Risk	.48*	.75**	.08	–		
	Relative Concern	.40	.32	.58**	.38	–	
	Income Gamble	.10	.39	-.13	.43	-.10	–
High Range (N = 20)	Investment Risk	–					
	Direct Risk	.50*	–				
	Direct Concern	.07	.08	–			
	Relative Risk	.37	.60**	.19	–		
	Relative Concern	.25	.17	.75**	-.12	–	
	Income Gamble	.00	-.09	-.41	-.44	-.11	–

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

One implications of this result is that directly asking people about their risk preferences may be as useful as apparently more sophisticated risk-diagnostics in helping people to choose financial products. Moreover, our results concerning prospect relativity, and the literature on framing effects more generally, suggest that attempting to associate individuals with an 'economic' risk preference (e.g., the curvature of the utility or value functions) is likely to be ineffective, because results will depend substantially on the framing of the question, rather than reflecting an underlying attribute of the consumer. Nonetheless, simple direct risk measures may still be used to help in the design of the financial products offered by financial advisers, by making the range of offered investment options to vary depending on the risk profile of the consumer. For example, in order to prevent a risk-averse client to make an investment which is too risky for him/her, one could offer a relatively safe range of investment options, and thus utilise the powerful effect of the context in order to accomplish a better match with individual risk preferences.

GENERAL DISCUSSION AND CONCLUSIONS

Our results demonstrate that when people make financial decisions the attractiveness of the choice options significantly depends on the other available options. In particular, the set of options offered as potential savings and risk options was shown to have a large effect on the selected options. In general, it seems that the context provided by items that are considered simultaneously does affect decisions about saving and investment risk. These results could be considered as an example of the prospect relativity principle, which suggests that risky prospects are judged relative to accompanying prospects.

We believe that our new result reflects what would be likely to occur were these choices being made for real, e.g., in a session with a sales person or a financial advisor. The Financial Affordability questionnaire was designed to enforce the participants to make their decisions in light of their real financial circumstances. We believe that this type of test of context effects could not be done in more realistic conditions, because no regulating or legislative authority would allow people's choices of (real) investment products and saving rates to be manipulated in such a drastic way. Retirement investment and saving choices can fundamentally affect a person's quality of life (during her entire life span) and so ethical considerations preclude direct experimentation in the context of a real selling process.

We also tested whether people naturally tend to make decisions matching their individual risk preferences and whether they can be manipulated to pass this level by manipulating the context (which in reality might be necessary in order to accomplish some financial objectives). The results showed that very simple and direct risk preference measures are significant predictors of the responses within a particular context, although these responses differ between different contexts (i.e., the actual choices were very much influenced by the context). One conclusion is that the context has much more powerful effect on risky choices than the underlying risk preferences, but another possible conclusion is that risk preferences are consistently defined in relation to the particular context.

Theoretical Accounts

Range-frequency theory (Parducci, 1965, 1995) is consistent with the result in our experiment, which showed that preferences for saving and risk are very much determined by the range of offered choice options (in particular, preferences for the £3,000 saving option and 50% risk option were different in the high and low context conditions). The success of range-frequency theory (Parducci, 1965, 1995) in accounting for our financial prospective relativity results (and also for choice in gambling, financial, and game theoretic contexts as discussed at the beginning) suggests that the mental representation of utility is analogous to the representation of any other magnitude information, and in particular like the representation of simple perceptual stimuli, as discussed earlier. Nonetheless, as we have already noted, behaviour in our task is by no means entirely driven by context (i.e., choice behaviour is not completely insensitive to whether the choices are from the low, high, or full range). In range-frequency theory, this might be captured, for example, by allowing some prior knowledge about income, expenditure, or current savings, to determine the range of considered options.

The predictive power of the simple question measuring Direct Risk suggests that people are more or less aware about their risk preferences, but their risk perception is determined by the context. Weber (1997) argues the perceived risk is different from risk preferences and that people's perceptions of risk may be different from any theoretical risk measures. Weber also demonstrates that people act on the basis of the perceived risk and that they could have stable responses to perceived risk within particular domains (like health, finance, environment, etc.). In a similar vein, Weber and Milliman (1997) provide support for the hypothesis that factors that change and affect choice also affect risk perception and that inherent risk preference may thus be a constant for a given individual. Furthermore, Weber and Milliman suggest that risk perception may arise from a number of different concerns like for example the chance of injury or loss, magnitude and/or probabilities of losses, aspiration levels/disaster levels, controllability, gain/loss balance, different domains (situational differences), and so on.

In light of this evidence, our results presented here suggest that the context, or the set of choice options, is likely to affect how risky the available options are perceived to be. However, whilst such relative comparisons will allow people to evaluate which options are more risky than others, and even by how much, it does not provide information on how risky the overall set is; all of the options in the set may all be relatively less risky, relatively very risky, or span the entire range of risk (from 0% to 100% as in our study). Thus higher options within the available range would be seen as more risky while the lower options as less risky. Once this conceptualisation of the risky options has taken place, then people would choose according to their genuine risk preferences.

A small number of experiments have investigated the effects of context (defined as the set of available options) on decision making under risk in a way analogous to the effects described here. For example, the set of options available as potential certainty equivalents has been shown to affect the choice of certainty equivalent for risky prospects (gambles). In making a certainty equivalent judgment, participants suggest, or select from a set of options, the amount of money for certain that is worth the same to them as a single chance to play the prospect. Birnbaum (1992) demonstrated that skewing the distribution of options offered as certainty equivalents for simple prospects, whilst holding the maximum and minimum constant, influenced the selection of a certainty equivalent. When the options were positively skewed (i.e., most values were small) prospects were under-valued compared to when the options were negatively skewed (i.e., most values were large).

Benartzi and Thaler (1998, 2001) have found evidence of another effect of the choice set by studying how people allocate their retirement funds across various investment vehicles. In particular, they find evidence for a diversification bias, which they call the 1/*n* heuristic. The idea is that when an employee is offered *n* funds to choose from in her retirement plan, she divides the money approximately evenly among the funds offered. Use of this heuristic, or others only slightly more sophisticated, implies that the asset allocation an investor chooses will depend strongly on the array of funds offered in the retirement plan. Thus, in a plan that offered one stock fund and one bond fund, the average allocation would be 50% stocks, but if another stock fund were added, the allocation to stocks would jump to two thirds. Read and Loewenstein (1995) also reported that people tend to diversify equally between the set of available options.

These findings illustrate that investors have ill-formed preferences about their investments, which is consistent with Stewart et al.'s (2003) claims. Benartzi and Thaler (2002) asked individuals to choose among investment programs that offer different ranges of

retirement income (for instance, a certain amount of $900 per month versus a 50-50 chance to earn either $1,100 per month or $800 per month). When they presented individuals with three choices ranging from low risk to high risk, they found a significant tendency to pick the middle choice. For instance, people viewing choices A, B, and C, will often find B more attractive than C. However, those viewing choices B, C, and D, will often argue that C is more attractive than B.

In summary, in all these experiments, the manipulated context was related to certain properties of the distribution (like range and rank) of the magnitude attributes of the choice options (like for example, the risk of a financial prospect); and the main purpose was to investigate whether decisions are affected by such manipulations of the simultaneously and sequentially presented options. In these cases, range-frequency theory (Parducci, 1965, 1995) is consistent with the results and could serve as a plausible model explaining the data. Note also that all these experiments were also based on an range-frequency type of account, which assumes that when people judge the attractiveness of, and thus their preferences for, a risky option, they do this by comparing this option with the other available options, instead of matching it with some stable internal scale for (absolute) judgment. Thus, these findings and the evidence that we have reported and reviewed here, present another challenge to the standard rational choice theory.

Practical Applications

The results presented here also show that we can increase savings and risky investment by manipulating the range of the choice options. Therefore we accomplished our practical goal to find a way to encourage people to save as much as possible, which is important because current saving rates are much less than the necessary level (see the report by Oliver, Wyman and Company, 2001, which details the UK savings gap), and at the same time also to stimulate them to invest at a higher risk in the capital markets. The rational behind the second aim is that by investing at a higher risk people would experience the least possible decrease in their current consumption because higher market risk would bring higher expected returns and therefore would require less income portion to be saved. In order to accomplish these goals, we used manipulations of the context in which the choice options are presented. We also investigated whether risk preferences affect these financial decisions and the test indicated that these characteristics should not be ignored when financial advice is provided.

The practical relevance of such results can be utilised by using such context manipulation methods during real financial advise, because financial advisers can take a normative stance and encourage people to behave in a direction that is expected to maximise their expected welfare. This also means that our assumption is that people are in principle unable to independently and autonomously make optimal decisions about their financial future, which is what the existing empirical evidence demonstrates as well (e.g., Benartzi and Thaler, 2002). Therefore, the presented results are also a direct test of whether the various documented context effects can be used (in combination) in order to produce certain desirable social objectives. Our results also serve as a good example of how psychological phenomena and decision-making theories could be applied to solve real-world problems.

AUTHOR'S NOTE

Ivo Vlaev and Nick Chater (Department of Psychology, University College London, London, UK).

This work was supported by a grant from the Institute of Actuaries (London) to the Institute for Applied Cognitive Science, Department of Psychology, University of Warwick. We thank members of the actuarial profession Alan Goodman, Martin Hewitt, Ian Woods, and John Taylor for their comments and suggestions. Nick Chater was partially supported by European Commission grant RTN-HPRN-CT-1999-00065, the Leverhulme Trust, and the Human Frontiers Science Program.

APPENDIX A
FINANCIAL AFFORDABILITY QUESTIONNAIRE

The following questions ask you about various facts and preferences related to your personal finances. We also expect you to provide absolute numbers on your income and expenditure (in pounds). The purpose of this test is to investigate to what extent your real financial circumstances, at the moment, affect your choices in the saving and investment experiment that follows. It is essential to be as accurate and honest as possible. We greatly appreciate your cooperation and we guarantee that the information that you provide will remain strictly confidential.

Please answer the following questions:

1. What is your annual income: _____
2. Which of the following statements reflect your financial circumstances (circle the appropriate one and provide the appropriate figures):

    a) You spend less than you earn; state by how much: _____
    b) You spend exactly the amount that you earn: _____
    c) You spend more than you earn (for example by borrowing or living on credit); state by how much: _____

In order to answer these questions try to estimate to what extent your current annual income is sufficient to cover your necessities, and in particular try to figure out *by how much* your income is sufficient or insufficient to cover your annual expenses (for example, you could say that you spent around £2000 more than your current salary in order to cover your necessities).

3. Try to estimate how much you are able to save at the moment. Please write down here your average annual savings: _____
4. Here we provide you with a list of various types of spending and you have to answer how much of your current annual income is spent on each of these expenditures. There are two types of expenditure examples – essential (e.g., food and rent) and discretionary (e.g., leisure activities), and you have to give estimates of your annual

spending across these categories (for example, you can say that you spend usually £200 on food, £250 on rent, and so on).

a) Essential expenditure

- Food _____
- Rent / Mortgage _____
Utilities (electricity, gas, heat, light, water) _____
Car _____
Other transport (train busses) _____
Debt repayment _____
Communications (telephone, etc.) _____
Childcare and Schooling _____
Health _____
Repairs and Maintenance _____
Other (e.g., health and life insurance, etc.) _____

b) Discretionary expenditure

- Holiday _____
- Entertainment (e.g., cinema) _____
- Sport _____
- Hobbies _____
- Meals and Drinks _____
- Other _____
TOTAL EXPENDITURE: _____

5. What is the maximum amount that you would like to save per year: _____
6. Can you give up some of your discretionary spending in order to increase your current savings rate if it is below your preferred maximum amount indicated in question 5?
YES / NO (circle the appropriate)

In order to answer this question you need to focus again on your discretionary spending and estimate the degree to which you can readily reallocate money towards a pension. Here we also aim to test how important and essential some of these discretionary expenditures are for you (e.g., some people might be unwilling to give up certain hobbies, sport activities, etc.). Note that in the following experiment we will ask you a series of questions about your preferred savings; if you answer values that are above your current savings (provided in question 3), then we will ask you to give up some of your essential or discretionary spending in order to provide the additional capital that is required to cover the difference between your real current savings rate and the savings rate that you have indicated in some of the test questions.

APPENDIX B

Description of each question in the Prospect Relativity Experiment. The questions are grouped by the key variable that the participants were asked to select (savings or investment risk).

I. *Savings.* The were five questions asking people to choose between savings options formulated as percentage that is saved out of the hypothetical income of £25000 per year.

1. Choose how much to save without information about other variables.
2. Choose how much to save and see expected retirement income.
3. Choose how much to save and trade it off with retiring at different age and see the expected retirement income.
4. Choose how much to save and see the retirement income and its minimum and maximum variability happening because assume that 50% of the savings are invested in the stock market.
5. Choose how much to save and take different levels of risk starting from low savings and investment risk and then increase both in parallel.

II. *Risk.* Next are the five questions asking people to choose levels of risk formulated as percentage of saving invested in risky assets:

1. Choose how much to invest without information about other variables.
2. Choose how much to invest and see expected retirement income and its variability.
3. Choose how much to invest and trade-off it with retiring at different age and see the expected retirement income and its variability.
4. Choose how much to invest and tradeoff it with amount to be saved (increasing investment corresponding to decreasing savings) and see the retirement income and its variability.
5. Choose between levels of variability of the retirement income. Variability reflects different investment strategies and is increasing with the income (higher variability corresponds to higher income).

APPENDIX C
MEASURES OF RISK AVERSION USED IN THE STUDY

(1) Please indicate here how much risk you are prepared to take on a scale from 1 (not at all – only sure outcomes) to 5 (very much):

Answer: _____

(2) How much are you concerned about your financial future? Indicate on a scale from 1 (not at all) to 5 (very much):

Answer: _____

(3) Are you more or less willing to take risks than the average person?
Indicate using the following scale:
1 - much less
2 - less
3 - the same as the average
4 - more
5 - much more
Answer: _____

(4) Are you more or less concerned about your financial future than the average person?
Indicate using the following scale:
1 - much less
2 - less
3 - the same as the average
4 - more
5 - much more
Answer: _____

(5) Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50–50 chance it will double your (family) income and a 50–50 chance that it will cut your (family) income by a third. Would you take the new job? Answer with YES or NO here:
Answer: _____

If your answer to this question is “yes,” then answer only question (a) and if your answer is “no,” then answer only question (b).

(a) Suppose the chances were 50–50 that it would double your (family) income, and 50–50 that it would cut it in half. Would you still take the new job? Answer with YES or NO here:
Answer: _____

(b) Suppose the chances were 50–50 that it would double your (family) income and 50–50 that it would cut it by 20 percent. Would you then take the new job? Answer with YES or NO here:
Answer: _____

REFERENCES

Barsky, R. B., Juster, E. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics, 112,* 537-579.

Benartzi, S., and Thaler, R. (1998). Illusory diversification and retirement savings. Unpublished manuscript, University of Chicago and UCLA.

Benartzi, S., and Thaler, R. (2001), Naive diversification strategies in defined contribution saving plans, *American Economic Review, 91,* 79-98.

Benartzi, S., and Thaler, R. (2002). How much is investor autonomy worth? *Journal of Finance, 57,* 1593-1616.

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choicebased certainty equivalents. *Psychological Science, 3,* 310-314.

Camerer, C. F. (1995). Individual decision making. In J. Kagel and A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 587-703). Princeton, NJ: Princeton University Press.

Fechner, G. T. (1966). *Elements of psychophysics* (Adler, H. E., Trans.). New York: Holt, Rinehart, and Winston. (Original work published 1860).

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46,* 373-380.

Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology, 48,* 218-224.

Gigerenzer, G., and Todd, P. (Eds.) (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper and Row.

Holland, M. K., and Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception and Psychophysics, 3,* 409-414.

Jesteadt, W., Luce, R. D., and Green, D. M. (1977). Sequential effects of the judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 92-104.

Kagel, J. H., and Roth, A. E. (1995). *The handbook of experimental economics*, Princeton University Press.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47,* 313-327.

Kahneman, D., and Tversky, A. (Eds.) (2000). *Choices, values and frames*. New York: Cambridge University Press and the Russell Sage Foundation.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge University Press, Cambridge.

Kapteyn, A., and Teppa, F. (2002). Subjective measures of risk aversion and portfolio choice. Discussion Paper 11, Tilburg University, Center for Economic Research.

Laming, D. R. J. (1984). The relativity of 'absolute' judgements. *British Journal of Mathematical and Statistical Psychology, 37,* 152-183.

Laming, D. R. J. (1997). *The measurement of sensation*. London: Oxford University Press.

Lockhead, G. R. (1984). Sequential predictors of choice in psychophysical tasks. In: Kornblum, S., and Requin, J. (Eds.), *Preparatory states and processes* (27-47). Hillsdale, NJ: Erlbaum.

Lockhead, G. R., and King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 461-473.

Loewenstein, G. (2003). *Preferences*. Oxford University Press.

Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal, 92,* 805-824.

Luce, R. D., Nosofsky, R. M., Green, D. M., and Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception and Psychophysics, 32,* 397-408.

Oliver, Wyman and Company (2001). *The future regulation of UK savings and investment*. Research report conducted by Oliver, Wyman and Company and commissioned by the Association of British Insurers.

Parducci, A. (1965). Category judgment: A range-frequency theory. *Psychological Review, 72,* 407-418.

Parducci, A. (1995). *Happiness, pleasure and judgment: The contextual theory and its applications*. Mahwah, NJ: Lawrence Erlbaum.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87–131.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation, 3,* 323-343.

Shafir, E., and LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53,* 491-517.

Sher, S., and McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition, 101,* 467-494.

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

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153-181.

Stevens, S. S. (1975). *Psychophysics*. New York: Wiley.

Stewart, N., Brown, G. D. A., and Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 3–11.

Stewart, N., Brown, G. D. A., and Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review, 112,* 881–911.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1-26.

Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General, 132,* 23-46.

Vlaev, I. and Chater, N. (2006). Game relativity: How context influences strategic decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 131-149.

Vlaev, I., Chater, N., and Stewart. N. (2007). Financial prospect relativity: context effects in financial decision-making under risk. *Journal of Behavioral Decision Making*.

Ward, L. M., and Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology, 84,* 27-34.

Weber, E. (1997). The utility of measuring and modelling perceived risk. In A. A. J. Marley (Ed.). *Choice, decision and measurement: Essays in honour of R. Duncan Luce* (pp. 59-73). Mahwah, NJ: Erlbaum.

Weber, E. U., and Milliman, R. A. (1997). Perceived risk attitudes: Relating risk perception to risky choice. *Management Science, 43*, 123-144.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 37-63

Chapter 2

CHARTING THE COURSE OF SELF EVALUATIONS AND SOCIAL COMPARISONS OVER TIME

Janet A. Sniezek*[1], Phanikiran Radhakrishnan[2] and Tahira M. Probst[3]

[1](posthumous) University of Illinois at Urbana-Champaign
[2]University of Toronto at Scarborough
[3]Washington State University Vancouver

ABSTRACT

Self-evaluations of performance are important in theory and practice. In contexts with multiple persons performing the same task, the evaluation of one's own performance is expected to be a process involving judgments about the performance of others, and comparisons between one's own and others' performance. We conducted a longitudinal study tracking 79 participants' evaluations of their own and others' performance on five repetitions of a task over a four-month period. Three temporal factors that Radhakrishnan, Arrow, and Sniezek (1996) identified as influences on self evaluations of performance were examined: Temporal Perspective, Time Horizon, and Experience. In the present study, we investigated in more detail, the role of these factors, on judgments evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance as well as on social comparisons. Participants made evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance and social comparisons both had a positively leniency bias. This bias in self evaluations and social comparisons decreased when estimates were made (a) after performance than before; (b) closer to the performance event than farther away from it; and (c) with increasing experience. However, evaluations of only one's own performance were more variable with changes in the temporal factors. Further, the increase in bias with longer time horizons was reduced considerably with increasing experience. Changes

* Address Correspondence to: Phanikiran Radhakrishnan, Management Department, University of Toronto at Scarborough, 1265 Military Trail, Scarborough ON Canada M1C1A4, Fax 416 287 7363. phanira@utsc.utoronto.ca. www.utsc.utoronto.ca/~phanira. Voicemail 416 287 7319.

in inter- and intra-individual validity followed those for bias. Interestingly, changes in solo evaluations over time were similar to those for social comparisons.

Charting the Course of Self Evaluations and Social Comparisons over Time Evaluative responding is assumed to be an automatic part of human behavior (e.g., Markus and Zajonc, 1985; Jarvis and Petty, 1996). Festinger (1954) made the astute observation that the evaluative process tends to involve social comparison when others are present. The act of comparing some aspect of oneself to another is meaningful, and may change behavior. For example, it can fulfill a wide variety of needs (Helgeson and Mickelson, 1995; Sedikides and Strube, 1997; Wood, 1989). And when the comparisons to others concern performance, they can change one's performance, for example by cueing competition (Seta, 1980). Thus, it should be natural and useful for people to evaluate their own performance in comparison to that of others performing the same or similar tasks.

Several empirical studies indicate that social comparison is indeed important to those in social performance settings. For example, Levin and Levin (1973) gave students the choice of obtaining either one of two pieces of information about their performance on a test: percent correct or score in percentiles. In general students preferred the social comparison feedback. In a study conducted in a similar setting, Suls and Tesch (1978) found that students were particularly interested in learning the average class score on the test.

Not only are people interested in social comparison information, they are affected by it in ways that may alter subsequent behavior concerning the task. For example, perceived competence has been shown to be affected more strongly by information about one's competence relative to others than information about absolute performance quality (Sansone, 1986). People are particularly motivated to compare their performance to that of others when there is a mix of knowledge of feedback and uncertainty (Brickman and Berman, 1971)-a condition that characterizes most performance settings.

But what do these social comparisons of performance look like? How do people make them? In this work, we seek to understand more fully the way in which people evaluate their performance in social contexts. We examine social comparisons of performance, and compare them to solo evaluations, i.e., judgments of one's performance without explicit reference to any other performers. In addition, we include the study of social performance judgments, the appraisals of the performance of others that are inherent to social comparisons of performance. The picture of self evaluation that we attempt to provide illustrates the relationships among these three types of judgments-solo evaluations, social performance judgments, and social comparisons. It also shows the correspondence of the judgments to actual performance levels.

To obtain a more complete picture of the way in which people evaluate their performance, we ask how social comparisons of performance-as well as self evaluations and social performance judgments--change over time as performance itself changes. We turn next to a discussion of the role of time in shaping self evaluations of performance.

SELF EVALUATION OVER TIME

The importance of comparing one's performance to that of others prior to the task is illustrated in a line of research initiated by Spence and Helmreich (1983). They identified

differences in achievement motivation that manifest themselves in terms of either mastery or performance goals. Whereas the relevant standard with mastery goals is internal, it is external with performance goals. People with performance goals desire to be better than others. While those with such a competitive personality may enjoy superior performance (Harackiewicz, Barron, Carter, Lehto, and Elliot, 1997), they also suffer particularly at lower levels of perceived competence (Dweck and Legget, 1988). This research area affirms the importance of the performance of others as a comparison standard before performing the task. But it does not reveal whether those with such goals have a good sense of how others will perform, and how they will stand in comparison. Further, there is no indication of how their views will change as the performance event draws near, passes, and reappears.

Few studies provide thorough data on social comparisons of performance, and none that we have found examine social comparisons of performance over time. In those studies where the standard of comparison is social performance, time is typically ill defined. That is, the performer is asked about his or her relative standing with respect to the average or entire distribution of co-performers for some indefinite period (e.g., Klar and Gilladi, 1997). It is not customary to define the relevant period of performance to the participant, or to report elapsed time between the task and the social comparison judgment in research articles.

Yet, time can make an enormous difference in the evaluation of one's own performance. Although neither their theory nor data concern social comparisons in particular, Radhakrishnan, Arrow, and Sniezek (1996) address the role of time in detail. They present a model proposing that the accuracy of self-evaluations of performance varies with temporal variables due to shifts in motivational and informational influences. Specifically, Radhakrishnan et al. hypothesize that the performer's goals and information available to the performer vary with two time variables. These are (a) *Temporal Perspective,* or the direction of time between evaluation and task performance (i.e., before vs. after task performance) and (b) *Time Horizon,* or the amount of time between evaluation of performance and task performance.

According to the Radhakrishnan et al. model, information and motivational states shift as the performer moves from a future to a past perspective. Prior to the task, the performer has some uncertainty about task requirements, yet wants to maximize performance. Both information and motivation change once the task has been completed. Then the performer has information about the task but can no longer make a difference in the quality of performance. One salient motive at this point is to defend one’s self esteem. These shifting motivations and information levels combine to make people more overconfident about their performance evaluations prior to the task, and much less confident once the task has been completed. Radhakrishnan et al. report an empirical study that shows support for the effects of Temporal Perspective predicted by their model. Students overestimated their score on the first quiz before taking it, and while still overconfident in their estimates, became significantly more accurate afterwards.

As for the role of Time Horizon, Radhakrishnan et al. (1996) hypothesized that the magnitude of optimistic bias increases with time horizon. That is, self-evaluations should be most optimistic long before and long after the time of the task-when the least information is available from the environment or memory. Their data show support for their prediction: the students were most optimistic well in advance of the performance event with the amount of bias decreasing as the event approached. Data on related phenomena point to similar conclusions. Nisan (1972) reports that people were more risky in their predictions and had higher expectations of success when they made predictions four weeks before, than immediately before task performance. In a study examining the effects of time horizon on prospective and retrospective judgments of performance

on hypothetical tasks, Gilovich, Kerr and Medvec (1993) found that evaluations were most optimistic as time to, and form a performance event increased. A final prediction from the work of Radhakrishnan et al., (1996) was a dampening of the effects of Temporal Perspective and Time Horizon with repeated task experience. This was expected to occur as long as experience allowed one to become well-informed about task demands as well as one's performance. This too was supported in their empirical work; the negative effects of time horizon and temporal perspective on accuracy were strongest for the first of three quizzes and weakest for the third quiz.

Two limitations of the previous study (Radhakrishnan et al., 1996) are noteworthy. First, time horizons for self-evaluations before and after the quizzes were not symmetrical. Consequently time horizon and temporal perspective were confounded. Second, there was no manipulation of time horizon following the task. But considering the parallel performance of multiple people, perhaps the most important restriction of our study was that the self-evaluations were made with respect to only a performance scale, and not to the performance levels of others. Although participants were asked to estimate the average for the sample, they never provided comparisons of their performance to this average. This is particularly unfortunate given that social comparisons were likely to have been automatically generated and quite important to the participants. It is therefore unclear how similar solo evaluations of performance are to social comparisons in terms of the effects of time horizon and temporal perspective. The present research is designed to provide an empirical answer. Next, we turn to theoretical bases for differences between these self evaluations.

SOLO EVALUATIONS VS. SOCIAL COMPARISONS

There are four reasons to suspect that the two types of self-evaluations- solo evaluations and social comparisons -- differ. The first concerns the availability of a standard for comparison; the second concerns motives regarding one's performance in comparison to the standard. Social comparison, by definition, involves only comparative judgment. Absolute judgment is possible only with non-social self-evaluation, as when one assigns a value to performance to represent the number of products produced. It is well known that absolute and comparative judgment differ (Biernat, Manis, and Kobrynowicz, 1997). Thus we can assume that the self-evaluation process is unique when solo evaluation results from absolute judgment. Of course this is not always the case; a comparison process is inevitable when one has defined performance goals with the same metric. But direct comparison is inhibited by vague representations of goals (Scheier and Carver, 1983), thus comparative judgment is not necessarily part of self-evaluation for tasks with great personal importance. In other words, solo evaluation of the number of books read results from a comparative judgment process if one's goal is to "Read 25 books", but not if one's goal is to "Read a lot".

Nevertheless, the processes leading to solo evaluation are different from those leading to social comparisons even when both kinds of judgments are comparative in nature. This is because, with a solo evaluation, behavior regarding the comparison is directed at reducing the discrepancy between the goal and judged performance levels. From a control theory perspective on self-regulation, performing above or below the standard triggers a change in behavior directed at achieving the standard (Carver and Scheier, in press). In contrast, there is no inherent motive regarding self-other discrepancies. Seta (1982) reports data showing how one can be

indifferent to assorted discrepancies: performance levels of coactors had no effect on one's own performance if the level was inferior, identical, or very superior. Alternatively one could strive to achieve a large, or small discrepancy. That is, one might prefer to be like everyone else, to be much better, or to be much worse.

A third reason to consider that the processes of solo evaluations and social comparisons may be distinct concerns one's degree of control over the comparison standard. One's goal is under one's control while the performance of the referent group is not (except, perhaps, under extraordinary conditions). Finally, there is a general difference in uncertainty surrounding the comparison standard: One knows one's goal but one must estimate the performance (or the goal) of the referent group. To summarize the differences between solo evaluations and social comparisons, the latter always involve comparative judgment while the former may or may not. Even if comparative judgments are involved in some types of solo evaluation, it is still a distinct process from social comparison. Specifically, preferences and behaviors concerning the self and a standard, and the degree of control and uncertainty regarding the standard differ greatly for solo evaluations and social comparisons. Thus, there are reasons to expect that different pictures of self evaluation of performance will result from solo evaluation and social comparison.

Social Performance Judgments. Another set of research questions concerns social performance judgments, or the performer's appraisal of the co-actors' performance. In the present study, we compare solo evaluations with social performance judgments. As Prentice (1990) commented, it is surprising how few social psychological studies directly compare judgments about the self and judgments about others. Here we define a solo evaluation as estimate of one's own score and a social performance judgment as an estimate of the average score for the class. Our general task is to identify differences between participants' solo evaluations and social performance judgments. Specifically, we examine how these judgments pertaining to the self and others vary with Temporal Perspective, Time Horizon, and Experience. Although Radhakrishnan et al. (1996) showed that these factors could combine to influence solo evaluations, there are no comparable data for social performance judgments.

CURRENT STUDV

The current research effort attempts to understand more fully the pattern of changes in self-evaluations of performance overtime and in a social context. Whenever one has co-performers, there are really two social components in the process of evaluation: How one judges the performance of others, and how one compares one's own performance to that of others. Thus, we obtain both social performance judgments (evaluations of the performance of the reference group in the form of estimates of the class average) and social comparisons (evaluations of one's performance relative to that of others in the form of estimates of one's percentile rank). The social judgments and social comparisons are in addition to evaluations of one's own performance with respect to an absolute standard. Data for all three types of judgments make it possible to determine whether social comparisons change with time and experience due to perceived changes in judgments about the performance of oneself, the reference group, or both. For example, suppose I decide that I will do more poorly than the average member of the sample will. This might be because I now think less of my performance, or because I now judge the sample even better than I previously thought.

The first goal of this research is to report systematic variations judgments of performance within the period surrounding the task performance event (i.e., temporally close vs. temporally farther away from the event and before vs. after the event) as well as over repetitions of similar tasks. Put another way, our purpose is to describe and explain the three types of performance judgments -solo evaluations, social performance judgments, and social comparisons as a function of Time Horizon, Temporal Perspective, and Experience. Of special interest are differences in the patterns for solo evaluations, social comparisons, and social performance judgments.

Another goal is to evaluate the quality of these three types of judgments with respect to actual performance. Of course, this requires an objective criterion measure of performance for the entire sample of performers. Given a criterion measure, it is possible to describe the correspondence a judgment and the criterion in two different ways. First is by calculating the accuracy or distance between the judgment and criterion. The second way is to assess validity by computing the correlation between the judgment and criterion. Although correlational and distance measures are related, they are not the same (Sniezek and Reeves, 1986) and will not necessarily be affected by the same variables. Thus we use both to assess the correspondence between each performance judgment and the appropriate criterion.

Our study was conducted in a 15-week advanced course in psychology. Several researchers (Carver and Scheier, 1994; Campion and Lord, 1982; Radhakrishnan et al., 1996) have noted the advantages of conducting longitudinal research in an academic, classroom setting. A university course provided an ideal setting in the case of the present study because it involved a large number of persons repeatedly performing identical tasks in parallel.

More importantly, the classroom setting permits broad generalization. Consider how natural settings in which coactors perform identical tasks independently but concurrently tend to offer feedback, whether formal, as in statistics in sports performances, or informal, as in observations or personal communications. The course used for this study provided substantial feedback about participants' own quiz scores, and the distribution of quiz scores for the sample the participant was a part. This feedback promoted learning about the task as well as about the quality of evaluations. Both absolute and relative standing contributed to the determination of valued outcomes for participants -- course grades -- feedback about one's own score, the class average, as well as each participant's place in the distribution were all meaningful to participants. A final useful feature was that the classroom context allowed for social interactions among participants over the time-period during which the study was conducted. In sum, the classroom setting made social comparisons and social performance judgments meaningful and possible for participants.

Participants made three types of performance estimates - solo evaluations,. social performance judgments, and social comparisons. These judgments were made for each of five quizzes at multiple points over the duration of the course. This setting made it possible for us to vary Temporal Perspective, Time Horizon, and Experience. We varied Temporal Perspective by asking participants to make estimates of their performance before and after each task performance event, but before receiving feedback about that particular task performance. In addition, Time Horizon, or the amount of time elapsed between the point of evaluation and the performance event was varied by having participants make judgments about their performance several weeks, three weeks, a few days, and a few min. before, and after, each performance event. The third factor, Experience, concerned systematic changes in evaluations of performance over five discrete performance events of similar tasks.

The performance task was an in-class, non-cumulative quiz consisting of 20 multiple-choice items with four-alternatives each. The task was highly involving for the student-participants enrolled in the course who were presumably motivated to maximize performance. Although a multiple choice quiz was not entirely novel for upper-class university students, the combination of a new course, a new instructor, a new teaching-assistant, and novel content produced sufficient uncertainty about task performance, made performance evaluations non trivial, and feedback about task performance informative. Five regularly scheduled quizzes provided multiple performance trials so that we could investigate the dynamic aspects of performance evaluations. A special advantage of this task was that quiz scores provided an objective external criterion for evaluating the quality of performance evaluations, and offered accurate and unambiguous feedback that promoted favorable learning conditions. This in turn enabled subsequent evaluations more accurate, less biased and more valid.

METHOD

Participants

The 79 undergraduates enrolled in the upper-level course made multiple evaluations of their performance on each of the five in-class quizzes as partial fulfillment of course requirements. Participants were informed that their responses would have no impact on their grades and that their evaluations would be kept sealed until the course was completed and final grades were submitted to the registrar. This information was provided to ensure, and assure participants of the confidentiality and consequences of their evaluations. All student-participants consented to the use of their data for research purposes. Participants were assured that providing consent to the use of their evaluations for research purposes was not part of the class requirement. The course was taught by one of the authors. However, a research assistant unrelated to course administrative duties collected students' performance evaluations.

Procedure

Several estimates were collected before and after performing each quiz. The usual sequence was to obtain the initial judgments for a quiz two-three weeks before (on the date of the previous quiz, with the exception of the first quiz). Additional evaluations were made two days before, a few min. before on the day of the quiz, a few min. after the quiz, and finally, again two days after the quiz. Note that all evaluations made after the quizzes were done prior to receiving feedback.

Measures

Elicited Measures. Participants provided three types of performance estimates at each time-point of data collection, by writing privately, on a form containing instructions for each type of estimate. One performance estimate was the solo evaluation, an estimate of one's own

raw score, or specifically, the number correct out of 20. A second type was the social performance iudgment, the estimated class average or specifically, the average number correct out of 20 for the entire class. A third type was the social comparison iudgment, an estimate of one's percentile rank, or the percentage of class members scoring below the individual's estimated raw score.

Derived Measures. To describe fully the quality of each type of judgment, we computed the accuracy and validity for each of the three types of estimates we elicited.

Accuracy. Both the absolute level of accuracy and the direction of inaccuracy are of interest in describing the quality of performance judgments. We defined accuracy to include Mean Absolute Deviation (MAD) and Mean Bias (MB). For a set of judgments $\{J_1, J_2,\ldots J_i,.. J_k)$ from k individuals with criterion values $\{Y_1, Y_2,\ldots Yi,.. Y_k\}$:

$$MAD = 1/k \sum_{i=1}^{K} | J_i - Y_i | \qquad [1]$$

Mean Absolute Deviation (MAD) quantifies the magnitude of error in a set of judgments. It is useful to show the extent to which estimates deviate from the criterion. However, it does not yield any information about the direction of the deviation.

In contrast, Mean Bias indicates whether the judgments overestimate or underestimate the criterion. It is equal to the mean signed deviation between the judgment and criterion:

$$\text{Mean Bias} = 1/k \sum_{i=1}^{K} | J_i - Y_i | \qquad [2]$$

MAD and Mean Bias were computed for solo evaluations (estimates of own scores) and social comparisons (estimates of percentile ranks). The criterion variables were, respectively, actual own score and actual percentile rank. For class averages, MAD and MB were obtained by taking the absolute difference between the estimates of the class average and the obtained class average and by taking the signed difference between the two respectively. Mean bias allowed us to examine the direction of inaccuracy and its changes over time and judgment type are examined which we report in detail in this paper.

Indirect social comparisons. We define indirect social comparison as the discrepancy between a solo evaluation and a social performance judgment. Whereas the performer makes a direct comparison with social comparisons, we make the comparison with indirect social comparison. We computed both the signed deviations and absolute differences between solo and social performance judgments. These discrepancies provided an alternative perspective on differences in judgments of self vs. others. In addition, they allowed us to diagnose whether social comparisons were inaccurate due to poor social judgment or poor comparison ability.

Validity. There are two meaningful types of validity coefficients. The correlation between the judgment and criterion over a set of persons, called *inter-individual validity*, tells whether the members of a sample can differentiate among their actual performance levels. As such, it is an index of quality at the level of the sample. An alternative, *intra¬individual validity*. is assessed by the correlation between a single individual's judgments and that individual's

actual scores over multiple tasks. This provides information about quality at the individual level -- whether an individual's judgment can capture true changes in his/her criterion performance over repeated evaluations of a performance event and over repeated performance of similar tasks.

We assessed *inter-individual validity* by correlating performance estimates with actual performance scores over persons within each trial. Inter-individual validity was computed for solo evaluations (by correlating estimated and obtained raw scores for the individual) as well as for social comparisons (by correlating estimated and actual percentile ranks). It is not possible to calculate an analogous validity coefficient for social judgments because the actual class average does not vary within a trial. For *intra-individual validity* coefficients, we computed a correlation between solo evaluations and the corresponding actual scores for each individual.

RESULTS

Overview of Analyses

Our analyses are organized around the three main questions motivating this study. While all the questions pertain to the effects of three factors--Temporal Perspective; Time Horizon, and Experience--on self-evaluations, they differ in terms of the particular self evaluation judgments of interest. The first question asks how solo evaluations (estimates of raw scores) and social performance judgments (estimates of class averages) varied in accuracy and validity with changes in Temporal Perspective, Time Horizon, and Experience. We analyzed solo evaluations and social performance judgments as two levels of a within-subjects factors because we were interested in whether the three factors affected the accuracy and validity of appraisals about the performance of oneself and that of others differently.

The two remaining sets of analyses help us identify when individuals see themselves as most different from, and most similar to their peers. Specifically, they address questions about the way in which Temporal Perspective, Time Horizon, and Experience influence social comparisons. The second question addresses the effects of these factors on the accuracy and validity of direct social comparisons (estimates of percentile ranks) while the third question concerns their impact on indirect social comparisons (discrepancies between solo evaluations and social performance judgments). The combination of these two analyses allows us to determine the extent to which social comparisons are inaccurate due to poor social judgment or due to poor comparison ability. In addition to the third question about the accuracy and validity of solo and social performance judgments, we were interested in the discrepancy between these two judgments as a dependent variable in and of itself. This discrepancy reveals the extent to which people judge their own performance to be different from that of the average other person. Thus, we also performed separate analyses on the discrepancy between solo and social performance judgments.

Question 1. Accuracy and Validity of Solo Evaluations and Social Performance Judgments

Changes in Bias. To examine the separate and interactive role of Temporal Perspective, Time Horizon, and Experience, we conducted three different sets of analyses on measures of mean bias'. First, we tested the role of a short Time Horizon and Experience on the mean bias in solo evaluations and social performance judgments. We used type of Estimate (Two Levels: Solo vs. Social Performance) x Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) x Experience (Five Levels: Quizzes 1 through 5) [6] This resulted in a 2 x 3 x 5 repeated measures ANOVA factorial design.

In addition, we tested the role of a long prospective Time Horizon and Experience on the bias of solo evaluations and social performance judgments [7]. This analysis had the same two types of Estimates (Solo vs. Social Performance), but a larger range of Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs. Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). This was a 2 x 4 x 3 repeated measures ANOVA [8].

Finally, we investigated the roles of Temporal Perspective, Time Horizon, and Experience. We conducted 2 x 2 x 2 x 4 repeated-measures ANOVA on the bias of two types of Estimates (Solo vs. Social Performance) with Temporal Perspective (2 Levels: Pre-task vs. Post-task), Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5) [9]. Due to different combinations of missing data, the three sets of analyses have different sample sizes. The design examining the effects of a short Time Horizon has 30 participants, the design exploring the effects of a long prospective Time Horizon has 38, and finally the design examining the effect of both Time Horizon and Temporal Perspective has 23 participants with complete data.

Temporal Perspective. Recall our prediction that post-task evaluations would be less biased than pre-task evaluations. As can be seen in Figure 1, both solo and social performance judgments had higher mean bias when they were generated before the task was performed (M=.92) than after it was performed (M=.50, $F(1,21)=31.69$, $p< .001$).

We predicted that solo and social performance judgments generated at points that were temporally closer to the performance event would be less biased than those made at temporally distal points. Again we found support for our predictions. Figure 2 shows that with a shorter time horizon (i.e., when evaluations were made at the three time-points before the performance event) mean bias decreased as individuals approached the task performance event, $F(2,56)=57.85$, $p<. 0001$. Similarly, Figure 3 shows that with a longer time horizon (i.e., when evaluations were made at four timepoints before the performance event) mean bias decreased much more dramatically, $F(3,108)=51.21$, $p<. 0001$. Figure 1 also shows that mean bias decreased with changes in Time Horizon even when evaluations were made before versus after the task performance event but before performance feedback from an objective source, $F(1,21)=20.61$, $p <. 0001$. In general, mean bias decreased as evaluations were made at points that were temporally closer to the performance event than those that were made farther away from it.

We predicted that evaluations made after experience and feedback (i.e., in later quizzes) would be less biased. We found that as students gained experience evaluating their performance and received feedback about each performance event, mean bias in their solo evaluations and social performance judgments decreased. Further, mean bias decreased with increasing Experience when the Time Horizon was short, $F(4,112)=30.08$, $p< .0001$ (see

Figure 2), when the Time Horizon was long, $F(2,72)=21.73$, $p< .0001$ (see Figure 3) and when Temporal Perspective varied with Time Horizon and Experience, $F(3,63)= 10.93$, $p< .0001$ (see Figure 1).

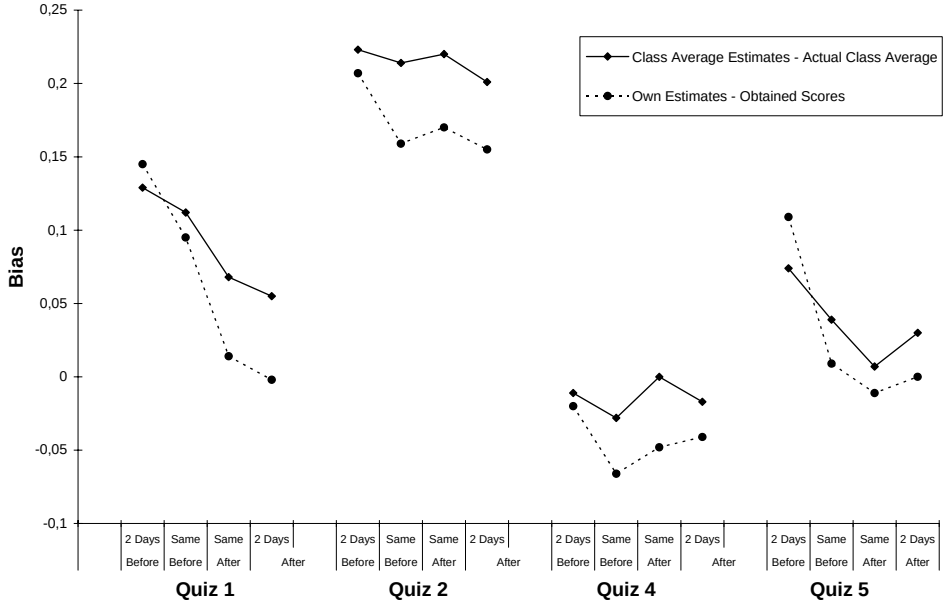


Figure 1. Bias in solo evaluations and social performance judgments as a function of Temporal Perspective and brief Time Horizons.

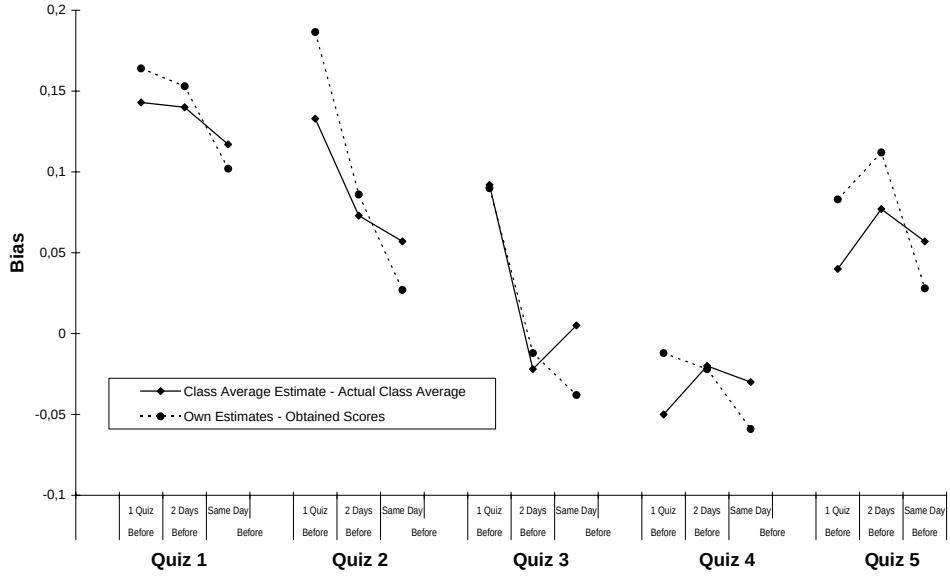


Figure 2. Bias in prospective solo evaluations and social performance judgments over short Time Horizons.

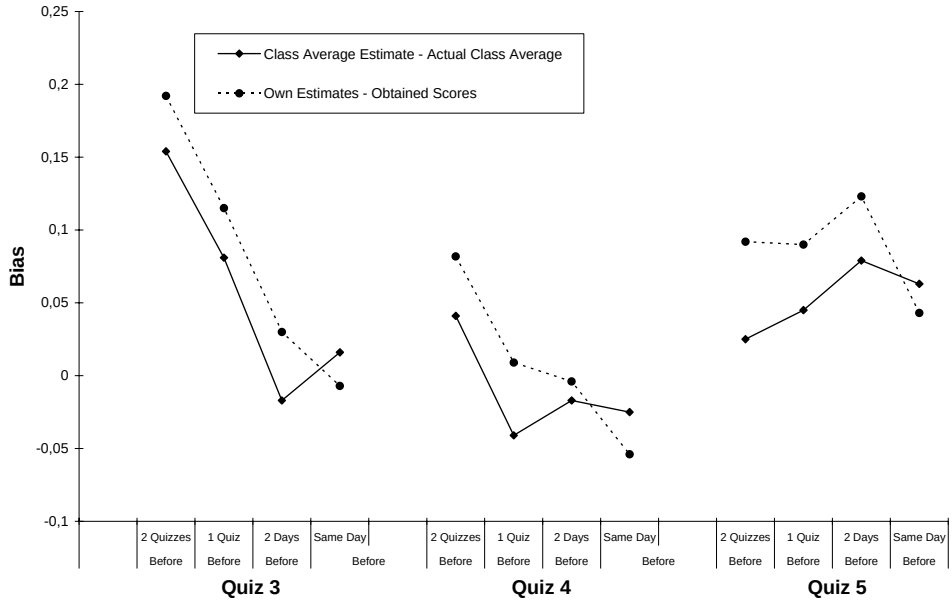


Figure 3. Bias in prospective solo evaluations and social performance judgments over long Time Horizons.

We predicted and found that the effects of Time Horizon on the bias of evaluations attenuated as individuals gained experience evaluating their performance (i.e., Temporal distance by Experience interaction). This was true for both solo and social performance judgments. Further, we found this pattern when the Time Horizon was short, $F(8,224)=9.97$, $p<.0001$ and when the Time Horizon was long, $F(6,216)=16.77$, $p< .0001$. Figures 2 and 3 show that at Quizzes 4 and 5, there was no significant difference ($p> .05$) in the mean bias of evaluations that were made farther away from the performance event than those that were made closer to the performance event. In contrast, at earlier quizzes (i.e., 1,2,3) there was a significant difference ($p< .05$) in the mean bias of evaluations that were made temporally closer, when compared to those that were made temporally farther away from the performance event.

In addition, we found a three-way interaction between Time Horizon, Experience and Temporal Perspective, $F(3,63)=3.63$ $p< .02$. As can be seen in Figure 1 post-hoc analyses showed that although increasing experience attenuates the bias caused by increasing by temporal distance, the bias does not completely disappear - it reappears when retrospective judgments are made. This suggests that memory distortion may also play a role in shaping evaluations, despite increasing experience with evaluating performance and receiving feedback regarding the quality of one's evaluations.

We found that across the different Time Horizons, solo evaluations had larger and more consistent decreases in their mean bias than did social performance judgments. This was supported by a significant interaction between Type of Estimate and Time Horizon when the Time horizon was short, $F(2,56)=22.71$ $p<. 0001$, when the Time horizon was long, $F(3,108)=19.38$ $p<.02$ (see Figure 1).

It is of interest to note that in general, people over-estimated their performance rather than under-estimated it (see Figures 1, 2, and 3). An exception is that overestimation

disappeared at points shortly in advance of Quizzes 3 and 4 (see Figure 3). One possible explanation for this is the difficulty of Quiz 2; the class average for Quiz 2 was lower than that of the other four quizzes ($p < .05$). Therefore, students may have under-estimated their performance on Quizzes 3 and 4 following feedback from Quiz 2 in anticipation that Quizzes 3 and 4 would be as difficult as Quiz 2. However, once Quiz 3 was completed and feedback was received, the expectation was revised.

In addition, we also found that the relative degree of bias in solo evaluations and social performance judgments changed with Time Horizon. The dominant finding prior to the quizzes was for more bias in solo evaluations than in social performance judgments. This pattern was evident when estimates were made two days before the quiz, one quiz before, and two quizzes before (see Figures 1, 2, and 3). However, directly before each quiz, the pattern reverses: solo evaluations show lower levels of positive bias than did estimates of the class average. This was true when for the relative differences in the bias of solo and social performance judgments analyzed with a short time horizon, $F(2,56)=22.71$, $p < .0001$ (see Figure 2), a long time horizon, $F(3,108)=19.38$, $p < .0001$ (see Figure 3) and in the context of temporal perspective $F(1,21)=6.63$; $p < .02$ (see Figure 1) [10]

Changes in Validity. To test our hypotheses about changes in the *inter-individual validity* we conducted a pattern analysis using binomial probabilities. We tallied the number of times the hypothesized pattern of correlations matched the obtained pattern of correlations by making multiple pair-wise comparisons. For example, if the correlations between estimates (and actual scores) made before Quiz 1 were .33, .21, .23 and those made after Quiz 1 (but before receiving feedback on Quiz 1) were .35, .29, .20, we concluded that five out of the six comparisons matched our predicted pattern. To test our hypotheses about the changes in the *intra-individual validity*, of solo evaluations we computed for each individual, a correlation coefficient between that person's solo evaluations and actual scores. These estimates were ones made at three different time horizons before the quiz: one quiz before the target quiz, two days before the target quiz, and immediately before the target quiz." We converted these coefficients to z scores and conducted a repeated measures ANOVA using Time Horizon as the 3-level within-subjects variable. For both intra and inter-individual validity, we used only those solo evaluations that were made before receiving feedback about actual performance [12].

We wanted to examine the effects of Temporal Perspective, Time Horizon and Experience on the *inter-individual validity* of solo evaluations. We hypothesized that the validity of estimates made after performing the task should be greater than that of estimates made before performing the task. As predicted, post-task solo assessments were highly correlated with actual performance than were pre-task predictions. The obtained correlation coefficients matched the expected pattern of correlations 41 times out of the 47 possible comparisons, resulting in a binomial probability of $p < .0001$, $({}^{47}C_{41})(.5^{41})(.5^{6})$. We also found that mean pre-task coefficients were lower than post-task coefficients. Using a Fisher r to z transformations we found that the mean pre-task correlation for solo evaluations was lower ($r = .30$) than the post-task correlation ($r = .45$).

We hypothesized that the predictive validity of self-evaluations should increase with decreasing Time Horizon. That is, estimates made temporally closer to the time of task should be more predictive of actual performance than should estimates made at temporally distant points. We expected that when estimates were made temporally closer to the quiz (in either direction), the correlation between those estimates and performance should correspondingly increase. We found that correlations between solo evaluations and actual scores (aggregated over Quizzes) decreased

as Time Horizon from the Quiz increased in either direction. These correlations followed the predicted pattern 50 out of the 63 comparisons, $p < .0001$, $({}^{63}C_{50})(.5^{50})(.5^{13})$.

We also predicted that as individuals gained experience predicting their performance, their estimates of performance should become increasingly valid. Our hypothesis regarding the effects of Experience on participants' ability to align their predictions of performance to their actual performance was also supported. Correlations across the five quizzes increased. Solo evaluations for later quizzes were more highly correlated with actual scores than were solo evaluations for earlier quizzes. The data for solo evaluations matched the predicted pattern 34 out of 48 times, $p < .01$, $({}^{48}C_{34})(.5^{34})(.5^{14})$.

Intra-individual validity coefficients were computed for each solo evaluation generated at three different time points (i.e., pre-performance, pre-feedback time points). There were no significant differences between the three time horizons. The validity of solo evaluations made one quiz before the target quiz was -.20, that of evaluations made 2 days before the target quiz was -.25 and the validity of solo evaluations made immediately before the target quiz was -.17.

Question 2. Social Comparisons

Changes in Bias. As for solo and social performance judgments, we conducted three sets of repeated measures ANOVAs on mean bias of social comparisons. One design examined the effects of Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) and Experience (Five Levels: Quizzes 1 through 5) on the bias of percentile ranks. The second examined the effects of a longer Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs.

Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). The third examined the role of Temporal Perspective (Pre vs. Post Task) a short Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5) on the bias of percentile ranks.

We predicted that percentile rank estimates made after task-performance should be less biased than those made prior to task-performance. We found support for this hypothesis. Figure 4 illustrates the general finding that students were less biased in evaluating their performance after task performance than before task-performance, $F(1,21)=10.86$ $p < .005$. In addition, we also found two significant interactions. Figure 4 also shows that the biggest difference in bias between pre- and post-task percentile ranks was in the first quiz. This was supported by a significant interaction between Experience and Temporal Perspective, $F(3,63)=6.69$ $p < .002$. When social comparisons are generated after task performance, mean bias in percentile ranks does not change much with increasing time horizon. This pattern was supported by a significant interaction between Temporal Perspective and Time Horizon, $F(1,21)=10.25$, $p < .005$.

We predicted that percentile rank estimates made at points that are temporally closer to the performance event should be less biased than those made at temporally distant time-points. We found support for this hypothesis. In general, when individuals estimated their percentile ranks at points that were temporally closer to the time of task performance they were less biased. We found that bias in percentile ranks decreased when Time Horizon was short $F(2,56)=12.98$, $p<.001$ (see Figure 5), long $F(3,108)=14.12$, $p<.001$ (see Figure 6), and when estimates of percentile ranks were made prospectively or retrospectively, $F(1,21)=4.62$, $p<.05$ (see Figure 4).

We predicted a reduction in bias as individuals gained more experience estimating percentile ranks. However, we did not find the predicted effects for Experience. We predicted that with increasing experience, the effects of Time Horizon on the bias of percentile ranks

should attenuate. Again, we did not find the predicted interaction between Experience and Time Horizon [14].

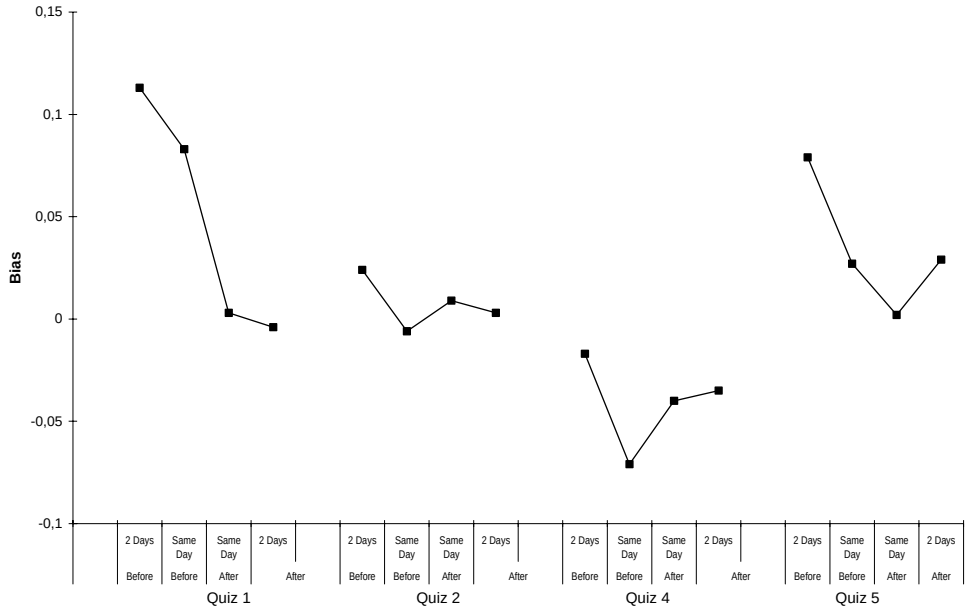


Figure 4. Bias in social comparisons as a function of Temporal Perspective and brief Time Horizons.

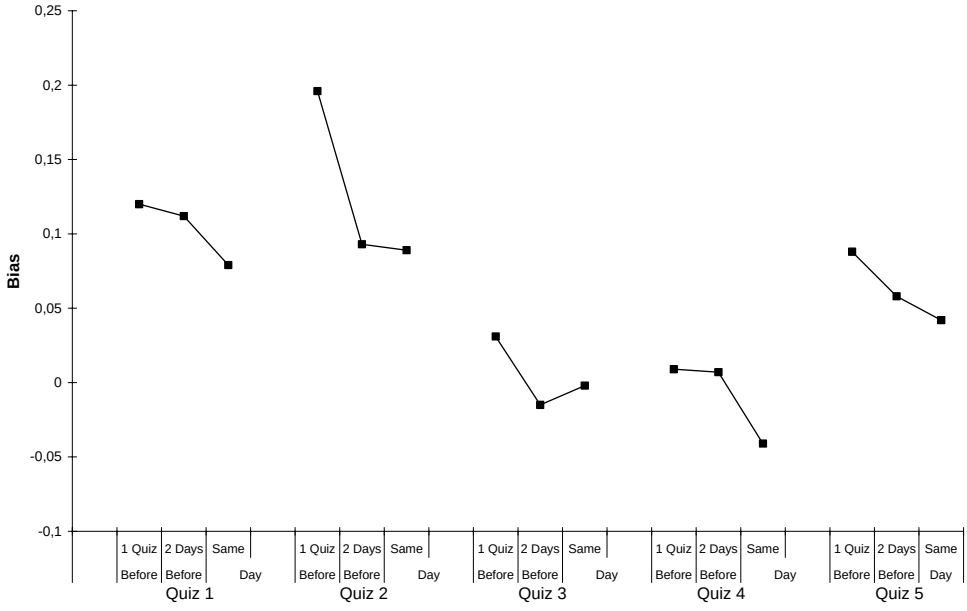


Figure 5. Bias in prospective social comparisons over short Time Horizons.

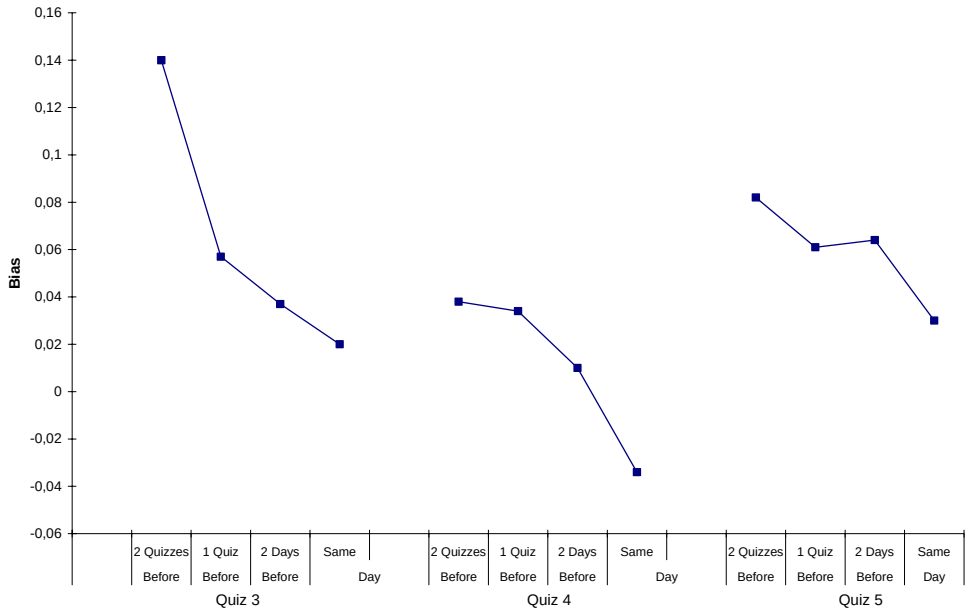


Figure 6. Bias in prospective social comparisons over long Time Horizons

Changes in Validity. As for solo evaluations, we conducted similar analyses to examine the changes in inter and inter-individual validity of percentile ranks. In general, our results supported our hypotheses for changes in the inter-individual validity. To summarize, the changes in the validity coefficients for solo evaluations and social comparisons were similar (see Figure 7). Temporal Perspective affected the validity of social comparisons. Estimates of percentile ranks made before task performance had lower correlations with actual percentile ranks than did estimates of percentile ranks made after task performance. Obtained correlations matched the hypothesized pattern of correlations 33 times out of the possible 46 comparisons, resulting in an binomial probability of $\underline{p}<.002$, $({}^{46}C_{33})\,(.5^{33})\,(.5^{13})$ [15]. Similarly, the Fisher r to z transformation showed that the pre-task correlation for percentile rank estimates was lower ($\underline{r}=.33$) than the post-task correlation ($r=.39$). Correlations between estimated and actual percentile ranks also increased as Time Horizon became shorter. The obtained pattern of correlations followed the pattern 40 times out of the 64 possible comparisons, $\underline{p}<.02$, $({}^{64}C_{40})(.5^{40})(.5^{24})$. Thus, the validity of percentile ranks increased with decreasing Time Horizon. Correlations between estimated and obtained percentile ranks increased with greater Experience. The validity coefficients for percentile ranks matched the predicted pattern 32 out of 49 times, $\underline{p}<.02$, $({}^{49}C_{32})(.5^{32})(.5^{17})$.

The changes in intra-individual validity coefficients for percentile ranks across the three Time Horizons were marginally significant, $F(2,1,04)=1,9$ $p=.15$. In general, there was an increase in the validity of percentile ranks as time horizon decreased. The validity of percentile ranks generated one quiz before the target quiz ($M=-.35$) were lower than those made 2 days before the target quiz ($M=.18$) which in turn were lower than those made immediately before the target quiz ($M=.34$). A post-hoc t-test suggested that the validity of percentile ranks generated immediately before the target quiz were significantly ($p<.05$) higher than from those made one quiz before the target quiz.

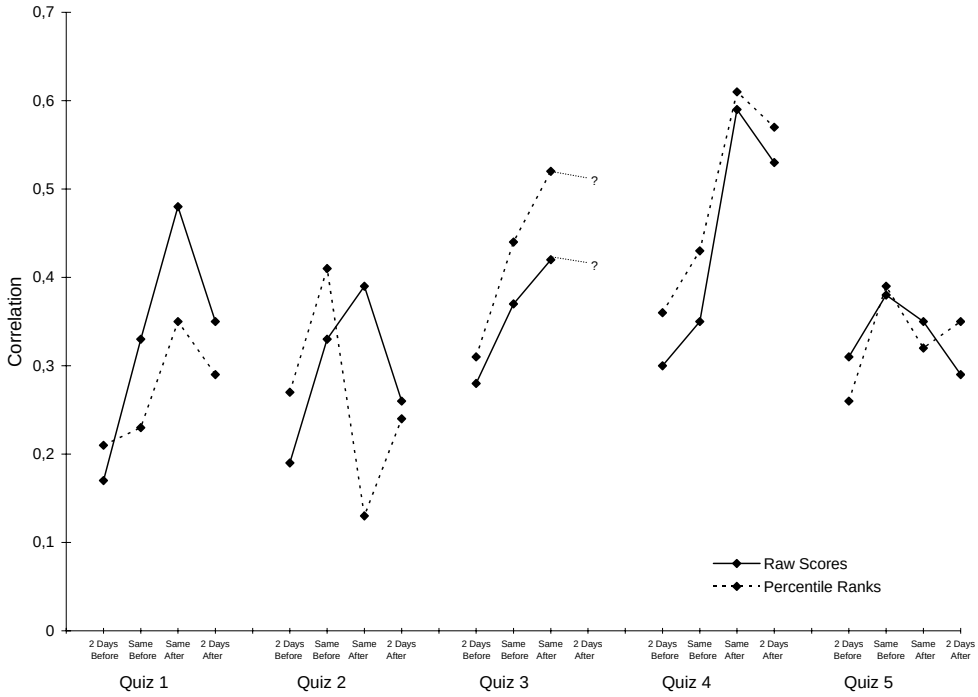


Figure 7. Inter-individual validity coefficients for solo evaluations and social comparisons as a function of Temporal Perspective and brief Time Horizons.

Question 3. Discrepancies between Solo and Social Performance Judgments

Changes in Bias. Here again, we conducted three sets of repeated measures ANOVAs on the discrepancy between solo and social performance judgments. One ANOVA examined the effects of Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) and Experience (Five Levels: Quizzes 1 through 5) to form a 3x5 factorial design. The ANOVA second examined the effects of a longer Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs. Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). The third AVOVA examined the role of Temporal Perspective (Pre vs. Post Task) a short Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5).

Figure 8 shows that discrepancies were lower when evaluations are made after the performance event than before. This was supported by a significant effect for Temporal Perspective, F (1,22)=9.12 p<. 01. Figure 8 also shows that the discrepancies between solo and social performance judgments decreased when evaluations were made at times closer to the performance event than those that were farther away from the performance event. This was supported by a significant effect for a short Time Horizon, F (1,22)=7.99 p <. 02. Finally, Figure 8 shows a significant interaction between Temporal Perspective and Time Horizon, F (1, 22)=6.18 p<. 05.

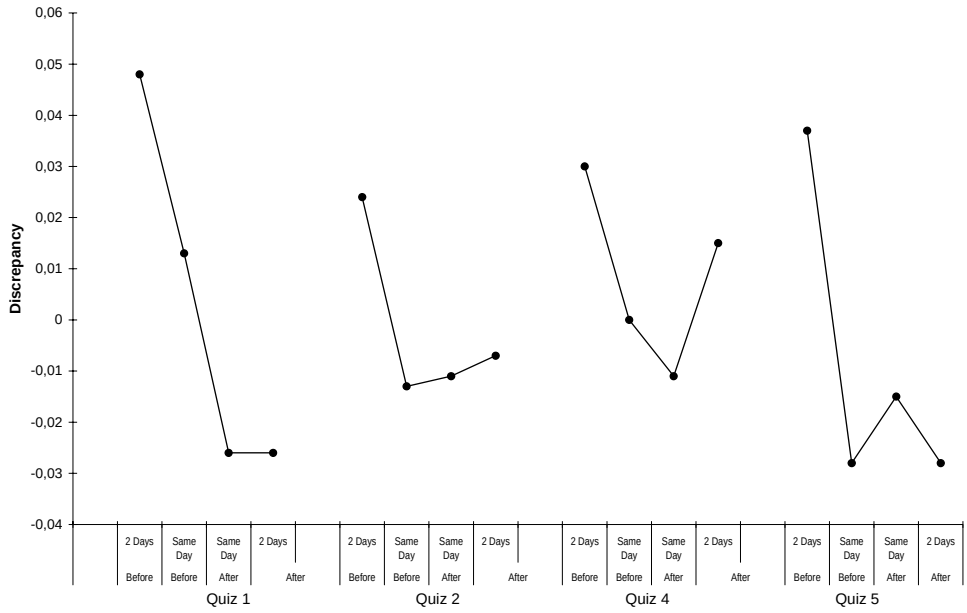


Figure 8. Discrepancies between solo evaluations and social performance judgments as a function of Temporal Perspective and brief Time Horizons.

That is, for the first two quizzes, pre-performance discrepancies between solo evaluations and social performance judgments were much higher than post-performance discrepancies between the two types of estimates. For Quiz 4 however, this pattern disappeared only to reappear for Quiz 5. We did not find similar results when we examined absolute differences between solo and social performance judgments. Only the Experience by Time Horizon interaction was replicated when the MAD between solo and social performance was the dependent variable, $F(3, 66)=3.17$ $p< .05$.

Figure 9 shows that there is a significant drop in the discrepancies between solo evaluations and social performance judgments from estimates made long before the performance event to those made right before the performance event. This was supported by a significant effect for short Time Horizon, $F(3, 111)=19.55$ $P<. 001$. Figure 9 also shows that with increasing experience, overall discrepancies between solo and social performance judgments are lower. This was supported by a significant effect for Experience, $F(2, 74) =3.69$, $p< .05$. Again, we did not find similar results when we examined the MAD between solo and social performance judgments. This can be explained by the mix of positive and negative discrepancy scores; the differences evident in signed scores cancel out with MAD.

Figure 10 shows that when evaluations were made at points that were temporally closer to the performance event, the bias between solo evaluations and social performance judgments reduces. Sometimes, solo evaluations were even lower than social performance judgments (e.g., right before quizzes 2, 4, and 5). This was supported by a significant effect for Time Horizon, $\underline{F}$ (2, 58)= 22.65 $\underline{P} < .001$. Again, we did not find similar results when we examined MAD between solo and social performance judgments.

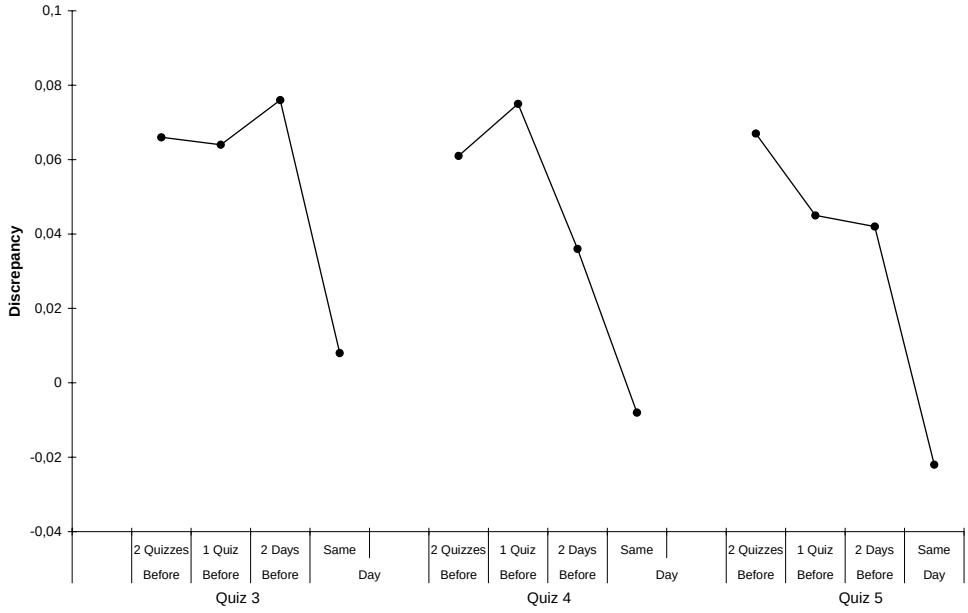


Figure 9. Discrepancies between prospective solo evaluations and social performance judgments over short Time Horizons.

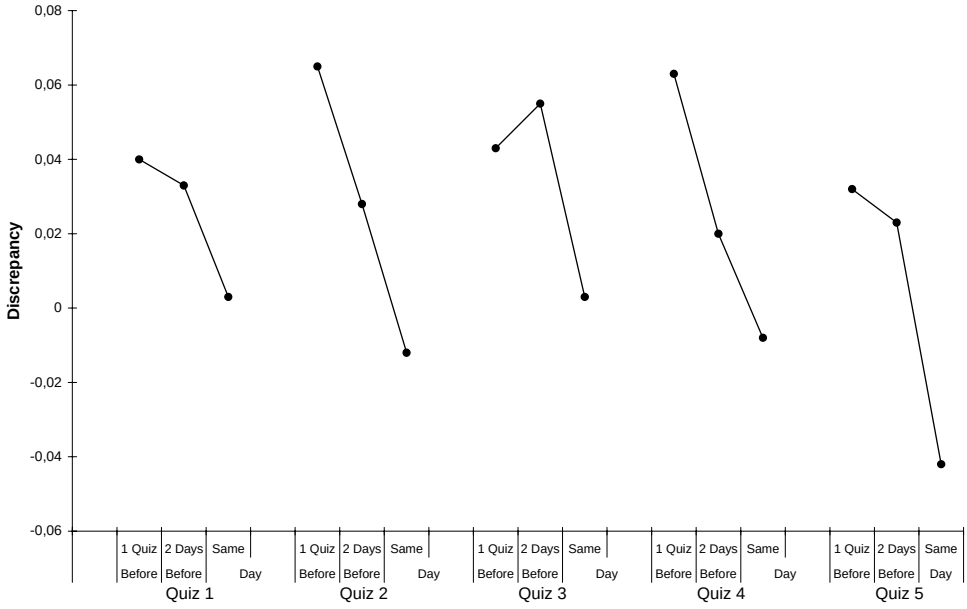


Figure 10. Discrepancies between prospective solo evaluations and social performance judgments over long Time Horizons.

DISCUSSION

Situations with many people performing essentially the same task are common in the military, education, business, government, and sports. People in these situations can be expected to care about, and frequently evaluate their performance. Their self-evaluations of performance are best examined as social behavior. Whether people perform concurrently or at different times, an implicit or explicit referent for judgments of one's own performance is the performance of others. Consequently, social performance judgments and social comparisons become part of the process of self evaluation of performance. It is of substantial interest to know how people in these domains evaluate their own performance and the performance of others, and compare the two.

This research endeavor uses a dynamic approach to describing three distinct judgments involved in self-evaluation of performance: social comparison, social judgment, and solo evaluation. It is worth pointing out several features of our effort that distinguish it from other work on self-evaluation and social comparison. First, the dimension of interest is performance, and not the ability and trait dimensions commonly found in social comparison research. The nature of the dimension under evaluation in social comparisons has been neglected by researchers (Wood, 1989); yet it may be an important distinction. Second, the self-evaluation made by the participant concerns the performance criterion variable. In our work, it is an estimate of actual performance level on the specified task and not a judgment of competence, ability, or satisfaction. Third, the comparison standard is a group and not an individual, and is offered by the environment rather than selected by the performer. It is a factual standard defined by the average of a group (i.e., a "social category reference point" Higgins, Strauman, and Klein, 1986), where that group provided the social context for performance and evaluation. Fourth, the judgments are evaluated with respect to three measures-bias, MAD, and validity. Whereas most social psychological research attends to questions of accuracy, applied research concentrates on the validity of self-assessments. We used both.

Of course, it is always risky to generalize across natural settings. But settings characterized by coactors, delayed performance feedback, and opportunities for social interaction are comparable to the setting of our study. Our results show what Festinger (1954) hinted and Wood (1989) documented - that people are not unbiased self-evaluators. People in the settings described above are likely to have distortions in their perceptions of their performance relative to that of others. What is more important is that our study shows that the direction and magnitude of the distortion changes with temporal perspective, time horizon, and experience.

Long before the task, most people estimate performance levels for themselves that are far superior to those of the referent group. This illusion fades as the task approaches, and is replaced by a far more pessimistic evaluation following performance. Once experience with the task has provided sufficient information about performance and the accuracy and validity of self-evaluations, realism replaces bias. One interesting Explanation for the pre and post-task change in evaluations is suggested by Feather (1969). In his study, Feather found that unexpected successes were rated as more satisfying than expected ones and although the opposite pattern was found for unexpected failures (i.e., that they were less satisfying than expected failures), the pattern was not significant. These findings suggest that after task

performance, people may lower estimates in order to avoid the negative affect accompanying unexpected failures and to approach positive affect that accompanies unexpected successes.

Sheppard, Ouellette and Fernandez (1996) examined changes in predictions as temporal duration of time to feedback decreased (4 months to .2 weeks before graduating in Study 1 and 1 month before quiz, 3 days, 50 min and 3 seconds before feedback in Study 2). Although they did not distinguish between pre and post-task performance changes in predictions they predicted and found that estimates decreased as time to feedback approached.

It is not typical for participants in social comparison studies to provide explicit judgments about others. This is disappointing given that examination of social judgments of performance helps diagnose the quality of self-evaluations of performance. Bias in social comparisons can be attributed to biased evaluations of others, biased evaluations about oneself, or both. By assessing social performance judgments in addition to solo evaluations and social comparisons we can identify whether inappropriate social comparisons are due to a flawed perception of others' performance or a flawed comparison process. The general answer is that all three judgments, solo evaluations, social performance judgments, and social comparisons, were positively biased. Whether people are more accurate about themselves or others may depend on when the judgments are made. Our data showed relatively less bias about one's own performance compared to that of others after the task. The reverse was true before the task. But with short time horizons there was comparable bias in self and other judgments of performance. Because the bias in direct social comparisons generally paralleled the bias in solo evaluations, there is no reason to suspect biased evaluation processes. That is, direct assessments about one's performance relative to that of others in the form of estimated percentile ranks neither minimized nor magnified the bias in solo evaluations.

We proceed to discuss implications of our results for people in natural settings with the specified characteristics, limitations of this study, and potentially fruitful directions for research on self evaluations of performance.

Implications for Behaviour

While this picture is interesting in itself, it is more valuable because of its implications for behaviour in these settings. People's evaluations of their performance relative to the each other may exert a broad influence on their behaviour. For example, perceptions of performance superiority enhance feelings of inequity (unless accompanied by perception of superior outcomes). Deleterious effects of inequity such as decreased motivation may be most severe long before the critical performance event. Ironically, because one forecasts superior performance one shifts to a strategy what will reduce performance, and ultimately performs far below the original forecast. If one is seeking to restore equity by conforming in terms of performance, the self-evaluations should decline monotonically up to the time of the event. Of course, the fact that this is the observed pattern in our data does not mean that perceptions of equity were responsible. Nevertheless, this pattern can provide useful insights into the judgments underlying perceptions of equity.

Other consequences of social comparisons for behavior are of considerable interest. For example, they might stimulate competition (Beck and Seta, 1980), produce assorted emotions (Higgins et al., 1986), encourage task persistence, or promote learning and performance. The important point is that for each consequence of social comparison of performance, there are likely

to be changes over the course of the time period leading up to and following the critical performance event. Only with a dynamic view such as we have presented can this source of variance be appreciated.

Limitations of the Study

Our work follows that of Radhakrishnan et al. (1996) by using manipulations of temporal perspective, time horizon, and experience. It has some of the same limitations, namely restriction to an environment with extrinsic rewards based on a mix of absolute and relative performance. Self-evaluation processes may differ in contexts such as athletic contests and markets where the competitive edge determines all. But the present study's design overcomes several deficiencies in their work because it includes multiple time horizons with the past temporal perspective, and symmetrical time horizons for both temporal perspectives. In addition, this study has more participants, more task trials, and measures than did theirs.

Nevertheless, our research design and execution was not ideal. Methodological problems include the failure to obtain all post-measures for quiz 3, and a lack of control over fluctuations in actual scores over the five quizzes. Although missing data were relatively few in number, there may have been a disproportionate exclusion of data from students at the lower end of the performance continuum. Finally, it is possible that the repeated requests for self-evaluation estimates may have induced unnatural levels of self-monitoring. The picture is considerably brighter when viewed from the perspective of educators: mere self-evaluation could have been beneficial to participants' motivation

(Bandura, 1986) and performance (Sanna and Pusecker, 1994). A between-subjects research design would minimize the problem of artificially increasing self-evaluation frequency or salience (as well as its likely benefits to the performers) by restricting self-evaluation to only one of the time periods of interest. Note that in the present study of two temporal perspectives, three time horizons, and five task repetitions, this approach would require 30 experimental conditions.

PROPOSED DIRECTIONS FOR RESEARCH

While this study has provided a detailed picture of self evaluations of performance over the course of time preceding and following performance, many questions remain. Attempts to understand the causes and consequences of social vs. solo evaluation seem especially worthwhile. Past research activity on the antecedents of social comparison has long focussed on studying the choice of a comparison standard (Wills, 1981). In light of the significance of temporal perspective and time horizon, it would be useful to ask how social comparison processes change over time. More specifically, it is desirable to know when social comparisons occur, and whether preferences for the comparison standard shift with time independently of other factors. According to the "similarity hypothesis" of Festinger's (1954) theory, people seek to compare themselves to others who are similar on the evaluation dimension. The time data from this study suggest the preferred comparison standard will change as one's self-evaluation of performance changes. Another question that can be

addressed with the dynamic approach exemplified in this study is whether social performance judgments precede social comparison or are formed only in service of the comparison process. All forms of evaluation may be automatic, but they may also vary in timing and influence each other differently.

A prominent theme in studies of the consequences of social comparison has been peoples' reactions to upward and downward comparisons (Pyszczynski, Greenberg, and LaPrelle, 1985). No doubt that data on this matter are extremely valuable. But many social comparisons are not between oneself and a single individual whom one has selected for comparison purposes. The comparison standard often pertains to an entire group, and is a given for all coactors (Wood, 1989). Such is the case with distributional information, such as mean performance for the group. It is not clear that the reactions people have to comparisons with an individual apply similarly to a group. It is one thing to aggrandize the individual outperformer (Alicke, LoSchiavo, Zerbst, ad Zhang, 1997), and quite another to imagine that an entire group consists of geniuses. In general, judgments about individual persons or events are fundamentally different from judgments about a set of persons (Klar and Giladi, 1997) or events (Sniezek and Buckley, 1991).

Recent research identifies a number of individual difference variables that may lead to differential use of social comparison as well as differential reactance to social comparison information. One variable that seems to hold special promise for explaining variance in self-evaluations of performance is gender. Beyer (1990) found that gender differences in self evaluations existed for certain kinds of tasks. These may be present with time and social comparisons, if as Cross and Madson (1997) argue, that women have more interdependent self orientations and therefore have more elaborate and available information about others. This in turn implies that social comparison information may be more accurate for women in general, and may be more sensitive to informational differences over time. One may also predict that women may consistently give lower estimates of the discrepancy between own vs. others' performance because they have more information about others than men do.

A personality variable that may divide self-evaluations over time into two distinct patterns is uncertainty orientation. Uncertainty-oriented people are motivated to acquire new knowledge about themselves while certainty-oriented people have the desire to avoid ambiguity and thus attempt to maintain existing beliefs. According to Roney and Sorrentino, (1995), self-assessment is more important to the former and self-verification is more important to the latter. The difference between the two personality types becomes most intriguing over time. Certainty-oriented persons are expected to seek information about their performance in the absence of feedback. But once it is obtained, they should be resistant to inconsistent information. In contrast, those who are uncertainty-oriented are predicted to seek information about their performance from multiple sources, and to do so as long as there is something new to discover. They cease self-assessment when information becomes redundant. Thus we speculate that for the certainty-oriented, solo evaluations and social comparisons will be highly correlated, with fewer changes with temporal perspective, time horizon, and experience. The predicted for uncertainty-oriented persons is less correspondence between social comparisons and solo evaluations, and steeper slopes for evaluations as a function of time horizon magnitude, and direction of temporal perspective direction.

A newcomer to the list of individual differences variables involved in self-evaluation is happiness. A study of the hedonic consequences of social comparison by Lyubomirsky and

Ross (1997) shows self-rated happy persons to be less sensitive to social comparison information than unhappy persons. They suggest that the relationship is bidirectional, meaning that happiness is a cause of selectivity in making social comparisons as well as a consequence of minimal attention to social comparison information when the comparisons are made. To establish the true nature of this link, it will be necessary to manipulate mood and affect, and to observe social comparison processes. It would be valuable if such research would include multiple temporal perspectives and time horizons, and assess the accuracy and validity of the social comparisons. It may well be that an even more intriguing pattern of differences between happy and unhappy people emerges. The present research shows how much can be learned by examining the quality of self-evaluations, and by tracking them over time.

Finally, Sheppard, Ouellette, and Fernandez (1996) found that self-esteem affected differences in how drastically people changed their estimates as the moment of feedback approached: low self esteem individuals made lower estimates as time for feedback approached. They suggest that people may lower estimates to avoid disappointment (rather than for other reasons such as to regulate, explain pre-feedback anxiety or to escape anxiety).

Although individual differences and personality variables hold promise for explaining variance in self evaluations of performance, it is unlikely that they can be understood adequately without reference to the effects of the temporal factors we have shown to have such consistent effects. The challenge will be to extend theory to understand how individual variables will or will not combine with time horizon and temporal perspective, and experience to alter the patterns observed in the study.

ENDNOTES

[1] We conducted similar analyses on measures of MAD. We only report the results for mean bias in detail to simplify our presentation and explanation of our results. However, when the results for mean bias differ from those for MAD which were few, we report these and provide possible explanations for such differences.

[2] For Quiz 1, the estimate termed "1 Quiz Before" was collected during base-line data collection, that is, before the set of all five Quizzes.

[3] This design was thusly constructed because we had a larger range of pre- feedback estimates for Quizzes 3, 4, and 5. Participants generated estimates on all five Quizzes at each time-point of evaluation. We only analyzed pre-feedback estimates because it was not possible to elicit an equal number of post-quiz and pre-feedback evaluations for these three quizzes.

[4] Due to experimenter error, "2 Days after Quiz 3" performance assessments were not obtained. Therefore, separate analyses were conducted using the four other quizzes (1,2,4,5) to test for Temporal Perspective effects.

[5] We conducted similar analyses on measures of MAD. We only report the results for mean bias in detail to simplify our presentation and explanation of our results. However, when the results for mean bias differ from those for MAD--which were few, we report these and provide possible explanations for such differences.

[6] For Quiz 1, the estimate termed "1 Quiz Before" was collected during base-line data collection, that is, before the set of all five Quizzes.

[7] Note that comparable long retrospective time horizons were not feasible in this study because they would have required delaying performance feedback to students for several weeks.

[8] This design was thusly constructed because we had a larger range of prefeedback estimates for Quizzes 3, 4, and 5. Participants generated estimates on all five Quizzes at each time-point of evaluation. We analyzed only pre-feedback estimates because it was not possible to elicit an equal number of post-quiz and pre-feedback evaluations for these three quizzes.

[9] Due to experimenter error, "2 Days after Quiz 3" performance assessments were not obtained. Therefore, separate analyses were conducted using the four other quizzes (1,2,4,5) to test for Temporal Perspective effects.

[10] We found that the patterns for MAD were similar to those depicted by measures of bias.

[11] These were the only 3 pre-feedback estimates we had for all 5 quizzes.

[12] We did not compute the validity coefficients for social performance judgments - - the criteria for these judgments do not change across persons.

[13] We did not find similar patterns for changes in MAD - they were non-significant across variations in Time Horizon, Temporal Perspective and Experience.

[14] Number of comparisons for raw score estimates and percentile rank estimates do not always match due to ties in one or more comparisons of the correlations.

REFERENCES

Alicke, M. D., LoSchiavo, F. M., Zerbst, J. and Zhang, S. (1997). *Journal of Personality and Social Psychology*, 73 (4), 781-789.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Beck, H. P., and Seta, J. J. (1980). The effects of frequency of feedback o a simple coaction task. *Journal of Personality and Social Psychology*, 38. 75-80.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, 59(5), 960-970.

Biernat, M., Manis, M., Kobrynowicz, D. (1997). Simultaneous assimilation and contrast effects in judgments of self and others. *Journal of Personality and Social Psychology*, 73(2), 254-269.

Brickman, P., and Berman, J. J. (1971). Effects of performance expectancy and outcome certainty on interest in social comparison. *Journal of Experimental Social Psychology*. 7. 600-609.

Carver; and Scheier, (1999). Self-Control Theory. In R.S.Wyer (Ed.) Themes and issues in the selfregulation of behavior: *Advances in social cognition*, vol. 12. Mahwah, NJ: Erlbaum.

Dweck, C. S., and Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Feather, N. T. (1969). Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology,* 13, 129-144.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.

Gilovich, T., Kerr, M. and Medvec, V. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, 64, 552-560.

Harackiewicz, J. M., Barron, K. E., Carter, S. M., Lehtc, A. T., and Elliot, A. J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, 73(6), 1284-1295.

Helgeson, V. S., and Mickelson, K. D. (1995). Motives for social comparison. Personality and Social Psvchologv Bulletin, Vol. 21, No. 11, Nova 1200-1209. Higgins, Strauman, and Klein, 1986. Standards and the process of selfevaluation: Multiple affects from multiple stages. In *Handbook of Motivation and Cognition: Foundations of Social Behavior,* ed. R.M. Sorrentino, E. T. Higgins, 1, 350378. New York: Guilford.

Jarvis, W. B. G., and Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology,* 70(1), 172-194.

Klar, Y. and Giladi, E.E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, 73(5), 885-901.

Lyubomirsky, S., and Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. *Journal of Personality and Social Psvchology*, 73(6), 1141-1157.

Markus, H. and Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey and E. Aronson (Eds.), *Handbook of Social Psychology* (Volt., pp. 137-230). Hillsdale, NJ: Erlbaum.

Nisan, M. (1972). Dimension of time in relation to choice behavior and achievement orientation. *Journal of Personality and Social Psychology*, 21. 175-182.

Prentice, D. A. (1990). Familiarity and differences in self- and otherrepresentations. *Journal of Personality and Social Psychology*, 59(3), 369-383.

Pyszczynski, T., Greenberg, J., and LaPrelle, J. (1985). Social comparison after success and failure: Biased search for information consistent with a self-serving conclusion. *Journal of Experimental Social Psychology*, 21, 195-211.

Radhakrishnan, P., Arrow, H., and Sniezek, J. A. (1996) Hoping, performing, learning, and predicting: Changes in the accuracy of self-evaluations of performance. *Human Performance,* 9(1), 23-49.

Roney, C. J. R., and Sorrentino, R. M. (1995) Self-evaluation motives and uncertainty orientation: Asking the "Who" question. *Personality and Social Psychology Bulletin*, Vol. 21 (12). 1319-1329.

Sanna, L., and Pusecker, P. A. (1994). Self-efficacy, valence of self-evaluation, and performance. *Personality and Social Psychology Bulletin*. Vol. 20 (1), 82-92. Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology,* 51, 918-931.

Scheier, M. F. and Carver, C. S. (1983). Self-directed attention and the comparison of self with standards. *Journal of Experimental Social Psychology*, 19, 205-222.

Sedikides, C. and Strube, M. J. (1997). Self-evaluation: To thine own self be good, to tine own self be sure, to thine own self be true, and to thine own self be better. *Advances in Experimental Social Psychology*, Vol.29, 209-269.

Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology*, 42(2), 281-291.

Shepperd, J. A., Ouellette, J. A.., and Fernandez, J. K. (1996). Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psycholoqy*. 70. 844-855.

Sniezek, J. A. and Buckley, T. (1991). Confidence depends on level of aggregation. *Journal of Behavioral Decision Making*. Vol. 4, 263-272.

Sniezek, J.A. and Reeves, A.P. (1986). Feature cues in probability learning: Database information and judgment. *Organizational Behavior and Human Decision Processes*, 37, 297-315.

Suls, J. M., and Tesch, F. (1978). Students' preferences for information about their test performance: A social comparison study, *Journal of Applied Social Psychology*, 8, 189-197.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.

Wood, J. V. (1989). Theory and research concerning social comparisons and personal attributes. *Psychological Bulletin*. 106, 231-248.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 65-86

Chapter 3

INSTABILITY AND RELATIVITY OF PREFERENCES: HOW CONTEXT DETERMINES UTILITIES AND DECISIONS

Ivo Vlaev*
Department of Psychology
University College London
26 Bedford Way, London, WC1H 0AP
United Kingdom

ABSTRACT

The aim of this paper is to draw attention to what is arguably a very general and pervasive feature of human cognition that may have important implications for our understanding of human decision making and also for some aspects of economics. The major claim, defended here, is that when people judge the attributes of choice options (like utilities, payoffs, and probabilities), they are not able to represent the absolute magnitudes of these attributes; instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. Also, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other sampled magnitudes. Such sampling of knowledge from memory and transferring it to the current situation produces certain biases in judgment because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there may be no ability to represent cardinal scales, for any magnitude and judgments involving such magnitudes are determined by the context. The core evidence for this claim comes from recent research in psychophysics on the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound, and also from research on the effects of context on decision making under risk and uncertainty.

* E- mail: i.vlaev@ucl.ac.uk; Tel: +44 207 679 7585

INTRODUCTION

Economic theory is often conveniently formulated to assume that economic actors have perfect information, a perfect grasp of their objectives, and the perfect ability to use that information to further their objectives. Markets comprising such agents, and games played by such agents, can be understood using particularly elegant techniques, and producing general and valuable theoretical results. It is common for social scientists outside economics to criticise the use of such strong assumptions as hopelessly indefensible. But such criticism is not, on its own, any more persuasive than any other criticism of scientific idealization. The analysis of the world routinely involves wild simplifying assumptions. Most prosaically, Newtonian celestial mechanics treats planets as point masses; the analysis of the path of a projectile may ignore friction. These idealizations seem rather harmless in comparison to those of the economist; they seem to involve ignoring factors that may seem, in some particular context, to be of rather limited significance; and in any case there is the hope that we may be able to add them in later, albeit at the cost of complicating our calculations.

In reality, though, science is replete with much more extreme simplifications---for example, the physics of spin-glasses is described in terms of the Ising model, which completely and knowingly falsifies the spatial structure of charges inside the spin-glass (Hertz, Krogh, and Palmer, 1991); models of avalanches in “sand-piles” use local rules for “collapse” which are known to be qualitatively different from the rules that govern actual sand-piles (Bak, 1997); “neural network” models of brain function knowingly and fundamentally falsify a range of key neuroscientific facts (Sejnowski, 1986; although see O’Reilly and Munakata, 2000). This explanatory phenomenon is so universal, even in physics, that it has led some to doubt the global coherence of physical science (Cartwright, 1983). The moral here, though, is that classical economists need not be defensive purely because they use wildly implausible assumptions; or that, indeed, this simplifying assumptions seem to be essential rather than a mere matter of convenience (because once they are abandoned the entire analytical apparatus becomes unworkable). Idealizations may be radical simplifications, and to a degree, falsifications of, reality---but if their resulting predictions provide elegant explanations and good predictions, then this is merely business as usual for scientific enquiry (Friedman, 1953).

The burden, then, rests with those who believe that classical economic theory’s idealizations are, in some crucial way, over-simple. The burden is, specifically, to show that by using a (presumably, slightly) more realistic model of economic agents, better economic explanations and predictions will be obtained. One theme in this special issue is a discussion of the ways in which economic idealizations may be over-simple, with an eye to providing the basis for a richer economic theory.

Note that economic and decision analysis can be made more complex along many different dimensions. One dimension is to allow that economic agents may have imperfect information; although analysis rapidly becomes difficult if we allow that each agent’s imperfect information may be idiosyncratic. Nonetheless, to understand the competitive behaviour of firms, the geographical spread of technology and innovation, or the role of “inside knowledge” in markets, a richer representation of the information available to the decision making process of each type of economic agent may be necessary (e.g., Tirole, 1988). Another dimension is to allow that economic agents may have imperfect cognitive

abilities to process the available information, and then to investigate whether these limitations are universal across all people, which would give us the possibility to derive general decision-making theories able to predict choice behaviour. This very important psychological dimension, which is the focus of numerous publications in psychology, economics, and all other social sciences, is "bounded rationality"---the fact that, even given certain information, people are not able to use it optimally (Simon, 1959, 1992). At a broad level, bounded rationality is a mathematical necessity---in general, the sophisticated calculations involving probability and decision theory that are invoked in economic analysis are known to be, in general, computationally intractable, and therefore, presumably, beyond the computational powers of the brain (Oaksford and Chater, 1998). But, in the light of my previous discussion, this general point is methodologically uninteresting. The key question is how far specific cognitive limitations, or systematic departures from the dictates of economically rational thought and behaviour, can be identified, and used to explain economic and social phenomena (see Camerer, 1998 for a recent review of applications in individual decision making). This general programme also motivates much of the sub-fields of behavioural decision theory (Slovic, 1977), experimental economics (Kagel and Roth, 1995) and behavioural finance (e.g., Sheffrin, 1999).

The aim of this article is to show that adding a new dimension (of complexity) related to certain fundamental aspects of human cognition may have important implications for some aspects of economics and decision sciences.[1] The cognitive claim, I argue for, is that people are not able to represent absolute magnitudes of stimuli of any kind (including a choice option's attributes like utilities, payoffs, and probabilities). Instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. This framework was first described in the *decision by sampling* theory proposed by Stewart, Chater, and Brown (2006), who also argued that, when people represent a magnitude, they can only do so on the basis of whether it is *larger* or *smaller* than other magnitudes sampled from memory or from the immediate context. Here, I argue that such sampling of knowledge produces certain biases in judgment, because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there is no ability to represent on any cardinal scale, the absolute value of a magnitude of any kind.

The core evidence for this claim comes from the study of the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound. Much traditional research in psychophysics has assumed the existence of some cardinal internal scale of intensities, onto which physical stimulation must somehow be mapped; and there has been consequent debate concerning the nature of this mapping (e.g., whether it is logarithmic, as argued by Fechner, 1966; or a power law, as argued by Stevens 1957). But more recent theory (reviewed and analysed in Laming, 1997) suggests a different point of view---that the very idea of an internal scale is incoherent. In the next section I digress briefly into psychophysics, before applying the resulting conclusion to an economic context in the following sections. Note, though, that the parallel between the two cases is relatively close. After all, just as perceptual theorists traditionally assumed that people had internal scales for the representation of loudness and brightness, so a traditional economic

[1] This is also a part of a more general program aiming to ground decision-making research more directly on the underlying cognitive mechanisms that produce choice behaviour (e.g., see Oaksford & Chater, 1998).

picture of an agent assumes that the agent must have internal scales for the representation of the utility of various outcomes; for representing the probability that they will occur; and so on. Without a scale for utility or probability, the model of the economic actor would look very different (and, as I will briefly consider later, perhaps different in the crucial way identified above---in providing the possible basis for a richer analysis of economic phenomena).

MOTIVATION FROM PSYCHOPHYSICS

So let me begin with a brief digression into psychophysics. A well-known and puzzling paradox of psychophysical magnitude perception is that people are rather good at *discriminating* the intensity of different magnitudes, but remarkably poor at categorizing them absolutely. That is, people are typically able to tell which is the louder of two sounds, or the brighter of two lights, with an accuracy that would lead to the naïve impression that they can tell the difference between of the order of a hundred different physical intensity levels. Yet when people are asked to explicitly associate intensity levels with category labels (e.g., 1 for the least intense stimulus, 2 for the next most intense, to, say, 7 for the most intense), they find this astonishingly difficult. Rather than having of the order of a hundred or so different cognitive "bins" into which items can be reliably categorized, performance tails off when about *five* different intensity levels must be categorized. Critically, it matters very little either what the absolute intensity level of the items to be categorized actually is; or, more shockingly, what the range of those items is. That is, so long as the items are readily discriminable from each other, to a first approximation, all that matters is the number of items that must be classified.

This is hard to reconcile with the idea that items are represented on an internal scale, according to which it would seem almost inevitable that performance would (a) be far higher than is observed; (b) would be degraded, if at all, in proportion to the crowdedness of items along the internal scale. The alternative viewpoint is that people do not construct any kind of internal scale---instead they are only able to make ordinal judgements, concerning which stimulus items are more intense than others (Laming, 1997). That is, I assume no more than the ability to make binary discriminations, rather presupposing the existence of internal psychological scales. The motivation for the restriction to discrimination comes from previous work on the direct judgment and the absolute identification of psychophysical magnitudes, such as luminance and sound pressure. In particular, Laming (1997) has shown that empirical data in line with Stevens' power law relating psychophysical variables and free numerical judgments can arise without assuming any representation of absolute information. Stewart, Brown, and Chater (2005) developed a theory of this psychological task (absolute magnitude identification), which embodies these assumptions and successfully predict the approximate limit of five items, as well as makes detailed predictions about the correlations across trials, and the nature of confusion errors.

Another example supporting this viewpoint is an elegant experiment conducted by Garner (1954), who asked participants to judge whether tones were more or less than half as loud as a 90 dB reference loudness. Participants' judgments were entirely determined by the range of tones played to them. Participants played tones in the range 55-65 dB had a half-loudness point, where their judgments were "more than half as loud" 50% of the time and

"less than half as loud" 50% of the time, of about 60 dB. Another group, who received tones in the range 65-75 dB had a half-loudness point of about 70 dB. A final group, who heard tones in the range 75-85 dB, had a half-loudness point of about 80 dB. Garner's experiment indicates, therefore, that people have no idea of the absolute intensity of the sound or what it means for one sound to be half as intense as another. Instead, it seems that people adjust their responses depending on the presented sound intensities from which they are asked to choose.

Other examples of similar context effects abound in psychophysics. Thus, empirical investigations in absolute identification (e.g., Garner, 1953; Holland and Lockhead, 1968; Lockhead, 1984; Luce, Nosofsky, Green, and Smith, 1982; Ward and Lockhead, 1970), magnitude estimation (e.g., Jesteadt, Luce, and Green, 1977), relative intensity judgment (Lockhead and King, 1983), and matching tasks (Stevens, 1975), have shown that perceptual judgments of stimuli varying along a single psychological continuum are strongly influenced by the preceding material. A robust finding is that current responses (judgments) tend to be contrasted (i.e. negatively correlated) with immediately preceding stimuli and assimilated (positively correlated) toward previous responses. Laming (1997) provides an extensive discussion of other similar findings and summarises many decades of psychophysical research, the results of which are consistent with the idea that participants are unable to make reliable decontextualised judgements of absolute sensory magnitudes. He claims that only relative judgements can be made – so whenever isolated stimuli are presented, and a judgement about the magnitude of the resulting sensation must be made, there is always a implicit comparison baseline of some kind. This might be stimulus presented on a previous trial, or may be some undifferentiated amalgam of remembered experience.

In summary, context effects, like those found by Garner (1954), are consistent with participants making perceptual judgments on the basis of relative magnitude information, rather than absolute magnitude information (see also Laming, 1984, 1997; Stewart, Brown, and Chater, 2002, 2005).

If the representation of utility is analogous to the representation of any other magnitude information, and in particular like the simple perceptual dimensions discussed so far, then the evidence presented above suggest that there is no fixed zero point on the utility scale and the experienced utility will shift depending on the context provided by the other experienced (consumption or risky) options. How worrying should this result be for economists?

> Theories of decision making under risk typically and historically start from a normative standpoint, which is an economic theory of how decisions should be made (standardly, expected utility theory, von Neumann and Morgenstern, 1947); then assess the degree to which people do make decisions as they should (e.g., Kahneman, Slovic and Tversky, 1982; Kahneman and Tversky, 2000); and finally attempt to modify the normative theory to bring its predictions into line with people's actual behavior (e.g., prospect theory [Kahneman and Tversky, 1979]; regret theory [Loomes and Sugden, 1982]; rank-dependent utility theory [Quiggin, 1982]). The result involves building a bridge from economics (the normative domain) to psychology (the descriptive domain). The core elements of the normative economic approach are maintained. Typically there is a numerical representation of value or utility; a kind of representation of probability, or some related notion; the value of an uncertain outcome is computed by multiplying its 'value' and 'probability' attributes (or something similar). Within these constraints, however, theorists have devised a range of elegant and important models that capture a great deal of empirical data.

The alternative approach that I presented in this article, however, takes a different stance. Rather than starting from a normative economic theory, and attempting to make modifications that render it descriptively acceptable, I start from assumptions about elementary cognitive processes, and attempt to construct an account that can address the economic problem of choice under uncertainty. That is, my attempt is to bridge from psychology to economics, rather than the other way round, with the ultimate aim of testing whether the resulting model predicts economic behaviour. So the context dependence of the utility scale should only be of relevance for economists as far as there is a model that explains how the context affects preferences and hence economic choices. Stewart, Chater, and Brown (2006) propose a theory based on such links between basic cognitive principles concerning the representation of magnitudes, and economic behaviour.

TRANSFER OF UTILITIES ACROSS SCALES AND CONTEXTS

Sampling from the past and from the current environment would make it almost impossible for people to create absolute representation of the utility of various choice options, which makes very difficult for people to make consistent choices matching the standards of the normative decision theories. However, these theories will be even less predictive if we realise that people are probably unable to build single utility scale along which to judge and compare options that are not very similar to each other (like for example various consumption goods).

There is recent evidence that if objects differ on more than one attribute then such a transfer could be problematic. In the psychology of perceptual judgment there is ongoing debate about the integrality and separability of psychological dimensions, arguing whether or not, for some pairs of psychological dimensions, judgment of the level of a stimulus on one dimension is interfered with by irrelevant variation on another dimension. There is evidence that for many pairs of dimensions, orthogonal variation on one dimension interferes with judgments of the level of a stimulus on the other dimension (e.g., Garner and Felfoldy, 1970; Ashby and Townsend, 1986; Lockhead, 1992). Such dimensions are said to be integral. It seems that, for integral stimuli, stimulus attributes are not represented independently of one another.

Stewart and Chater (2003) did loudness judgment experiment, in which on each trial, participants were presented simultaneously with a tone and a hiss and asked to judge which was louder. Their results showed that the assimilation of the current stimulus towards the previous stimulus was stronger when the current and previous stimuli were of the same type, and was attenuated when the current and previous stimuli were of different types, as was the accuracy of responding on the current trial. These data suggest that information about the loudness of a stimulus is not represented separately from the information about other stimulus attributes. If the information was represented separately then, for example, the effect of a tone or a hiss (of equal loudness) on the previous trial should have been the same. Stewart and Chater suggest that there may not be a single underlying scale representing loudness independently of other stimulus attributes. In particular, these data appear to suggest that people cannot consistently compare the loudness of two different types of sound: their binary

discrimination between categories of sound can be manipulated depending on the sounds they have heard on the previous trial.

Another interpretation is that the successful abstraction of a common scale depends on how similar are the two stimuli whose attributes are being judged. Note that common finding in research on sequential effects on judgment is usually an interaction between the previous stimulus and the previous response (two time-lagged variables). The assimilation towards the previous response seems to be modulated by the difference between the two consecutive (previous and current) stimuli (Jesteadt et al., 1977; Petzold, 1981). The closer the stimuli, the stronger the assimilation. Therefore, the conclusion is that if stimuli are very dissimilar then people contrast them so much so it becomes impossible to compare them realistically on a single scale. Probably the findings of Stewart and Chater can be explained in terms of the similarity between the stimuli, i.e., if the two stimuli, or their attributes, are very similar, then an ordinal loudness scale could be extracted and generalised independently of other stimulus attributes; while if the stimuli are very dissimilar (e.g., hiss and tone might be perceived as categorically quite dissimilar stimuli), then each type of stimulus is ordinally represented only on its own loudness scale. In support of this idea, Garner and Felfoldy (1970) showed that for certain dimension pairs that are rather dissimilar, like for example circle size and diameter angle, there was little or no facilitation or interference in direct stimulus sorting (scaling) task.

If these findings are transferred into the decision making domain and assume that judgments of value on any dimension are similar in nature and context dependent, as I argue here, then the conclusion is that the utility scale cannot be generalised over domains, situations, and product types, if products and domains are very dissimilar. For example, when people choose between various consumer goods it might be impossible for them to compare the utility from a holiday with the utility from CD player. Therefore consumer theory cannot be based on the standard indifference curves and needs some revision in light of the presented evidence about the locality of utility scales.

In summary, the findings presented here seem to be problematic for those accounts that suggest such sequential effects should be considered as a biasing of absolute judgment. Instead, I argue for a more radical alternative: that all there is, cognitively, is relative judgment. When there is a change in the stimulus type, removing at least part of the context used to make a relative judgment, then there will be a reduction in accuracy. The evidence presented in this section suggests that most of the representation of perceptual magnitudes is context dependent. In the next section, I move from psychophysics to behavioural decision making and present some relevant psychological evidence on the context dependence of judgments and choices under risk.

THE ROLE OF CONTEXT IN DECISION MAKING

A small number of experiments have investigated the effect of the context, i.e., the set of available options, on decision making under risk in a way analogous to the effects I have described in a psychophysical context, above. For example, the set of options available as potential certainty equivalents has been shown to affect the choice of certainty equivalent for risky prospects (gambles). In making a certainty equivalent judgment, participants suggest, or select from a set of options, the amount of money for certain that is worth the same to them as

a single chance to play the prospect. Birnbaum (1992) demonstrated that skewing the distribution of options offered as certainty equivalents for simple prospects, whilst holding the maximum and minimum constant, influenced the selection of a certainty equivalent. When the options were positively skewed (i.e., most values were small) prospects were under-valued compared to when the options were negatively skewed (i.e., most values were large).

Benartzi and Thaler (1998, 2001) have found evidence of another effect of the choice set by studying how people allocate their retirement funds across various investment vehicles. In particular, they find evidence for a diversification bias, which they call the *1/n* heuristic. The idea is that when an employee is offered *n* funds to choose from in her retirement plan, she divides the money approximately evenly among the funds offered. Use of this heuristic, or others only slightly more sophisticated, implies that the asset allocation an investor chooses will depend strongly on the array of funds offered in the retirement plan. Thus, in a plan that offered one stock fund and one bond fund, the average allocation would be 50% stocks, but if another stock fund were added, the allocation to stocks would jump to two thirds. Read and Loewenstein (1995) also reported that people tend to diversify equally between the set of available options.

Simonson and Tversky (1992) also reported strong context effects but their evidence was that there is a general preference for the central options in each choice set, which they explained with what they called the compromise effect. For example, when participants had to choose between $6 or famous brand pen, the introduction of a pen from a lesser known brand name increased the proportion of participants selecting the famous brand pen, and reduced the proportion selecting the $6. Plausible account for this type of data was the notion of trade-off contrast, where participants, who are assumed to have little knowledge about the trade-off between two properties, i.e., they do not have a clear idea what is the exact utility of each option, deduce what the average trade-off is from the current or earlier choice sets. These data may reflect a more general tendency to prefer central options when choosing amongst set of options (also called extremeness aversion), which might be due to the relativistic way people derive the utilities of the choice options. Similar trend was observed when people choose between products on a supermarket shelf as shown by Christenfeld (1995). These results back up my earlier suggestion that preferences between different types of good may not be stable.

Simonson and Tversky (1992) also provide several cases where preceding material significantly influences current judgments in decision making. For example, when choosing between pairs of computers that vary in price and amount of memory, the trade-off between the two attributes in the previous choice affects the current choice. This result shows that by varying the preceding products the preference can be reversed. Such an effect of the preceeding material is similar to the sequential context effects found in the psychophysical studies of perceptual judgment reported in the previous section.

In summary, the results reported above seem to indicate a variability of decision behaviour, which cannot be explained in terms of the existing normative economic theories. In the next section, I present a review of the research that contributed to the development of the idea about the relativistic nature of human judgments and decisions, and also present some recent models that could account for these data.

There have been quite a few recent publications on the role of cognition in individual decision making (Kahneman and Tversky, 2002, for a review) as well as interactive game

playing (e.g., Colman, 2003). These persuasive accounts describe psychological phenomena in economic behaviour by introducing various non-standard principles of judgment and reasoning processes (like for example, loss aversion, non-linear and weighting of probabilities, team reasoning, stackelberg reasoning, etc.). The goal has been to explain psychological phenomena in decision making that orthodox economic theory, and its conventional extensions, cannot explain. I argue that in order to better understand human preferences, in addition, a model is needed of how the economic agent perceives and mentally represents the decision problem initially before any judgment and reasoning, consequently, takes place.

As an illustration of such an account, here I offer some results from a general research program that aims to ground accounts of rationality in general, and decision theory in particular, on the underlying cognitive mechanisms that produce the seemingly paradoxical behaviour. Existing models of rational choice, like expected utility theory for individual decision making, and also game theory for interactive decision making, are typically based on the underlying assumption that only the attributes of the risky prospect or the game need be considered when reaching a decision. In other words, these theories assume that the utility of a risky prospect or strategy is determined by the utility of the outcomes of the prospect or game, and transforms of the probabilities of each outcome. Then the assumption is that the decisions are based on these utilities.

There is recent evidence, however, that the attributes of the previously or currently seen risky prospects and games influence the decisions in the current prospect and game, which suggests that prospects and games are not considered independently of the previously played ones (Stewart et al., 2003; Vlaev and Chater, 2003, 2006; Vlaev, Chater, and Stewart, in press). In particular, Stewart et al. (2003) have argued for the existence of what they call "prospect relativity": That the perceived value of a risky prospect (e.g., "p chance of x") is relative to other prospects with which it is presented. In particular, Stewart et al. studied peoples' perception of utilities in individual decision making tasks in gambling situations. The initial expectation based on the psychophysical studies described above, is that the option set (i.e., the context) will affect peoples' choices because there is no fixed internal scale according to which people make their judgements of the values of certain options. The results demonstrated a powerful context effect in judging the value of different risky prospects - the set of options offered as potential certainty equivalents for simple prospects was shown to have a large effect on the certainty equivalents selected. For example, when during judging the value a 50% chance of winning £200 people have options of 40, 50, 60, and 70 pounds, the most popular choice is 60 and then second choice is 50. When people have options of 90, 100, 110, 120 pounds, the most popular choice is 100, and then second choice is 110. So the set of alternatives affects valuation by a factor of nearly 2! This effect was replicated despite monetary incentives designed to encourage participants to deliver accurate and truthful certainty equivalents. In another experiment, the set from which a simple prospect was selected was also shown to have a large effect on the prospect that was chosen.

Vlaev, Chater, and Stewart (in press) further verified the prospect relativity principle, originally discovered with abstract gambles, by demonstrating relativity of human preferences in financial decision making under risk. This study investigated how the range and the rank of the options offered as saving amounts and levels of investment risk influence people's decisions about these variables. In the range manipulation, participants were presented with either a full range of choice options or a limited subset, while in the rank manipulation they

were presented with a (positively or negatively) skewed set of feasible options. The results showed that choices of saving rates and investment risk are affected by the position of each option in the range and the rank of presented options, which suggests that such judgments and choices are relative.

Similar context effects were also found in a sequential setting during interactive decision-making when people play many one-shot Prisoner's Dilemma games with appropriate anonymity (Vlaev and Chater, 2003; Vlaev and Chater, 2006), thus providing a new type of anomaly for orthodox game theory. In particular, I found that the degree to which people co-operate in these games is well predicted by a function of the pay-offs in the game, the cooperation index (Rapoport and Chammah, 1965).[2] In particular, the participants were asked on each round of the game to predict the likelihood that their co-player will cooperate, and then to make a decision as to whether to cooperate or defect. The results demonstrated that the average cooperation rate and the mean predicted cooperation of the co-player in each game strongly depended on the cooperativeness of the preceding games, and specifically on how far the current game was from the end-points of the range of values of the cooperation index in each session. In particular, the actual and predicted cooperation rate for a particular game was higher if this game was closer to the highest cooperation index value in the given sequence, compared to a condition in which the same game was closer to the lowest value in the sequence. Thus the perceived cooperativeness of a game did not depend only on the absolute value of its cooperation index, but also on the position of this index value in comparison with the minimum and the maximum index values in each experimental session (condition) as would be expected if the “cooperativeness” of a game could not be represented absolutely. Another results was that in games with identical cooperation indices, people cooperated more and expected more cooperation in a game with higher rank position (in terms of its cooperation index) relative to the other games in the sequence. These results present a challenge to game theoretic models that assume that the attributes of each game in a sequence are independently considered from the other games that are played.

In summary, these findings on the role of context in decision making, which are reviewed here, present another challenge to the standard rational choice theory and game theory. But they also challenge descriptive theories of decision-making under uncertainty, including rank dependent utility theory (Quiggin, 1982, 1993), configural weight models (Birnbaum, Patton, and Lott, 1999), and prospect and cumulative prospect theories (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), which all assign a risky prospect with a value or utility that depends only on the attributes of that prospect. In addition, the results by Vlaev and Chater (2006) present another challenge to the standard game theory, which assumes that games in a sequence are considered independently (e.g., Fudenberg and Tirole, 1991).

However, there are some theories, in which the utility or value of a prospect is not independent of the other prospects in the choice set, and hence they are potential candidates that could account for the prospect and game relativity phenomena described here. Thus these theories embody the relativistic idea (defended in this article) that choice alternatives are judged only relative to each other, instead of being based on some absolute internal cardinal

[2] The cooperation index was proposed by Rapoport and Chammah (1965) to give a measure of the probability with which human players tended to cooperate when playing the game. Specifically, the cooperation index is a simple function of the values in the payoff matrix in Prisoner’s Dilemma. Roughly, it depends on the degree of incentive the players have in playing “defect,” in terms of the size of the gain they may achieve, and the disincentive from an altruistic standpoint in terms of potentially harming the payoff of the other player.

scale (for value, utility, etc.), These theories are *regret theory* (Loomes and Sugden, 1982), the *stochastic difference model* (González-Vallejo, 2002), the *multi-alternative decision field theory* (Roe, Busemeyer, and Townsend, 2001), the *componential-context model* (Tversky and Simonson, 1993) and the *range frequency theory* (Parducci, 1965, 1974). I briefly present the basic ideas in each of them.

According to *regret theory* (Loomes and Sugden, 1982) when choosing between outcomes people expect and estimate possible feelings of regret they may have on experiencing each outcome of a prospect. Such anticipated feelings of regret modify the utility of an outcome that results from a particular choice with respect to the outcomes that would have resulted from other choices (which where not taken). Therefore in regret theory, the utility of a prospect is not independent of the other prospects in the choice set.

In the *stochastic difference model* (González-Vallejo, 2002) prospects are also judged relative to one another and the function comparing prospect attributes gives the difference between them as a proportion of the larger attribute (the theory also assumes that subjective prospect attributes are the real prospect attributes). This proportional difference strategy is a special case of the stochastic difference model. The proportions are summed over all attributes to give the overall preference for one prospect over another. Hence in this theory the utility of a prospect is also not independent of the other prospects in the set.

In the *multi-alternative decision field theory* (Roe, Busemeyer, and Townsend, 2001) attribute values are compared across options, and these differences are summed across all dimensions to produce the momentary "valences" for each option. Preferences are constructed for each option by integrating valences over time (the relative weight for each dimension is assumed also to vary in time). This process contrasts with the accumulation of absolute attribute values. Instead, valences represent the comparative affective evaluations. Thus, the choice between options is made in relative rather than absolute terms, as in the stochastic difference model.

Tversky and Simonson (1993) proposed the *componential-context model* as a model of context dependent preference devised to provide an account of trade-off contrast and extremeness aversion (Simonson and Tversky, 1992). According to the model, each attribute has a subjective value depending on its magnitude and the value of an option is a weighted sum of its attribute values. The background context is assumed to be the previous choice set, which modifies the weighting of each attribute (dimension) according to the trade-off between the attributes in that set. Thus, after the weighting of each attribute has been modified, the value of an option in the current set is then modified by the relative value of the option averaged over pair-wise comparisons with the other options in the choice set (i.e., the choice between options is made again in relative terms).

Range frequency theory (Parducci, 1965, 1974) models how people value or rate items that vary along a single psychological dimension. According to the theory, the subjective value given to an attribute is a function of its position within the overall range of attribute values, and its rank among the other attribute values (here I mean the attribute values considered, or available, when the judgment is made). Thus attributes are judged purely in relation to one another, which is what the prospect relativity phenomenon is demonstrating when the attributes of the risky prospects are compared; and Stewart et al. (2003) indeed claim that the range frequency theory can account for the effects of the choice set. In particular, their results showed that people choose consistently either the more risky or the less risky options in the choice set, i.e., prospects at a relatively less or more risky position

within the total range of risk, and low or high ranking prospect when prospects are ranked by risk. Birnbaum (1992) also found his data on context effects in decision under risk to be consistent with the theory.

In summary, in theories where prospects are judged in relation to one another, as in regret theory, the stochastic difference model, multi-alternative decision field theory, the componential-context model, and range frequency theory, the same effects of the choice set can, under some circumstances, be predicted. These relational theories all have in common the idea that preferences are constructed for a given choice set.

My account for the results presented in this section is summarised by the claim that people have poor notions of absolute cooperativeness, risk, and utility, and instead make their judgments and decisions in relative terms (analogously to the presented psychophysical and cognitive theories of perception and judgment of information about magnitudes representing intensities of stimulus attributes). Recall that the experiments by Stewart et al. (2003), Vlaev and Chater (2003; 2006), and Vlaev, Chater, and Stewart (in press), were all motivated by evidence from absolute identification and magnitude estimation paradigms, which demonstrates that participants typically have poor access to absolute magnitude information, and instead they rely upon comparisons with recent or concurrent stimuli, as evident from the strong effect of preceding material demonstrated in these paradigms. The key message here is that these experiments challenge the assumption that people have access to some stable internal cardinal scale representing the absolute magnitude/intensity of stimuli during perceiving and judging the differences between them.

Such account departs fundamentally from previous work in this field, by modelling the highly flexible and contextually variable way in which people represent magnitudes (like sums of money, probabilities, time intervals, cooperativeness, etc.), rather than assuming that these magnitudes can be represented on absolute internal (cardinal) psychological scales. My conjecture is that the results from the studies presented here suggest that people use the context in order to derive the utility of a risky prospect or a strategy. Thus, if absolute judgments are impossible and judging by how much one option is better/worse than another option changes depending by the other options that are available, then the only reliable judgment that can be made is that one option is *just* better/worse than the other (without being able to say by how much). Therefore, the best a decision maker can do is to rank order the available choice options on each dimension for judgment (i.e., constructing an ordinal scale). Note, however, that I do not even need to postulate the existence of a stable internal ordinal scale, because all the decision maker is doing is making ordinal binary comparisons between the choice alternatives available in the working memory (perceived from the environment or retrieved from memory), which was originally proposed by Stewart, Chater, and Brown (2006). Using this decision strategy, the decision maker can determine which option is the best one, for example, by simply counting how many times each option was better than another option (in the context). As a result, the option that was most often better off relative to the sample, is obviously the best option to select. Given people's inability to make absolute judgment, such simple judgment heuristic is probably the most efficient to use. And indeed, there is some recent evidence that similar fast and frugal mental heuristics can lead to near-optimal results in various cognitive domains (Gigerenzer and Goldstein, 1996; Gigerenzer, Todd, and the ABC Group, 1999).

In summary, any descriptive account of decision behaviour, should incorporate a model of agents' basic cognitive perceptual processes. In summary, I believe that the standard

decision theory needs to be supplemented by a more general "cognitive decision theory," which grounds decision-making in the underlying cognitive mechanisms that produce choice behaviour.

IMPLICATIONS FOR MODELS OF ECONOMIC BEHAVIOUR

One novel aspect of this approach is that, although its primary focus is decisions involving key economic variables, money, risk, and time, these variables are treated as undifferentiated 'attributes,' which must be traded off against each other (which was recently proposed by Stewart, Chater, and Brown, 2006). This is in stark contrast to a normative economic account, where the nature of the trade-offs between these dimensions is specific to the attributes involved. Thus, trade-offs between pay-off and probability should be governed by a multiplicative combination of (transformed) pay-off and (transformed) utility (von Neumann and Morgenstern, 1947; Starmer, 2000); and trading off the time delay after which a pay-off is received against the size of the pay-off is governed by an exponential, or perhaps hyperbolic, time-discounting function (Loewenstein and Prelec, 1992). However, although the different structure of these trade-off is normatively appropriate, it might not reflected in people's psychological processes. Thus, the principles governing such trade-offs might be the same, whatever attributes are being combined. This hypothesis should be tested in future research; but if it is correct, then the theory of decision under risk, and decisions concerning time, appear to be special cases of the multi-attribute decisions, where these attributes might just as well concern different qualities of an article of clothing (price, look, material), as key economic variables (Roe, Busemeyer, and Townsend, 2001). From a normative economic perspective, such an approach would be scandalous---because it blithely ignores vitally important normative criteria. But from a psychological perspective, this approach is not unreasonable, if the mental processes underlying decision making draw on common psychological mechanisms.

Stewart, Chater, and Brown (2005) argue that the assumption that people do not have internal scales for value rejects Bentham's (1789/1970) notion that utility is calibrated on an internal psychological scale, which is also in sharp contrast from the psychological theories derived from economics, which make a similar assumption. Note that economists have also shifted away from assuming the existence of internal utility scales. For example, the standard "revealed preference" interpretation of utility in economics (Samuelson, 1937) takes utilities to be revealed by observable choices without further specification about the psychological nature of these utilities. Recall also that Savage (1954) generalized this assumption to utilities and probabilities by showing that preferences over gambles could be used to "reveal" utility and probability information simultaneously. Thus, from the revealed preference perspective, the utility and probability scales are derived from choice preferences, rather than from assumptions about psychological scales.

The conceptual framework presented here has interesting similarities with respect to this traditional view in economics. In both, the reveal preference perspective and the decision by sampling approach proposed by Stewart, Chater, and Brown (2005), people are assumed to have access only to their own binary preferences (or more generally, to binary comparisons between perceptual magnitudes). Therefore, to the extent that people have broader grasp of their own, more global, values (probabilities, etc), this must be inferred from sampling their

own past choices and other memories revealing their preferences. So, for a given person to gain any "global" insight into how much pleasure is gained from consuming a specific product, this person has to sample from her memory some related, comparable events where she consumes that product. If the consumption episode in question is preferred to these events sampled from memory, this "reveals" to the person that this was a good experience; if it is preferred to some past episodes, but dispreferred to as many, this reveals to the same person that the experience was moderate, and so on. Thus, to the extent that people have any global grasp of their views concerning their perspective on how valuable or probable some event is, they must "reveal" this, by sampling from their own binary preferences, just as the economist attempts to reconstruct utility and probability values from the entire set of a person's' binary preferences.

Yet, the account presented here also has very different properties from those of economics. The psychological approach outlined here assumes that sampling from memory is extremely limited, and also stochastic, people's judgments concerning magnitudes will be strongly influenced by the particular items that they happen to sample (see Stewart, Chater, and Brown, 2006, for a precise specification of this sampling model and some simulation results). In this account, these may be drawn from long-term memory of recent events, but also from magnitudes that have been presented in the decision problem that the person faces. Hence, people's assessments of pay-offs, probabilities and intervals of time will vary capriciously, and may be highly malleable, rather than corresponding to a stable ordering, as in normative economic theory. Indeed, the effects of sequential and simultaneous context discussed above confirm this prediction.

It is evident from the discussion so far that crucial to the framework presented here is the process of sampling comparison magnitudes from long-term memory. A critical question will be how such magnitudes are distributed in memory. If, for example, there are many memories of small sums of money, and few of large sums, then a medium-sized sum will be likely to be judged as relatively large---because it is larger than most comparison items. Stewart, Chater, and Brown (2006) assume that the distribution of magnitudes in memory reflects the distribution of magnitudes in the natural environment. This work connects with a range of recent research which views cognitive processes as adaptive reflections of environmental structure (Anderson, 1990; Anderson and Schooler, 1991; Chater and Brown, 1999; Oaksford and Chater, 1998).

Thus, it seems that the cognitive system is naturally wired to be history and context dependent and to search for comparison points in the memory of recent events that relate to the current problem. I could even argue that such constant sampling from our memories, which are derived from our interactions with the social and natural environment, and the inbuilt tendency to rely on the information reflecting the statistical structure of the environment in order to make adaptive decisions, has vital consequences for the knowledge transfer in the economy and society at large. Thus, we sample from people around us by observing what they do, and they also observe our preferences, and then in the future we all sample from our memories that are reflection of this sampling. Such view suggests that preferences are unstable within the individual even at most basic level and are socially (environmentally) determined and transmitted, and my major claim is that the transmission of preferences can be viewed as a form of knowledge transfer. Note that the traditional view in economics is that knowledge is socially generated and transmitted while preferences are endogenous and stable. Of course, there are fashions and tastes but these appear to be marginal to the economic theory of rational choice. Is there some implication of this view for the functioning of markets and the geographical propagation of preferences and prices? Well, my answer is that because preferences are so socially malleable, they may sustain any fixed equilibrium, but perhaps would still tend to stabilize as markets are connected (e.g.,

globalization may not just help integrate markets from the point of view of allocation of resources and prices, but also in integrating preferences).

In summary, inconsistency in judgment and decision making can come from imperfect biased sampling, which happens because people sample in reality mostly from one side of the scale (e.g., only bad quality products of certain type) and they adapt to that level and might become unable to make adequate judgments of value and utility when they encounter the same product in a context where its value is out of the range of values experienced before. Therefore people might find it difficult to integrate or transfer absolute knowledge about utility between sequences of experiences and the transfer is biased by the past cases. I also conclude that because of the inseparability and locality of some judgment scales, people usually end up with at best separate ordinal utility scales for different classes of good. But if there is no a single utility scale along each different products are compared, then these products cannot be compared systematically in order to derive the indifference curves ubiquitous in consumer theory. Therefore consumer theory and marketing analysis have to take into account the particular scale that is characteristic for each consumer group and product type.

THE ROLE OF CONTEXT IN COGNITION

The approach and the evidence presented so far suggest that decision making is fundamentally context-dependent. Here I take a broader perspective by arguing that there is enough evidence that most of human cognition is context dependent and I believe that this is a consequence of the goal of the cognitive system to adapt flexibly to the dynamic environment. Thus, my account can be considered as an attempt to provide an adaptionist approach to decision making. This approach would require adaptive, efficient, robust, context-specific, domain-specific, species-specific behaviour. This approach is contrasted to the traditional rational approach that demands consistency, transitivity and content-independence and context-independence for the resulting decisions. Therefore, current and future research should investigate decision making, preferences, and utility functions, from an adaptively normative viewpoint. The aim is to explore the possible adaptive value of people's deviation from the prescriptions of the traditional rational choice theories by taking adaptiveness and context-dependence in natural decision environments rather than consistency in arbitrary problems as the ultimate standard for good judgement and decision making. Such framework promises also to propose an evolutionary-normative alternative to the standard rational choice theory. Here I also argue that context-dependence is nor restricted to decision making, but is a general feature of our cognitive system, which affects all cognitive processes, and which appears to serve adaptive purposes.

Advocates of such ecological views of rationality (Evans and Over, 1996, 1997; Gigerenzer and Goldstein, 1996; Gigerenzer and Todd, 1999; though see Chater, Oaksford, Nakisa, and Redington, 2003, for a critical analysis of some aspects of these views) emphasise the contrast between everyday human behavior, the success of which must be judged in the context of a specific and complex environment, and abstract classical principles of rationality, which appear to be justified a priori, but which may crucially ignore constraints imposed by the environment about which learning occurs. In short, the concern is that

classical principles of rationality, on their own, are inappropriate as standards of real-world reasoning. For example, in Anderson's rational analysis framework (Anderson, 1990; Oaksford and Chater, 1994), agent's goals and environmental constraints are all used to modify one's understanding of what is optimal behavior in a particular context, although rational principles still play a key role in determining what behaviour is optimal (given the agent's goals and the environmental structure). This emphasis on the environment does count against the decontextualized study of human inference, which ignore content and context.

Note that this emphasis on the environment is a reaction to the attempt of some researchers to devise empirical tests of descriptive rational theories that are independent of specific contexts, beliefs, or utilities; which has led to a focus on internal consistency of behavior in highly artificial conditions, rather than on how behavior meshes with the environment. For example, there are psychological models of inference where normative theories are interpreted as models of mental calculation, not merely behavioral description. The paradigm example of such models are "mental logic" theories in the psychology of reasoning, which regard the syntactic proof theory for logic as the basis of the algorithms that implement logical inference in the mind (e.g., Braine, 1978; Fodor and Pylyshyn, 1988; Rips, 1994). However, these algorithms appear to be intractable and therefore cannot apply to complexities of real-world contextualised inference (Chater and Oaksford, 1990; Cherniak, 1986; McDermott, 1987; Oaksford and Chater, 1991). On the other side, reasoning that may appear poor in an ecologically invalid laboratory context may be highly adaptive in the natural environment, as has been extensively argued (Gigerenzer, Hell and Blank, 1988; Gigerenzer and Hoffrage, 1995; Gigerenzer and Murray, 1987; Oaksford and Chater, 1991, 1993, 1998). Thus, it is important to stress essential role of the environmental context in which reasoning takes place in order to understand everyday human inference and rationality (Oaksford and Chater, 1995).

Note that people do not need to calculate their optimal behaviour functions in order to behave adaptively, because deriving the optimal behaviour function is frequently very complex. They simply have to use successful algorithms; they do not have to be able to make the calculations that would show that these algorithms are successful (Chater et al., 2003). This viewpoint is standard in rational explanations of human and animal behaviour across a broad range of disciplines. As I pointed out, economists do not assume that people actually make complex game-theoretic or macroeconomic calculations (Harsanyi and Selten, 1988); zoologists do not assume that animals calculate how to forage optimally (e.g., McFarland and Houston, 1981); and, in psychology, rational analyses of, for example, memory, do not assume that the cognitive system calculates the optimal forgetting function with respect to the costs of retrieval and storage (Anderson and Milson, 1989; Anderson and Schooler, 1991). Chater et al. (2003) claim that such behavior may be built in by evolution or be acquired via a long process of learning--but it need not require real time computation of the optimal solution.

According to this framework, one way to account for at least the prospect relativity effects described in this paper, is to assume that the internal scale used to represent the items in question (e.g., the scale representing the utility of a prospect, or the cooperativeness of a game) not fixed but is stretchable and adaptive to the environmental stimuli; as if people have fixed quantity of dimensional capacity, which can be stretched or contracted to accommodate the task demand most efficiently (although even the assumption of such a flexible scale may not be necessary as demonstrated by Stewart, Brown, and Chater, 2005). Therefore, the

resolution of the scale will depend upon the task demands. This flexibility has the advantage that sensitivity can be task dependent, and that adaptation can occur. As a consequence, people are unable to make reliable judgements of absolute magnitudes because they do not have direct access to information about absolute magnitudes.

There are at least two existing models of context effects in perceptual identification conforming to these principles. The first is the *adaptation level theory* (Helson, 1964), which states that the judgement of a particular event is proportional to its deviation from the mean value of all other events – the adaptation level – which itself is assigned a neutral value. This implies that the sum of the judgements of all experiences (e.g., pain and pleasure) will not depend on the shape of the distribution of events, because the sum of the deviations from the mean is zero. Note, though, that adaptation level theory does not explain how people's judgments of absolute magnitudes seem relatively sensitive to the *range* of the items involved, and would require some extension in order to deal with this. The second possible account is the *range frequency theory* proposed by Parducci (1968, 1974), which I have already discussed. Parducci found that the neutral point of the scale did not correspond to the mean of the contextual events (contrary to the adaptation level theory), but rather to a compromise between the midpoint (defined by the range) and median (depending on the skew of the distribution) of the distribution of contextual events. For example, satisfaction judgements depend on the skew of the distributions even when the means of the distributions are the same. Thus, the range principle reflects tendency to judge an event relative to the proportion of the range of stimuli lying below that event on the specified dimension of judgment, while the frequency principle reflects a tendency to judge an event relative to the proportion of contextual stimuli lying below that event on the specified dimension of judgment (which therefore depends on the rank order of the stimuli). In summary, the subjective value given to an attribute is a function of its position within the overall range of attributes, and its rank. Thus attributes are judged purely in relation to one another.

In summary, by claiming that judgments and decisions are based on the context rather than on some absolute judgment of value or utility, I also attempt to provide, a speculative general theoretical framework that will present the *context* as a explanatory concept which could account for the effects of different factors employed in the models of choice attempting to account for people's deviation from the predictions of the traditional rational choice models (such explanatory constructs for example are 'frames', 'categories', and 'discourse'). In the framework presented here, the major assumption is that every form of decision making happens in some cognitive context and this context plays the main causal role in people's behaviour.

CONCLUSION

In summary, the key difference between the approach I developed here and those approaches derived from normative economic accounts is that I do not assume that people internally represent values, probabilities, temporal durations, or indeed any other magnitudes. Instead, I assume only that people can sample items from memory, and can judge whether those items are associated with a higher or lower value (or probability, or duration) than the present item. That is, I assume no more than the ability to make binary discriminations, rather

presupposing the existence of internal psychological scales. The motivation for the restriction to discrimination, comes from previous work on judgment and absolute identification of psychophysical magnitudes, and also from recent work on context effects on judgments and decisions under risk and uncertainty (see Stewart, Chater, and Brown, 2006, for a model based on these assumptions). The assumption that people do not have internal scales for value, probability etc., constitutes a break from Bentham's (1789/1970) notion that utility is calibrated on an internal psychological scale; and thus a break too from the psychological theories derived from economics, that make a similar assumption. Finally, I argue that solving the heterogeneous problems arising from the standard rational theories of choice requires looking at how people represent the decision problems and the quantities that define them depending on the decision context.

Here one could raise the question whether this added level of detail is really going to matter to economists. After all, as I discussed at the beginning of this paper, although economists use radical simplifications, if their resulting predictions provide elegant explanations and good predictions, then this appears to be characteristic for every scientific enquiry. I argue that it may matter to incorporate the level of detail that I discuss here because by taking into account the factors governing perception and judgment of magnitudes, I can obtain better economic explanations and predictions.

For example, if the assumption is that judgments of value on any dimension are similar in nature and context dependent, as I argue here, then the conclusion is that the utility scale probably cannot be generalised over domains, situations, and product types, if products and domains are not very similar. The important implication of such a view for consumer choice theory is that it cannot be based on the standard indifference curves analysis, and needs some revision in light of the presented evidence about the locality and relativity of utility scales (for example, by taking into account the particular scale that is characteristic for each consumer group and product type). Note that context dependent models of magnitudes perception can be incorporated without substantially complicating the calculations (for example, the range frequency theory is formally much simpler than most of the standard normative and descriptive decision theories).

Another consequence of the view presented in this article is that preferences are unstable within the individual even at most basic level and are socially determined and transmitted, which also implies that they may sustain any fixed equilibrium. Therefore, the transmission of preferences can be viewed also as a form of knowledge transfer, and because markets are interconnected, preferences would still tend to stabilize at certain equilibrium points.

I believe that further research is required to integrate the findings presented in this paper with the standard economic theory in a way that might justify economists in replacing their current rather severe idealizations.

REFERENCES

Anderson, J. R. (1990). *The adaptive character of thought*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Anderson, J. R., and Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703-719.

Anderson, J. R., and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 1, 396-408.

Ashby, F. G., and Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 93, 154-179.

Bak, P. (1997). *How Nature Works: the science of self-organized criticality*. Oxford: Oxford University Press.

Benartzi, S., and Thaler, R. (1998). Illusory diversification and retirement savings, working paper, University of Chicago and UCLA.

Benartzi, S., and Thaler, R. (2001), Naive diversification strategies in defined contribution saving plans, *American Economic Review*, *91*, 79-98.

Bentham, J. (1970). *An introduction to the principles of morals and legislation* (J. H. Burns and H. L. A. Hart, Eds.). London: The Athlone Press. (Original work published in 1789),

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science, 3,* 310-314.

Braine, M. D. S. (1978). On the relation betwen the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1-21.

Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Chater, N., and Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition*, 69, 17-24.

Chater, N., and Oaksford, M. (1990). Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn. *Cognition*, 34, 93-107.

Chater, N., Oaksford, M., Nakisa, R., and Redington, M. (2003). Fast, frugal and rational: How rational norms explain behaviour. *Organisational Behavior and Human Decision Processes*, 90, 63-86.

Cherniak, C. (1986) *Minimal rationality*. Cambridge, MA: MIT Press.

Christenfeld, N. (1995). Choices from identical options. *Psychological Science*, 6, 50-55.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26,* 139-153.

Evans, J. St. B. T., and Over, D. E. (1996). *Rationality and reasoning*. Hove, Sussex: Psychology Press.

Evans, J. St. B. T., and Over, D. (1997). Rationality in reasoning: The problem of deductive competence. *Cahiers de Psychologie Cognitive*, 16, 1-35.

Fechner, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt, Rinehart, and Winston. (Original work published 1860).

Fodor, J. A., and Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.

Friedman, M. (1953). *Methodology of positive economics*. Chicago: Chicago University Press.

Fudenberg, D., and Tirole J. (1991). *Game Theory*. Cambridge, MA: MIT Press.

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, *46*, 373-380.

Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology*, *48*, 218-224.

Garner, W. R., and Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1*, 225-241.

Gigerenzer, G., and Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650-669.

Gigerenzer, G., Todd, P., and the ABC Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Gigerenzer, G., Hell, W., and Blank, H. (1988). Presentation and content: The use of base-rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 513-525.

Gigerenzer, G., and Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102,* 684-704.

Gigerenzer, G., and Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.

González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review, 109,* 137-154.

Harsanyi, J. C., and Selten, R. (1988). *A general theory of equilibrium selection in games. Cambridge*, MA: MIT Press.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper and Row.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley, Redwood City CA,

Holland, M. K., and Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception and Psychophysics, 3,* 409-414.

Holyoak K., and Koh K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15,* 332-340.

Holyoak, K., and Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Jesteadt, W., Luce, R. D., and Green, D. M. (1977). Sequential effects of the judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 92-104.

Kagel, J. H., and Roth, A. E. (1995). *The handbook of experimental economics*. Princeton University Press.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291.

Kahneman, D, and Tversky, A. (Eds.). (2000). *Choices, values and frames*. New York: Cambrige University Press and the Russell Sage Foundation.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Laming, D. R. J. (1984). The relativity of 'absolute' judgements. *British Journal of Mathematical and Statistical Psychology, 37,* 152-183.

Laming, D. R. J. (1997). *The measurement of sensation*. London: Oxford University Press.

Lockhead, G. R. (1984). Sequential predictors of choice in psychophysical tasks. In S. Kornblum and J. Requin (Eds.), *Preparatory states and processes* (pp. 27-47). Hillsdale, NJ: Erlbaum.

Lockhead, G. (1992). Psychophysical scaling: Judgments of attributes or objects? *Behavioral and Brain Sciences, 15,* 543-601.

Lockhead, G. R., and King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 461-473.

Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal, 92,* 805-824.

Loewenstein, G., and D. Prelec. (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. *Quarterly Journal of Economics, 107,* 573–597.

Luce, R. D., Nosofsky, R. M., Green, D. M., and Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception and Psychophysics, 32,* 397-408.

McDermott, D. (1987). A critique of pure reason. *Computational Intelligence, 3,* 151-160.

McFarland, D., and Houston, A. (1981). *Quantitative ethology: The state space approach*. London: Pitman.

O'Reilly, R. C., and Munakata, Y (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press.

Oaksford, M., and Chater, N. (1991). Against logicist cognitive science. *Mind and Language, 6,* 1-38.

Oaksford, M., and Chater, N. (1993). Reasoning theories and bounded rationality. In K. I. Manktelow, and D. E. Over (eds.), *Rationality,* (pp. 31-60). London: Routledge.

Oaksford, M., and Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608-631.

Oaksford, M., and Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning, 1,* 121-152.

Oaksford, M., and Chater, N. (1998). *Rationality in an uncertain world: Essays in the cognitive science of human reasoning*. Hove, Sussex: Psychology Press.

Parducci, A. (1965). Category judgment: A range-frequency theory. *Psychological Review, 72,* 407-418.

Parducci, A. (1974). Contextual effects: A range-frequency analysis. In L. Carterette and M. P. Friedman (Eds.), *Handbook of Perception,* Vol. II, 127-141. New York: Academic Press.

Petzold, P. (1981). Distance effects on sequential dependencies in categorical judgments. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1371-1385.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation, 3,* 323-343.

Quiggin, J. (1993). Generalized expected utility theory: The rank-dependent model. Boston: Kluwer Academic.

Rapoport, A., and Chammah, A. (1965). *Prisoner's dilemma: a study in conflict and cooperation*. Ann Arbor: University of Michigan Press.

Read, D., and Loewenstein, G. (1995). Diversification bias: explaining the discrepancy in variety seeking between combined and separated choices, *Journal of Experimental Psychology: Applied, 1,* 34-49.

Rips, L. J. (1994). *The psychology of proof.* Cambridge, MA: MIT Press.

Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370-392.

Samuelson. P. (1937). A note on the measurement of utility. *Review of Economic Studies, 4,* 155-161.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Sejnowski, T.J. (1986) Open questions about computation in cerebral cortex. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* Vol. 2: Psychological and Biological Models, ed. McClelland, J.L. and Rumelhart, D.E., eds. MIT Press.

Shefrin, H. (1999). *Beyond Greed and Fear: Understanding Bahvioral Finance and the Psychology of Investing*. Harvard Business School Press.

Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *American Economic Review, 49,* 253-283.

Simon, H. A. (1992). *Economics, bounded rationality and the cognitive revolution.* Aldershot, England: Elgar.

Simonson, I., and Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281-295.

Slovic, P. (1977). Behavioural decision theory. *Annual Review of Psychology, 28,* 1-39.

Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature, 38,* 332-382.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153-181.

Stevens, S. S. (1975). *Psychophysics*. New York: Wiley.

Stewart, N., and Chater, N. (2003). No unified scales for perceptual magnitudes: Evidence from loudness. In: *Proceedings of the 25rd Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.

Stewart, N., Brown, G. D. A., and Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 3-11.

Stewart, N., Brown, G. D. A., and Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review, 112,* 881-911.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1-26.

Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General, 132,* 23-46.

Tirole, J., (1988). *The theory of industrial organization*, Cambridge, Mass.: MIT Press.

Tversky, A., and Simonson, I. (1993). Context-dependent preferences. *Management Science, 39,* 1179-1189.

Vlaev, I., and Chater, N (2006). Game relativity: How context influences strategic decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 131-149.

Vlaev, I., and Chater, N (2003). Effects of sequential context on judgments and decisions in prisoner's dilemma game. In: *Proceedings of the 25rd Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.

Vlaev, I., Chater, N., and Stewart, N. (in press). Financial prospect relativity: Context effects in financial decision-making under risk. *Journal of Behavioral Decision Making.*

von Neumann, M., and Morgenstern, O. (1947). Theory of games and economic behavior. (2nd ed.). Princeton, NJ: Princeton University Press.

Ward, L. M., and Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology, 84,* 27-34.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 87-110

Chapter 4

PRESS MEDIA INFLUENCE ON PERCEPTIONS OF GENETICALLY MODIFIED FOOD IN THE UK AND SPAIN

Marta Vilella-Vila[a] and Joan Costa-Font[b]*

[a]nVision Europe, Future Foundation Group,
London, UK
[b]London School of Economics and Political Science,
UK and Universitat de Barcelona, Spain

ABSTRACT

The vast majority of the public learns about new risks to health and society predominantly from the media, including the press media directly or indirectly. However, little is known about the role and mechanisms through which the press media influences attitudes and risk perceptions. Some approaches stress the ides that risks are partly created while other state that the media plays a neutral role, however empirical evidence is hard to retrieve and still is scarce. This paper empirically examines both the role of press media coverage and reporting of new genetically modified (GM) foods between 1999-2004. We draw upon a combination of qualitative and quantitative evidence. First, evidence of content analysis of key press media in two countries - Spain and the United Kingdom (UK) - is examined to illustrate preliminary evidence and subsequently, quantitative evidence of survey data (Eurobarometer surveys) is examined to scrutinise for the existence of some media biases, inter-country differences in public perceptions as well as specific media effects connected to role of journalism in the country. Results point towards the existence of significant differences in media reporting and respect for journalism between the two countries, which correlate with public perceptions, although a similar lack of trust was identified. Furthermore, we find evidence suggesting some specific media biases depending on the press media readership.

* Contact address: Joan Costa-Font, London School of Economics, LSE Health and Social Care, Houghton Street, WC2A 2AE, Cowdray House (London), Fax: +44 (0)207 955 6803, E-mail: J.Costa-Font@lse.ac.uk.

Keywords: *media content analysis, Genetically Modified Food, risk perceptions, trust and media bias.*

1. INTRODUCTION

In most western societies the benefits and risks of specific applications of biotechnology have been increasingly communicated to the public. However, whilst in some countries biotechnology has captured significant pro-active media coverage in others the media reports on the issue have been modest. Given the limited information in hand of consumers, communication on biotechnology is an interesting case study. It is important whether public perceptions - supposedly underpinning individual's behaviour – are effected by the number of stories published in a certain press outlet. Similarly, an interesting questions refers to whether the way certain stories are portrayed (for example, whether reporting systematically stressed the risky nature of biotech) has any influence on the readers of certain press media?. Given that press media outlets act as opinion leading sources, and that there sis always some degree of representation of reality (Hornig, 1994, 1992; Siegrist and Cvetkivich, 2001; Bauer et al, 2001; Gaskell et al., 1999; Marks et al, 2002), the mere fact that a story is written down in a newspaper could potentially affect readers perceptions as well as and those interacting with them (Siegrist and Cvetkivich, 2001). Though little is known about the process whereby the media influences public perceptions in the area of biotechnology, the link between public 'acceptance of biotechnology' and the extent of media coverage (Gutteling et al., 2002; Gaskell et al., 1998; Bauer et al., 2001; Marks et al., 2002) is an important contemporary question to address empirically.

In claiming that the media exerts a specific influence in shaping public perceptions, it is unclear whether it simply drives public perceptions (for example, by influencing people's emotions and stressing certain values) or whether merely reflects an information demand to respond to an 'ongoing debate'. For latter to be the case, the public should trust the media as a key information source on biotechnology. On the other hand, it may well be that press media readers, regardless of their understanding of the issue and media reporting, confer greater value to information reporting risks as opposed to others reporting potential benefits, following what could be expected from prospect theory (Viscusi, 1998). However, limited evidence has been provided on the association between the media and risk and benefit perceptions at the news outlet level. Finally, when reporting about GM food, it has been pointed out that journalists and communication professionals are liable to dramatise the reality (Bauer et al, 2001), and consequently are liable to bias readers' perceptions.

Media biases are typically classified into 'reporting biases' and 'selection biases'. The first refers to the "slant" of the story, that is the content of the information reported, which according to Mullainthan and Shleifer (2002) can take the form of an 'ideological bias' resulting from a news outlet's desire to influence readers opinions (for example, due to the traditional political affiliation of a newspaper) and/or a 'spin bias' resulting from the attempt of journalists to create a 'memorable story', which results from the current mood and the country specific competitive environment in media reporting. Selection biases' could result from the journalistic tendency to select certain stories (for example, those reporting high and salient risks) in setting the press agenda. Finally, another potential selection bias is that

resulting from the placement of a story within the newspaper, which could arguably affect the intensity of individuals' perceptions.

To explore these issues further, we believe that an exploratory analysis of the press coverage and agenda-setting rules should be undertaken at the media outlet level looking at the press media coverage and its content intensity in communicating risks regarding the genetic modification (GM) of food. Because our aim is to provide in depth analysis, we have selected two European Union countries, the UK and Spain, where differences in media reporting and coverage were expected. In particular, we look at qualitative and quantitative evidence from two of the mainstream press media outlets in each country, whether there are significant differences in public perceptions between the two countries — and specific press media outlets — and at potential country heterogeneity in media trust and respect and trust in journalism.

We argue that media biases take place in the communication of GM food risks and benefits in Spain and the UK. With our data we attempt to scrutinise a potential 'broad association' between the number of press media messages (coverage) and its qualitative content and public perceptions of GM food (RQ1). Secondly, we examine again evidence of potential media biases in both countries (RQ2). Finally, we explore the possible determinants of media specific influences in attitudes, and in particular journalists respect public awareness and understanding of the issue under discussion (RQ3). This evidence contributes to the debate on the issue of socially formed preferences of risks.

This study is structured as follows. The following section deals with risk perception and the media. Section 3 explains the methodology of the study. Section 4 reports the results of a content analysis of UK and Spanish printed media on GM food. The reason for examining these two countries is rooted in the fact that while the media has reported intensively on the risks and benefits of GM food in the UK, Spain stands as a contrasting example within the European Union, as the media reporting there has been significantly limited. Moreover, Spain is particularly interesting because it is one of the few EU member states producing GM food and an interpretation of Spain's attitudes to biotechnology is generally absent from previous European studies (Gaskell et al., 1998) . After identifying the key patterns in each country's press media, section 5 develops an empirical analysis of attitudes and perceptions to tease out possible trends in acceptance and risk perception, among other issues. Possible links between coverage in the press media and public perceptions of GM food can be identified, which might offer some broad clues on the role that the media plays in the communication of issues around GM food. The final section is devoted to a discussion of the results attained and their policy implications.

2. PERCEPTION AND REPRESENTATION OF GM FOOD IN THE MEDIA

Although recent research has evolved towards the 'active audience' tradition of mass communications — which widely accepts that individuals 'actively' construct their own interpretations of reality — when it comes to communication of scientific ideas, the general lack of knowledge and experience of this area challenges such an interpretation. Indeed, the so-called 'pluralist theory of the media' (Blumler, 1977; Harrop, 1987; and Becker and McCombs, 1975), suggests that the media, by reflecting the balance of forces within society, aims to influence public perception, especially through the agenda-setting process — thus

determining 'what is/what is not news', as well as prioritising information reporting. Similarly, Innis (1991) argues that each mass medium is controlled by an elite who has some monopoly power over the information disseminated in its outlet. Both journalists and press editors are argued to adjust the story frame to their ideology, professional and knowledge limitations, as well as time and space constraints, which is argued to determine the potential socio-political and ethical implications of a story (Hornig, 1992). To this extent, it is natural to hypothesise that some specific 'media biases' could be in place, and establish relationships between the coverage and content of newspapers and television and citizens' perceptions regarding GM food.

Inappropriate prioritisation of a certain feature in the press news may potentially lead to an exaggeration (or amplification) of the social risks in people's perceptions. The notion that frequently prevails is that the media exaggerate some risks and ignore others (Slovic, 1987) or, alternatively, sacrifice objectivity for sensationalism (Johnson and Covello, 1987). Similarly, Palfreman (2001) notes that the media effects on public perceptions might be explained by pressures on journalists to tell a story 'as simply and dramatically as possible', as well as the fact that good news and bad news do not affect people to the same degree. Indeed, Siegrist and Cvetkivich's (2001) research on media content analysis indicates that public information about potential risks appears to be more trusted, possibly as Nisbet and Lewenstein (2001) suggest due to the framing of the news. Furthermore, vocalising and thereby legitimating some points of view while ignoring others plays a key role in structuring public debates around risk (Hornig, 1993).

Yet, the social amplification framework indicates that socially amplifying risks (Kasperson et al., 1988) entails more than simply publishing stories, but rather it refers to the content and the language used in a particular story. Modern techniques enable through content analysis an examination of this feature. Indeed, in the area of new technologies, Petts et al. (2001) suggests that the media can only amplify or attenuate risk if they capture or resonate with an existing public mood. Thus, numerous studies over a long time period, of both newspapers and television, fail to identify any strong link between media consumption and public perceptions of risk (Gunter and Wober, 1983). Other studies suggest that the media, although potentially influencing our risk perceptions, is only one factor among many and that people own experience is indicated as a much stronger factor (Wahlberg and Sjoberg, 2000). In the economics literature, Swartz and Strand (1981) found that the media reporting on a keepone contamination exerted a negative and stable impact on the demand for oysters in certain US markets. Verbeke and Ward (2001) found that media coverage of the Bovine Spongiphorm Ecephalopaty (BSE) crisis led to a two per cent reduction in beef consumption expenditure in Belgium. However, the effects of the media tend to be temporary and of a limited size. Indeed, Kalaitzandonakes et al. (2004) found that whilst in a case base of substantial and continuous media coverage there were no media effects, in another case of acute and brief media coverage the media effects were substantial.

In setting where indicial handle limited individual information (for example, genetically modified (GM) food), it has been shown that individuals in recalling information on new technologies look for 'simplifying summaries' from trusted sources (Fischhoff and Fischhoff, 2001). As stated by the European Commission in 2003, 'for the greater majority of Europeans, biotechnology is not a part of everyday personal experience'. As a result, lay public understanding on conflicting issues draws on 'second hand' or mediated information rather than personal experience. However, individual's information and perceptions depends

on the 'information frame' (Kanheman and Tversky, 1984), which is arguably different across countries and across newspaper outlets.

In comparing European perceptions of biotechnology with those of the USA, Gaskell et al. (1999) explain European resistance on the basis of trust in regulatory bodies — which is higher in the USA than Europe — and knowledge determining public perceptions of biotechnology. Hornig (2001) finds that media coverage in the USA before 1996 exhibited mainly positive frames. Studies that specifically deal with the media's role in influencing public perception across European Countries (Bauer et al., 2001; Gutteling et al., 2002), although inconclusive, identify an association between media coverage and public perceptions, although this association was not found for press framing. Furthermore, they found that between 1973 and 1996, in European countries in which 'scientists' enjoyed more coverage biotechnology was framed in a more positive way than in those countries where 'politicians' captured the coverage of biotechnology.

Marks et al (2002) find that between 1990 and 1999, both for the USA and the UK the focus of the media has been risks rather than benefits, and the newspaper coverage is found to be as 'negative' in the UK as in the USA. More general studies dealing partially with GM food in the UK identify specific media (Sheehy et al., 2002), as well as a lack of individual understanding and reliance on moral and ethical issues as being significant (Pett et al., 2001). Finally, Frewer et al. (2002), on examining the media effects regarding genetically modified food in the UK, found that although media reporting occasionally affected risk perceptions, it did not influence trust in regulators. However the media alone are not found to determine the risk amplification process.

Finally, one might well argue that the specific media impact might depend on the specific structure of the press media in a specific society. Compared to other European communication systems there is a low circulation of the daily press in Spain. Although only some 100 daily newspapers are sold per 1,000 inhabitants, 25 per cent of the country's citizens read the daily press and the two leaders in sales are El Pais and El Mundo. Despite the proliferation of print and broadcast media, and their diverse political stances, concerns have been raised about political influence in the media, and particularly in public broadcasting. Compared to the Spanish media, the British media are freer and able to report on all aspects of British life. The British press media are divided into tabloids (for example, The Sun) and broadsheets (for example, The Guardian).[1] The variety of publications on sale reflects the full spectrum of political opinion, as well as the British public's voracious appetite for newspapers. For a previous analysis of the British press for GM food using content analysis see Frewer et al. (2002) and Sheehy et al. (2002).

3. METHODS

This study examines the media coverage of GM food in a rather ample way by combining two techniques. We first undertake a *content analysis*, and secondly an empirical *examination of attitudes*, information and public perceptions to scrutinise whether the media have an impact on the way that individuals perceive GM food Yet, in examining press coverage over

[1] Accordingly, whilst in Spain, differences among newspapers might reflect differences in outlet specific ideology; in the UK there might be addition differences depending on tabloid and broadsheet readership.

time, there it is possible that other than press coverage are in place. On the other hand, if one examines differences across many jurisdictions that vary in press coverage, one can be challenged by the question of whether differences in public perception are the result or the cause of press coverage differences. This study aims to combine both time and country dimensions depending on the information available. Accordingly, we have selected two European countries in which there were substantial differences in the number of messages and in the way that the media tackled the issue of GM food, namely the United Kingdom (UK) and Spain. The latter is a GM food producer and only recently has there been any debate there on the controversies around GM food, whereas the former is paradigmatic as regards social debate on GM food. Therefore, a close look at these two cases might shed some light on the role that the media plays in informing society about technology and innovation.

3.1. The Press Content Analysis

The first empirical aspect that we focus on in our study is the frequency of media coverage and the framing used in selected print media, using classical press content analysis. We take some key publications as baseline indicators of possible trends in coverage, we then examine two of the most popular (or elite) publications — which are likely to be opinion leading sources for other newspapers — in each country from January to June 1999, and from April to June 2003. In Spain these were *El Pais* — a daily newspaper with a circulation of 435,300 readers and E*l Mundo* — a daily newspaper with a circulation of 300,300 readers. For the UK we examine *The Guardian* — a daily newspaper, with a circulation of 380,000 readers and *The Sun* — a daily newspaper with a circulation 3,500,000 readers.[2] Both newspapers in both countries represent a fairly similar contrast.

Although one could well point out as a methodological limitation the fact that these newspapers are not read by the whole population, an examination of the press media is important in that these exert an influence as an opinion-leading source for other media. Indeed, although a large number of people do not regularly read a newspaper, the press media does influence the agenda-setting and the information portrayed by other media sources, as well as key political stakeholders such as politicians and policy-makers, experts and academics, the industry and key lobbies. Due to the technical complexity of the GM food issue, the newspapers selected in this study for each country might be of great interest because they inform politicians and other journalists and, over time, reflect the tone of the national debate. In fact, some other prior content analysis of the UK press media on several areas including GM food does not include the entire number of newspapers (Sheehy et al., 2002). On the other hand, because the newspapers selected are influenced by 'different ideologies', the press content analysis might provide some evidence of the 'ideology effect' in driving press content.

[2] El Pais is chosen because of its status as an elite national newspaper, of left-wing orientation, which arguably informs opinion leaders as well as policy-makers. El Mundo acts as El Pais' counterpart from a right-wing position, employing a more populist and sensationalist style. It acts as the 'voice of the government', currently dominated by the right-wing Partido Popular, in government since 1996. The Guardian, although independent, can be regarded as of left-wing orientation and close to Labour Party elites. As a well-positioned broadsheet, it performs as an opinion leader that establishes the frame for some debates and has a major influence on policy-makers and other journalists. The Sun, owned by Rupert Murdoch's News Corp. Empire, the clearest icon of the popular British tabloids, reaches daily the outstanding number of 3.5 million readers.

The two time periods under scrutiny serve to provide some evidence from the field of the media treatment of the issue, and in particular allow us to examine whether potential 'spin effects' took place between 1999 and 2003. The reasons for focusing on this period are twofold. On the one hand, 1999 stands out as an important date due to the explosion of coverage and social debate (Bauer et al., 2001) and the debate once again has taken centre stage in Europe since 2003, with the end of the moratorium and the commercial battles at the core of the European Union and the World Trade Organisation. Furthermore, other earlier periods have been already examined for several European countries (Bauer et al., 2001), although no specific comparative analysis has been carried out for the two countries under examination and, in particular, the information on the Spanish media is limited.

Finally, the search terms employed were wide as all articles were revised on a piece-by-basis, as is normal practice in classical content analysis. Other studies — which included GM food as one area of analysis — have employed lexical content analysis (Sheehy et al., 2002) and qualitative analysis of the media using focus group discussions (Petts et al., 2001). However, given the widely acknowledged drawbacks of computer based methods (for example, their lack of qualitative depth) as well as the specific scope — in terms of time, topic, number of countries and media outlets examined — of the study, we have employed classical media content analysis, although it is arguably more time intensive and less sophisticated in producing indicators of media impact. Other alternative methods include the use of focus groups, which has been employed intensively in media impact analysis. However, some prior studies suggest that individuals' limited knowledge serves to constrain the production of qualitative evidence, especially due to the lack of real debate that the issue has generated in some countries (for example, Spain). The accessibility of the press in both countries was similar at the time of the study given that all the newspapers have digital online editions. A population of articles has been constructed using the Internet archives of each of the newspapers and searching for key words related to GM foods. The search resulted in a final selection of 570 articles across the four publications for the two periods of study. Hard copies and microfilms of the newspapers were also reviewed in order to check the location of the news and the front pages devoted to the topic. This evidence was used to examine the salience of the GM food debate, the number of articles published, the position of the stories in the newspaper (for example, whether they made the front pages) and how the GM issue was portrayed (for example, as a local or international issue).

3.2. Empirical Analysis of Survey Data

Empirical analysis of public perceptions has been based on a number of different sources, including European as well as country-based surveys. Other alternative methods to country-based surveys include extensive research using a manageable sample of individuals, such as that carried out by Frewer et al (2002). However, because our study has comparative purposes a representative survey of the populations of the two countries was deemed to be more appropriate. First we used the Eurobarometer surveys, which contain several questions on general attitudes towards and perceptions of biotechnology products in the 15 European

Union states.[3] Normally, the survey is carried out with a representative sample of 16,500 respondents, approximately 1,000 in each EU member state. This survey is the best instrument available to examine individual cross-country attitudes, risk and benefit perceptions and trust in information channels, as well as acceptance of and public knowledge on the issue.[4] Yet, because most of research employs the originals surveys we especially compare data from 1999 and 2002, which is the period in which the salience of GM debate was greater in all the European Union countries. Ideally, we should have measured public perceptions in time periods preceding upsurges in press coverage (in at least one country) and then again in periods after such surges. Unfortunately, the latest data available on public attitudes in of 2002, which although not ideal is envisaged as already advancing the patterns of media coverage observed in 2003. Indeed, the events that stimulate the most negative press coverage such as the moratorium on GM foods took place before 2002. Thus, thus visual timeline in the UK and Spain that could conceivably have stimulated press coverage and when press coverage and public perceptions were measured remained unaffected.

4. PRESS CONTENT ANALYSIS FOR THE UK AND SPAIN

4.1. Description of the Media Reporting on GM Food in UK and Spain

Along with parallel food scandals, GM food captured the attention of the press as the result of a research project commissioned by the Scottish Office from the Aberdeen-based Rowett Research Institute. Dr Arpad Pusztai was in charge of the project, which had the objective of investigating the effect of GM crops on animal nutrition and the environment. This included, for the first time, feeding GM potatoes to rats to see if it had any harmful effects on their metabolism and health. This feature alone arguably signalled some mistrust towards industry information and the American FDA (Food and Drug Administration) as regards their assessment of the safety of GM food. However, the story caught the media's attention when in January 1998 Dr Pusztai appeared on BBC2's *Newsnight* and voiced his scientific concerns about the weakening of the rats' immune systems. In April 1998, Granada TV's *World in Action* programme approached Dr Pusztai. He told viewers that he would not eat GM food. He insisted that he found it 'very, very unfair to use our fellow citizens as guinea pigs. We have to find (them) in the laboratory'. Two days later, Dr Pusztai was summarily suspended and subsequently forced to retire by the Rowett Institute's director. The outburst made headlines around the country.

The debate moved up a gear when the first GM food products became commercially available in the UK. Because of the ensuing controversy, in August 1998, a bill was presented in the Commons seeking a moratorium on GM food sales, and in October 1998 the

[3] The Eurobarometer 58.0 is one of the series of Eurobarometer surveys on biotechnology and the life sciences. The advantage of this survey is that can be compared to other standardised surveys conducted in 1991 (Eurobarometer 35.1); 1993 (Eurobarometer 39.1); 1996 (Eurobarometer 46.1); and 1999 (Eurobarometer 52.1), and occasionally, we employ data from Eurobarometer 55.2 (2001). All of these surveys are available online from the European Commission gateways.

[4] We draw upon the report on the survey: 'Europeans and Biotechnology in 2002', A report to the EC Directorate General for Research from the project 'Life Sciences in European Society', by G. Gaskell, N. Allum and S. Stares, Methodology Institute, London School of Economics, London WC2A 2AE, UK.

government announced a one-year moratorium and set up a cabinet committee on biotechnology and GM foods. In the first three months of 1999 the public debate with regard to GM foods was a key issue in Britain. On 14 February 1999, the bio-safety convention to establish international regulations governing GM organisms began in Cartagena, Colombia. Again, GM foods came dramatically to the public's attention with the publication in 1999 of a statement signed by 126 influential food writers and journalists condemning the use of GM foods. This initiative was promoted by the environmental organisation, Greenpeace. Furthermore, a set of articles were published in the UK newspapers stating the opinions of His Royal Highness, the Prince of Wales, who came out strongly against GM food. Several EU member states also voiced concerns about the safety of GM foods and, for five years, imposed a de facto moratorium on GM imports. In October 2001, German Agriculture/Consumer Affairs Minister Renate Künast and Environment Minister Jürgen Trittin wrote to the European Commission stating that the moratorium should remain in place until the revised GMO deliberate release directive came into force in October 2002 and future traceability/labelling regulations were clarified. In May 2003, President Bush launched a legal challenge to the EU at the World Trade Organisation, in an effort to force Europe to accept imports of US GM crops. In July 2003 the European Parliament passed laws intended to end a European Union-wide ban on new genetically modified foods.

In the UK, the government launched a public debate in June 2003. The aim was to listen to the public's views before deciding whether to license GM crops in the country. This was arguably the country's first nationwide public discussion around GM issues. Local authorities and network groups organised meetings to weigh up the pros and cons. The findings, along with the views submitted via an Internet site, were then fed back to the government in order to inform their policy-making on GM products. Therefore, the situation is now rather different to that of the late-1990s, when the story of GM food began to be written. Almost all the food in the shops is non-GM and it could be argued that it is those who would wish to buy GM food that are being denied a choice. The new (2003) EU labelling regulations are now even stricter than before and consumers will finally have a choice about whether or not they intend to consume products containing modified crops. As noted, in July 2003, the European Parliament approved a new set of regulations that impose additional labelling requirements on those wishing to sell any foods derived from GM sources. Public debate on genetically modified foods has become highly polarised. Some groups have set out to establish a less partisan perspective by involving a range of people with different viewpoints and working towards consensus. In the meantime, the amount of media coverage has been extensive, especially in the UK. Indeed, for 1999 we found 301 articles in *The Guardian*, in contrast to the 35 found in *The Sun*. Similarly, we identified 90 articles in *El Pais*, while only 40 in *El Mundo*. Yet, the picture is significantly different for 2003, for which period we identifed 63 articles in *The Guardian* and only one in *The Sun*, while in Spain we found 15 articles in *El Pais* and four in *El Mundo*. On the basis of this evidence, there would seem to be a pattern indicating that an ideological bias might be present in determining the coverage of GM food, whereby left-wing newspapers report more on the issue than papers of other affiliations. On the other hand, there is evidence of the spin effect; that is to say that reporting increased with the mood of the issue.

4.2. Content Analysis Results

4.2.1. The UK Press

GM food has been associated in the UK with a great deal of media attention, particularly in early 1999. Commercially publicised pressures generated public concern, which in turn forced the major UK supermarkets to announce a withdrawal of products containing GM ingredients from their own brand ranges. In addition, numerous regulations for the labelling of GM foods were introduced in the UK. They extended the range of products for which labelling must be applied, including food in restaurants and from 'fast food' outlets.

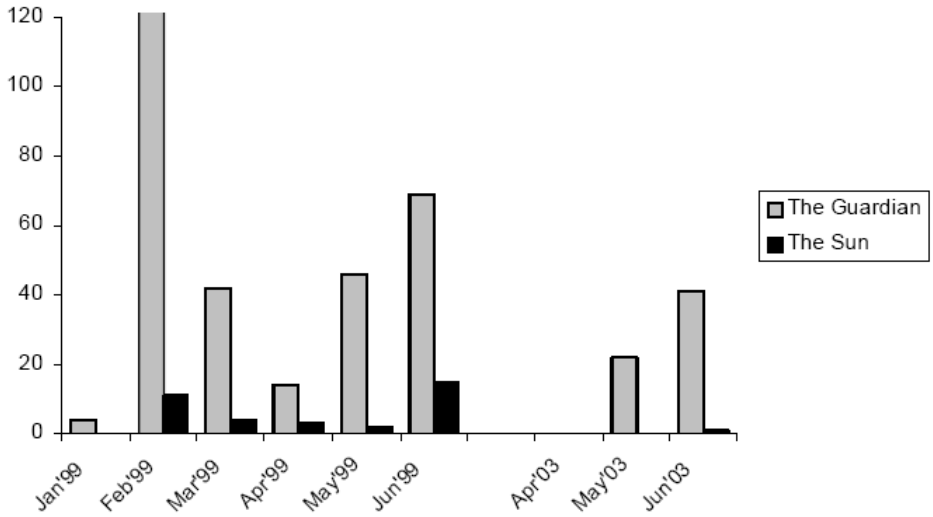


Figure 1. Number of News Stories in the Elite British Press (1999/2003).

The Guardian dedicates extensive coverage and debate to the issue of GM food. Indeed, the GM issue appeared as a *front-page topic 13 times* from January to June 1999. Seven of these 'top stories' happened to be in February, when the subject reached the level of 'food scandal' as Dr Pusztai was forced to retire. *The Sun*'s most extensive and complete articles came out mainly during February and June 1999. Nevertheless, in contrast with *The Guardian*, the GM food debate never merited front-page treatment there; instead it was normally kept for the second page. Interestingly, this evidence suggests that a potential 'selection bias' may be in place that might translate into a difference in public perceptions. In contrast, during 2003, although GM food was still a prominent and unresolved debate, it was not granted sufficient importance to emerge as a front-page issue either in *The Guardian* or *The Sun*, although *The Guardian* still offered more extensive reporting on GM food. As we show in Figure 1, the newspaper has devoted ample space in its analysis section to discussing the issue of risks and benefits. Indeed, the former refers to political and environmental correspondents' reports and public conferences inviting major experts to discuss the topic. Finally, it is noteworthy that the analysis and commentary articles in *The Guardian* amount to as much as 15 per cent of the total coverage, while in *The Sun* the maximum extension devoted to GM food, and only on key occasions, is three quarters of a page, lacking any deep analysis. However, this feature suggests that some public debates run parallel to the press

media reporting. Thus, the UK press is likely to suffer from 'spin biases', given that certain 'memorable features' are likely to affect the outlet readership. On the other hand, the potential 'ideological biases' are modest, insofar as *The Sun* has not undertaken a campaign against Labour over the topic of GM food in the same way it has with other 'hot debates ' such as the Euro.

The sensitivity of the UK public over the issue of GM food is linked to prior concerns such as the BSE and dioxin scandals, together with other food scandals that have taken place in Europe (Frewer et al., 2002). Food security and consumer trust have been dented and thus have become essential themes. The confidence of consumers and their right to choose have raised the news about GM food to a central concern. However, there is a significant variability in the wording and alarmist content of the news published by broadsheet and tabloid newspapers, according to the case studies selected. While the former present a complex picture of the issue, outlining diverse opinions and both positive and negative viewpoints, the latter focuses its attention on news stories that stress the need for strict regulation of or even a ban on the products, the potential health risks and shocking pronouncements from academics, scientists and popular figures.

The press reporting tends to reinforce a negative perspective of GM food, by stressing the risks and potential damage, without paying much attention to the possible benefits that such technology might bring (Gaskell et al., 2003), although from 1999 onwards the tone becomes more positive, as no new 'big controversies' have arisen. Other issues that were identified were c*onsumer trust and rights*, the extremes of the *political debate* and *communication controversies* amongst *s*cientists, environmentalists, politicians and the industry. *Positions* are neither similar nor homogeneous. Therefore, one might expect the public to develop ambivalent attitudes if they confer equal importance to all the information reported by the media. However, even though some specific articles exhibit straightforward opposition to GM food, normally the ongoing discussion does not show a two-sided antagonism between those 'in favour' of GM food and those 'against' it. In contrast to *The Guardian*, commentaries in *The Sun* are almost non-existent. Accordingly, assuming that the media exerts an influence, one would expect the reader of the tabloid to be less informed or less media-influenced in the GM food debate.

4.2.2. The Spanish Press

GM food security was one of the main health topics to emerge in the press in 1999, though not the one to be given the most extensive coverage. Nevertheless, questions on the safety of GM food 'exploded' in Spain in February 1999. From January to June of that year, the five main daily newspapers devoted a total of 65 features, 17 editorials and two front pages to GM food. The number of articles in *El Pais* and *El Mundo* evolved irregularly during 1999, as shown in Figure 2. Interestingly, none of the stories in these two newspapers featured on the front page. They were normally located within the 'Society' section, or, specifically in *El Mundo*, in the weekly Health Magazine. The coverage of the stories tended to be much shorter than the articles in *The Guardian*. Comments and analysis amounted no more than two per cent of total press coverage. This feature is consistent with the novelty of the issue and the fact that in Spain attitudes towards GM food were likely to be less politically influenced. Possibly, the end of the EU moratorium might have been the most important piece of information there regarding GM food in the last decade.

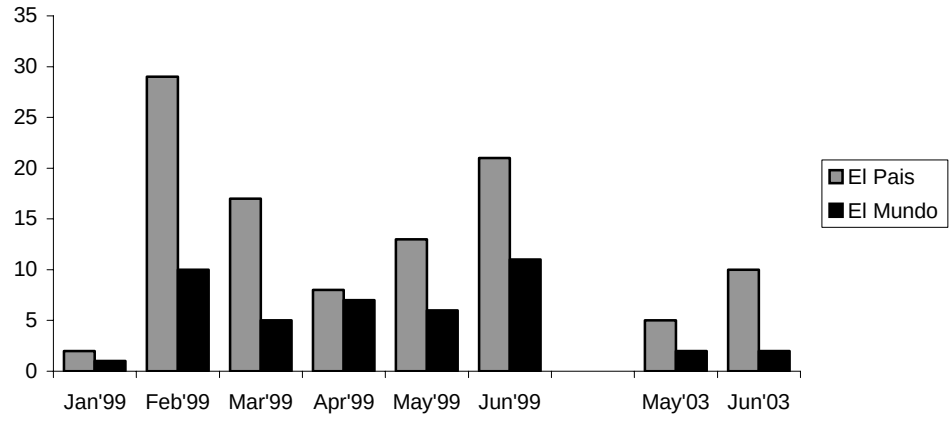


Figure 2. Number of News Stories in the Elite Spanish Press (1999/2003).

The Spanish press gives neither a 'pedagogic' nor a complete picture of this important public debate. There are no statements of opinion, nor on the position of the government, nor is the hostility of NGOs mentioned. Thus, there are no signs of a serious public debate within Spanish society on the issue of GM food. Accordingly, it is unlikely that the media reporting there results from the public demand for information. The 'Comments and Analysis' or 'Letters to the Editor' sections are almost devoid of any contribution, reinforcing the idea that the information on GM food is mostly straightforward information, lacking profound analysis.

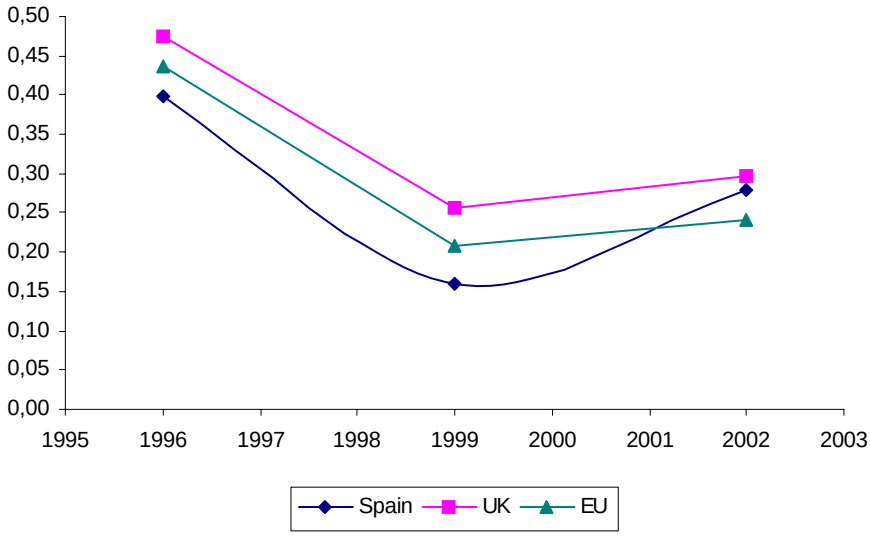
Unlike in Britain, GM food is not a 'hot topic' in the Spanish media. The debate has tended to be portrayed as one affecting the international arena rather than that country. The main coverage of the topic refers to stories from the UK. Yet Spain is one of the European countries most affected by GM crop cultivation, though the media does not appear to reflect this reality. Risk and uncertainty seem to affect those in the UK more than the Spanish. Stories on the debate in Spain and discussion of legislation represented less than 40 per cent of total coverage. The fears of environmentalists or consumer organisations did not generate great coverage.[5] Press reporting has tended to reinforce a negative view of GM food. Of the total amount of stories covered from January to June 1999, 68 per cent involved negative considerations, with only 14 per cent reporting the potential benefits; the remainder were considered neutral. However, comments, analysis and debates are few and far between in the Spanish press. Only *El Pais* offers some comments on the debate, in particular two editorial pieces have been published, suggesting potential 'selection biases'. On the other hand, *El Mundo's* limited reporting is always focused on questions of security, informing the readers about the opposing studies that have appeared on the topic.

[5] Some examples illustrate these assertions: 'Blair forms two commissions for GM food control (…) Government expects to recover consumer confidence' (El Pais, 22 May 1999); 'Spanish scientists guarantee GM food. Experts think this is a commercial war. The Lancet and the British Medical Journal published articles last week that GM foods do not have different effects to the traditionally produced food' (El Pais, 17 March 1999).

5. Is Public Perception of GM Food and Biotechnology Associated with the Media Coverage?

5.1. Correlation between Public Perceptions and the Press Media

After examining the press media content and coverage the next important question is whether there is a particular link with public perceptions. Evidence on this feature might provide suggestive evidence on the role of the media in influencing public perceptions. A first empirical feature is revealed in Figure 3, which indicates that optimism with regards to biotechnology declined steadily, reaching its lowest level in 1999 in line with media coverage patterns (Gaskel et al., 2003). Indeed, it is in this period (from February to June 1999) that the press media experience relative increases in coverage in Europe, especially during 1999, which evolved concurrently with greater public concern over the production and consumption of GM food. Interestingly, the same association was found in previous studies (Gutteling et al., 2002, although a causal relationship cannot be drawn from this or previous evidence. In line with a reduction of media coverage and reporting, the climate of opinion towards food biotechnology has improved since 1999. While a majority of Europeans were opposed to such technology in 1999, a small majority now support the planting of GM crops, while the public is split 50/50 on GM foods (Gaskell et al., 2003). Therefore, o although public perceptions have been hypothesized to deteriorate in light of increased press coverage, the UK and the Spanish public have become more optimistic about biotechnology after 1999, although recovery indexes appear to be weaker in the UK, while Spain seems to have experienced such an increase in public optimism that in 2002 it stands above the EU average (Figure 3).



Question: 'I am going to read out a list of areas in which new technologies are currently being developed. For each of these areas, please state whether you think it will improve (value=1) our way of life in the next 20 years, it will have no effect (value=0) or it will make things worse (value=-1)'.

Figure 3. Optimism towards Biotechnology in EU Countries.

This leads us to distinguish between 'cumulative' and 'time specific' exposure. According to this evidence, cumulative exposure in 2003 captures the press coverage all over the period. However, what we do only observe is press coverage in one point in time instead. Yet, disentangling the effect of the media is a complex issue, bearing in mind the intensity of media exposure from 1999 to 2002, evidence indicates that awareness over GM food might have led to what is referred to as 'media created knowledge', presumably affecting public perceptions. Indeed, the 'time specific' media coverage peaked in 1999, as did literacy over the genetic modification of food (Gaskell et al. 2003).

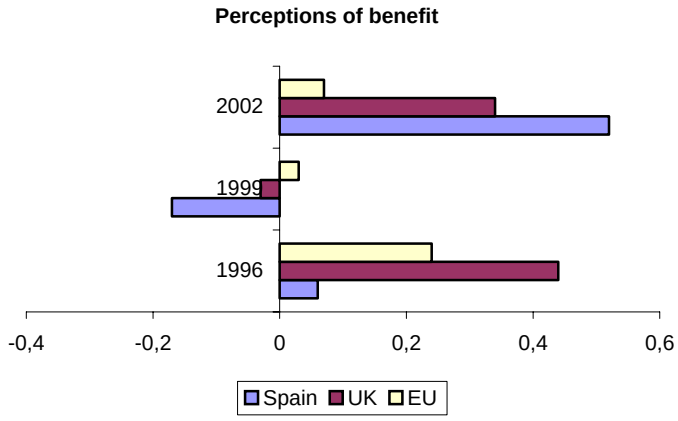
Finally, another key issue with regard to the role of the media is the individual's willingness to devote time to updating their information on biotechnology (demand for information). Comparing survey results for 1999 and 2002 we find that although in the UK a larger share of the population would take time to read articles or watch TV programmes (71 per cent) on the advantages and disadvantages of recent developments in biotechnology as compared to Spain (59 per cent),[6], in 2002 we find a significant decline in the Spanish demand for information (41 per cent) while there is a less pronounced delcine in the UK (67 per cent). This result, when viewed in the context of the less intensive reporting in Spain, indicates that whilst there is an ongoing public debate in the UK the potential debate in Spain is less intensive and potentially declining, although both countries exhibit a similar share of the population that demand information on GM food.

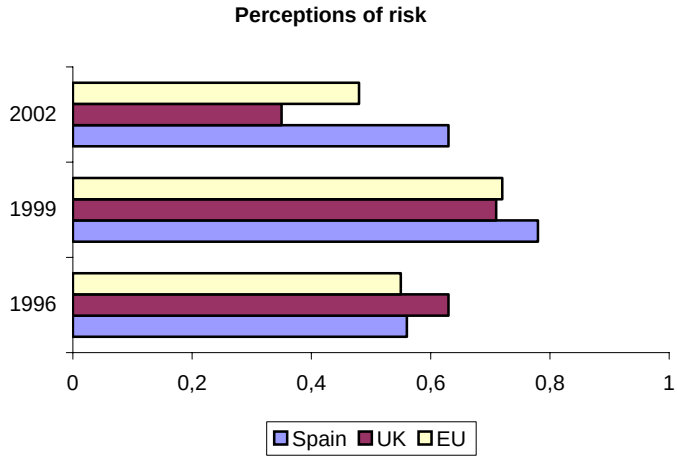
5.2. Media Coverage Versus Perception of Risk

A number of hypotheses have been drawn which point to the media as being responsible for public risk perceptions regarding several scientific developments. The evidence examined suggests that people in the UK and in Spain consider the use of modern technology in the production of foods to be a decidedly risky business (Figure 4). On the other hand, research shows that the lay public see some biotech procedures as beneficial, since their use may make it possible to feed more people in a more efficient way, ultimately benefiting consumers, but at the same time the potential human risks and ethical concerns are recognised to be causing anxiety (Gaskell et al., 1997).

According to Figure 4, the perceptions of benefit of GM food in Spain and the UK varied markedly across the period 1996–2002, especially in Spain. Indeed, although in 1996, the perceptions of benefit in Spain were lower than those in the UK and the EU average, by 2002 the perceptions of the Spanish population were higher on average than the other two. Finally, on average those perceiving greater benefits outweighed those that did not at all times in the EU, even in 1999. The perceptions of risks show a similar pattern to that of optimism, peaking in 1999 whilst by 2002 risk perceptions in the UK were even lower than those of 1996, before the controversy blew up in the press. A remarkable across-country difference emerges as regards moral acceptance, which is lower on average in the EU countries as compared to the UK and Spain.

[6] Whilst 21% in the UK disagreed with the sentence 'I would make time to read articles or watch TV programmes on the advantages and disadvantages of recent developments in biotechnology' and 8% didn't know, in Spain 27% disagreed and 14% provided a don't know response. Interestingly, we found that the UK ranked first in having the highest share of the population (30%) who answered that newspapers 'do not do good work'.





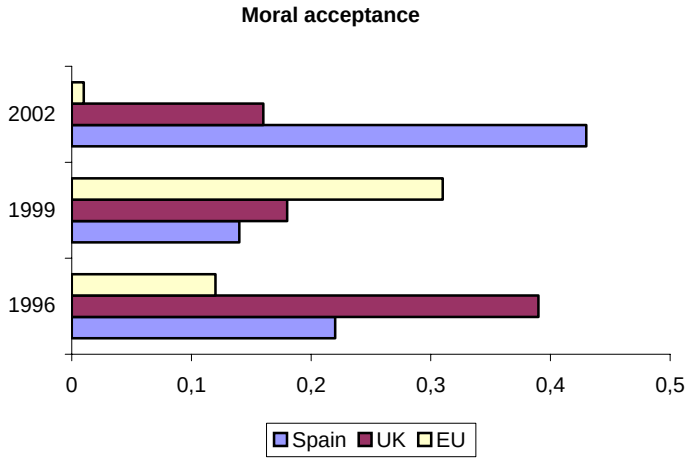
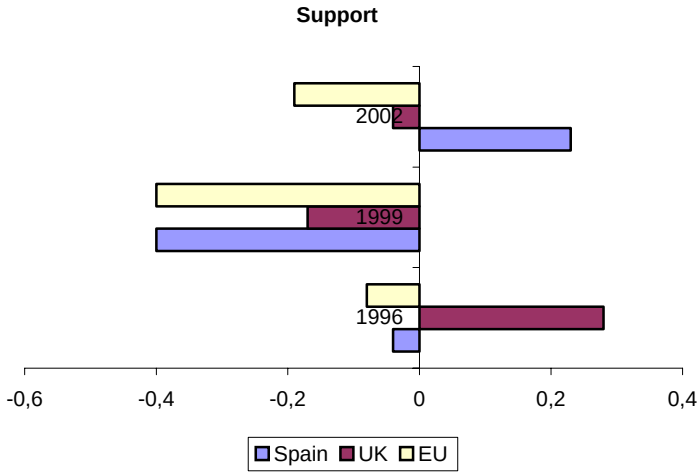


Figure 4. (Continued).



Question: To what extent do you agree / disagree that this application should be encouraged/ is useful/risky/morally acceptable?

Note: index estimated by giving a value of +2 to definitely agree answers, 1 to tends to agree, =1 tends to disagree and –2 to definitely disagreesSource: Eurobarometer, various years.

Figure 4. Public Perceptions on the Use of Modern Technology in the Production of GM Foods.

Finally, support for biotechnology exhibits significant changes over time. Whilst in 1996 only the UK revealed high levels of support, in 1999 there was an agreement over the three areas of study not to support GM food and in 2002 only Spain supported GM food.

Finding potentially unambiguous explanations for perception changes is an issue that requires additional information. In another Eurobarometer survey, Eurobarometer 55.2 collected in 2001 (see Appenedix A1) it was found that whilst in the UK and Europe about 60–66 per cent of the population understood the information disseminated by the press on GM food approximately half (33 per cent) of that amount did so in Spain. In both the UK and Spain, the sources of information most trusted 'as telling the truth about modern biotechnology' by the public were the medical profession (26 per cent and 29 per cent, respectively), consumer (17 per cent and 22 per cent, respectively) and environmental organisations (16 per cent and 17 per cent, respectively)[7]. On the other hand, the media ranked among the least trusted (three per cent and four per cent, respectively). Therefore, it is worth pointing out that although the media are often regarded as responsible for socially amplifying risk, if they are among the least trusted stakeholders one might expect to find readers to be more critical with the information reported. Thus, a lack of trust might be a key element in the role of the media in influencing attitudes.

Interestingly, journalism is more highly respected in Spain than in the UK and the rest of Europe (Appendix A2), and in the event of a disaster 25 per cent of the Spanish public would trust the media, whilst only 12 per cent of Europeans would do so. Therefore, although one potential explanation may be linked respect for journalism and trust in the media in reporting on catastrophic issues is higher in Spain, and thus the Spanish public could, arguably, be

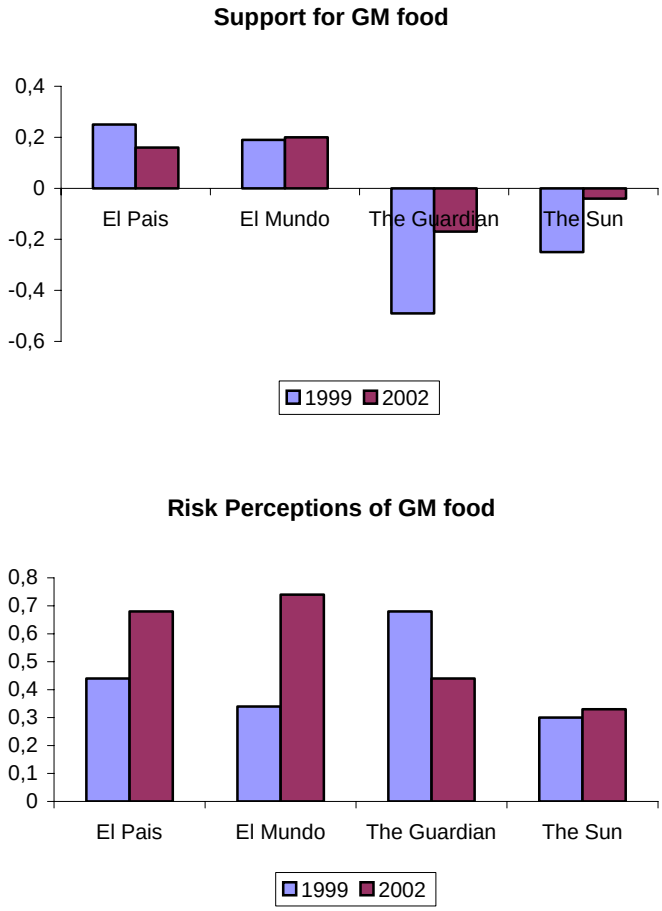
exhibiting a higher sensitivity to the media, which is consistent with the lower literacy of their population. Distrust is a particularly relevant factor among UK society, where 30 per cent declare that newspapers '*do not do good work*'. The evidence from the 2002 Eurobarometer survey shows that whilst in Europe and Spain only 15 per cent and 12 per cent respectively disagreed with the statement that 'newspaper and magazine reporting on biotechnology' are doing a good job for society, in the UK the share reached 30 per cent. Furthermore, in the 2002 survey, 33 per cent of UK respondents and 31 per cent of Spanish (and average EU) respondents disagreed with the statement 'I am sure about my opinions about genetically modified food'. Similarly, 53 per cent and 56 per cent of Spanish and UK citizens, respectively, disagreed with the statement 'It is easy for me to form an accurate judgement on genetically modified food', compared to an average of 49 per cent for the EU in 2002.

Finally, in the 2002 Eurobarometer survey there was a question in which dimension was more important in determining public perceptions. Interestingly, whilst in Spain and across the European Union the perceptions of benefit outweighed perceptions of risk, in the UK the same share of the population judged GM food on the basis of risks as those who focused on its usefulness (44 per cent). Accordingly, specific public perception changes might result from a greater proportion of information on risks within the information system.

Figure 5, provides evidence of potential media outlet specific effects. Whilst in Spain there are no significant differences in support and risk perceptions among newspapers, in the UK differences were marked, although by 2002 they had become less prevalent. Systematically, support was lower among the readers of *The Guardian* as compared to readers of *The Sun* in 1999 and 2002. However, whilst in 1999 risk perceptions were greater for the readers of *The Guardian*, in 2002 risk perception differences between the media outlets were insignificant. Thus, evidence from the newspapers examined highlights how different journalistic treatment and coverage has an influence on public perceptions of GM food and biotechnology in general. A potential reason for the difference in public perceptions between Spain and the UK lies in the fact that the tabloid press does not exist in the former. Therefore, the disparities among consumers of different newspapers appear to be weaker. Only very small differences emerge, suggesting that *El Mundo* readers tend to perceive fewer risks than *El Pais* readers.

This perception can be linked to the fact that *El Pais* gives more coverage to the topic and, consequently, greater controversy is generated. Another reason might lie in the fact that environmentalists' opinions may be given more space in left-wing newspapers. However, unlike in the UK, given the limited effect there is no evidence of an 'ideology bias' affecting public perceptions. Finally, one can imagine that even though the intensity of coverage might decrease over time, public perception might still be strongly influenced by coverage if it becomes decidedly more positive or negative over time.

[7] *Question:* Now I would like to ask you which of the following sources of information, if any, you trust to tell you the truth about modern biotechnology. a) Please choose the source of information YOU TRUST MOST, if any, from the following list.



Question: To what extent do you agree /disagree that this application should be encouraged (acceptance) or is risky? Source: Eurobarometer 52.1

Figure 5. Acceptance and Risk perceptions of GM Food and Newspaper Readership in 1999 and 2002.

6. CONCLUSIONS

This paper has intended to show that risk communication is a multifaceted process that involves individual psychology as well as specific effects resulting from the way that an issue is framed, the extent of understanding that individuals have on an issue, as well as trust and respect for the media in a specific country. As predicted by the renowned sociologist Ulrich Beck, individuals will be increasingly dominated by large-scale environmental dangers, forcing us to make decisions on the risks that we are, or are not, willing to take. Information sources and organs of mediation, the media among them, will be key actors in this process. Therefore, substantive efforts need to be made to communicate risks in order to allow individuals to make informed decisions. In this study, we have sought to shed some light on this complex process. The results are exploratory and they should be taken as "suggestive" rather than as "definitive". The comparison between some UK and Spanish newspapers, as well as the existing divergences in the knowledge and optimism of the UK and Spanish

populations about GM food has led to some arguably intriguing evidence and further questions.

Overall, qualitative content analysis reveals that coverage in the UK and Spain has been characterised by an extreme focus on *risks and the potential hazards to public health* (conversely, US media coverage has been predominantly positive and one-sided). Overall, European coverage is to some extent driven towards negative effects, framing the reality of GM food as a highly controversial issue (R1). Thus, the British and Spanish media coverage is driven mostly by controversy, as the dominant themes rarely display a positive emphasis on the potential benefits of GM food. The empirical evidence on media content reveals that *GM food is a topical issue within the British press, as a high level of reporting led GM food to become a front-page news story 13 times during 1999*. Within a press that influences opinion leaders and policy-makers, the theme emerges as a complex debate that involves many stakeholders, namely scientists, politicians and state headquarters, food companies and other interested parties. All employ the media as a risk communication tool. In sum, GM food is a *multifaceted topic with many interest groups trying to defend their own positions and all leading to divergent published commentaries*.

In the UK we find a remarkable variation between broadsheet and tabloid reporting. Whereas *The Guardian* acts as a significant agenda-setter in the GM food debate, *The Sun* courts extensive popular attention. Contrary to *The Guardian*'s broad and somewhat pedagogical coverage, it was noticeable that a detailed debate was rarely found in *The Sun*. Unlike the UK, GM food was not reported as a topical issue by the Spanish press, where there was less coverage, resulting in less controversy, with comparatively little public debate, which confirms the idea that 'no news is good news'. Possibly, some cross-country differences might have to do with the structure of the press media, insofar as a tabloid press does not exist in Spain.

As noted by Kalaitzandonakes et al. (2004), media effects might be short term. Indeed, although an examination of coverage in the UK and Spain indicates a significant *increase in the salience of GM food in the early 1999*, after a peak in 1999 there was certainly a reduction in the volume of newspaper coverage, although the intensity of the debate was still significant (R2). Moreover, although we find media specific biases resulting from specific media outlet ideology, the media is not the preferred (and most trusted) information-updating tool and significant differences are found across countries in the understanding of the GM food debate (R2). However, information disseminated in the media might succeed in establishing a debate and affecting other information stakeholders that might be more trusted by the population. Therefore, when *the media provides extensive coverage on an issue involving important risks to society, this results in greater perceived uncertainty, which might encompass some individuals' distrust*. The more limited support for GM food in the UK as compared to Spain after 2002 can be understood as a consequence of a major social debate played out in the media, which has negatively influenced public awareness of GM food (R3). This shared sensitivity is plainly expressed in the words of a journalist in *The Guardian* who claims; '*it's something the public keeps asking for. It's certainty*' (*The Guardian*, 23 February 1999).

On the other hand, the population's perception of risk is complex, involving a vast array of factors besides the media (R3). Although survey data reveal that public perceptions of risk regarding GM food have increased in line with news reporting, we argue that there is *no conclusive evidence to establish a direct cause-effect relationship between negatively biased news and the lack of public trust in the field.* Available data do not prove that treatment of the

issue in the media is the only determining factor in the perception of risk. Furthermore, research on lay people's risk perception reveals the influence of the media in enhancing 'ambiguity', when reporting both positive and negative information. It has long been acknowledged that people are ambiguity averse. That is, people fear the unknown and prefer known risks to uncertain ones, although these may result in lesser objective risks than the former. The relative loss for a low probability event is perceived as higher when the alternative contains ambiguous risk information (Viscusi, 1998). Therefore, as long as information on GM food remains technical and risk communication is unable to provide clear-cut information on the risks, one might expect that individuals will suffer from ambiguity aversion. This effectively implies that the safe although conservative alternative of an unsupportive public with regard to GM food is expected to continue.

Historical, socio-political and cultural backgrounds must also to be acknowledged. For instance, Spanish ignorance with regard to science remains as an historical reminder of 40 years of an authoritarian regime, during which time certain information channels, such as the media, were subject to censure. This may have led an important cohort of the Spanish population to suppress their intellectual and creative freedom to express their views on strategic issues for Spanish agriculture such as GM food. Spain's 'new' democracy, in contrast to the 'long-running' version in the UK, exhibits some cultural anxiety that has been fuelled by the erosion of collective values resulting from the rapid ascendancy of a market economy and enterprise culture alongside a process of European integration that challenges the Spanish economy. Cultural conditions should not be overlooked when analysing specific social controversies such as that around GM food, as the agenda setting provided by the media is to a large extent sustained by specific national cultural and political backgrounds (McCombs et al,1997). What remains unclear is the way that the media can best contribute to a democratic debate on major social issues. Informing the public about the risks and benefits on issues that constrain innovation should be undertaken in the light of known individual psychological responses. Finally, a key information source is education, given that the technical nature of biotechnology information might generate the need to provide the public with a better education on scientific developments.

APPENDIX

Table A1. Understanding of GM food in the European Union, Spain and UK (%)

	Europe	Spain	UK
Think that understand	60.9	35.3	66
Do not think that understand	33.4	60.2	26.8
Don't know	5.7	4.5	7.2

Source: Eurobarometer, 55.2, 2001.

Question: In recent years, newspapers and TV have regularly dealt with the following issues (Genetically Modified Food). Could you tell me if you think you understand or not each of the following?

Table A2. Journalist's appreciation

	Europe	Spain	UK
Respect for journalists [a]			
-	85.38	73.09	94.68
+	14.62	26.91	5.32
Trust in journalists in case of a major disaster [b]			
-	79.65	74.95	87.78
+	20.35	25.05	12.22

Source: Eurobarometer, 55.2, 2001.

[a]Question *:* For which of the following professions do you have the highest regard? (Journalism).

[b]Question: Suppose that there was a major disaster in your area. Who would you trust most to explain the reasons for the disaster? (The Media).

REFERENCES

Alsem, K-J., Brakman, S., Hoogduin, L. and Kuper, G. (2004) *The Impact of the Newspapers on Consumer Confidence: Does Spin Bias Exist?* CESIFO, Working Paper No 1328.

Anderson, A. (2000) 'Environmental Pressure Politics and the Risk Society,' in S. Allan, Adam, B. and Carter, C. (eds.), *Environmental Risks and the Media* (Routledge).

Bauer, M.W., Kohring, M., Gutteling, J. and Allansdottir, A. (2001) The Dramatisation of Biotechnology in the Elite Mass Media, in G. Gaskell and M.W. Bauer (eds), *Biotechnology 1996-2000 — The Years of Controversy* (London, Science Museum), pp. 35–52.

Beck, U. (1992) *Risk Society: Towards a New Modernity* (Sage Publications).

Becker, P. and McCombs, M. (1975) 'The Development of Political Cognition,' in S.H.Chaffee (ed.) '*Political Communication: Issues and Strategies*,' (Beverly Hills, California, Sage).

Besley, J.C and Shanahan, J (2005). Media Attention and Exposure in Relation to Support for Agricultural Biotechnology. *Science Communication*, 26 (4): 347-367.

Blumler, J.G. (1977) *The Political Effects of Mass Communication* (Open University Press).

Camerer, C.F. and Weber, M. (1992) 'Recent Developments in Modelling Preferences: Uncertainty and Ambiguity,' *Journal of Risk and Uncertainty*, no. 5, pp. 325–70.

Covello, V.T. (1992) 'Risk Communication: an Emerging Area of Health Communication Research,' in S. Deetz (ed.), *Communication Yearbook* (15th ed.) (Newbury Park: Sage Publications).

Dosman, D.M., Adamowicz, W.L. and Hurdey, S.E. (2001) 'Socioeconomic Determinants of Health and Food Safety Related Risk Perceptions,' *Risk Analysis*, vol. 21, pp. 307–17.

European Commission (EC) (2003) 'Europeans and Biotechnology in 2002,' *Eurobarometer 58.0* (Brussels, The European Commission, March).

Fischhoff, B. and Fischhoff, I. (2001) 'Publics' Opinions about Biotechnology,' *AgBioForum*, vol. 4, pp. 155–62.

Frewer, L.J., Miles, S. and Marsh, R. (2002) 'The Media and Genetically Modified Foods: Evidence in Support of Social Amplification of Risk,' *Risk Analysis*, vol. 22, no. 4, pp. 701–11.

Gaskell., G. (2000) 'Agricultural Biotechnology and Public Attitudes in the European Union,' *AgBioForum*, vol. 3, pp. 87–96.

Gaskell, G. et al. (1997) 'Europe Ambivalent on Biotechnology,' *Nature*, vol. 387, pp. 845–7.

Gaskell, G. et al. (1999) 'Worlds Apart? The Reception of Genetically Modified Foods in Europe and the U.S.,' *Science*, vol. 285, pp. 384–6.

Gaskell, G. et al. (2000) 'Biotechnology and the European Public,' *Nature Biotechnology*, vol. 18, pp. 935–42.

Gaskell, G. et al. (2003) 'Ambivalent GM Nation? Public Attitudes to Biotechnology in the UK, 1991–2002,' *Life Sciences in European Society Report: London School of Economics and Political Science*, July 2003.

Gaskell, G., Bauer, M. and Durant, (1998) 'The Representation of Biotechnology: Policy, Media and Public Perception,' in M. Drant, W. Bauer and G. Gaskell, *Biotechnology in the Public Sphere*, A European Source Book (Science Museum, Exhibition Road, London).

Gunter, B. and Wober, M. (1983) 'Television Viewing and Public Perceptions of Risks to Life' *Journal of Environmental Psychology*, vol. 3, 325–35.

Guteling , J.M. et al. (2002) 'Media Coverage 1973–1996: Trends and Dynamics,' in M.W. Bauer and G. Gaskell (eds.) *Biotechnology — the Making of a Global Controversy* (Cambridge: Cambridge University Press), pp. 95–128.

Hallin, D. (1994) *We keep America on the Top of the World. Television, Journalism and the Public Sphere* (Routledge).

Harrop, M. (1987) 'Voters and the Media,' in J. Seaton and B. Pimlott (eds), *The Media in British Politics* (Aldershot: Gower).

Hoban, T.J. (1997) 'Consumer Acceptance of Biotechnology: An International Perspective,' *Nature Biotechnology*, vol. 15, pp. 232–34.

Hornig Priest, S. (1992) 'Science Stories: Risk, Power and Perceived Emphasis,' *Journalism Quarterly*, vol. 67, no. 4, pp. 767–76.

Hornig Priest, S. (1993) 'Reading Risk: Public Response to Print Media Accounts of Technological Risk,' *PublicUnderstanding of Science*, vol. 2, pp. 95–109.

Hornig Priest, S. (1994) 'Structuring Public Debate on Biotechnology: Media Frames and Public Response,' *Science Communication*, vol. 16, no. 2, pp. 166–79.

Hornig Priest, S. (2001) *A Grain of Truth: The Media, the Public and Biotechnology* (Lanham MD: Rowan and Littlefield).

Innis, H. (1991) T*he Bias of Communication* (Toronto: University of Toronto Press).

Johnson, B.B. and Covello, T.V. (1987) 'Agenda Setting Group Conflict, and the Social Construction of Risk,' in B.B. Johnson and V.T. Covello (eds.), *The Social and Cultural Construction of Risk* (Dordrecht: D. Reidel Publicising).

Kalaitzandonakes, N., Marks, L.A. and Vickner, S.S. (2004) 'Media Coverage of Biotech Foods and Influence of Consumer Choice,' *American Journal of Agriculture Economics*, vol. 5, pp.1238–46.

Kanheman, D. and Tversky, A. (1984) 'Choices, Values and Frames,' *American Psychologist*, vol. 39, no. 4, pp. 341–50.

Kasperson, R.E., Renn, O. and Slovic P. et al. (1988) 'Social Amplification of Risk: A Conceptual Framework,' *Risk Analysis*, vol. 8, pp. 177–87.

Keeney, R.L. and von Winterfeldt, D. (2001) 'Appraising the Precautionary Principle — a Decision Analysis Perspective,' *Journal of Risk Research*, vol. 4, pp. 191–202.

Leisinger, K.M. (2002) 'Science and Public Trust,' *1st Dialogue on Science* (Engelberg, Switzerland, 23–25 October).

Logan, R.A. (2001) 'News Compartmentalization: Implications for Food Biotechnology Coverage,' *AgBioForum*, vol. 4, pp. 194–98.

Lujan, J.L. and Todt, O. (2000) 'Perceptions, Attitudes and Ethical Valuations: The Ambivalence of the Image of Biotechnology in Spain,' *Public Understanding of Science*, vol. 9, pp. 383–92.

Marks, L.A. and Kalaitzandonakes, N. (2001) 'Mass Media Communications about Agrobiotechnology,' *AgBioForum*, vol. 4, pp. 199–208.

Marks, L.A., Kalaitzandonakes, N., Allison, K. and Zakharova, L. (2002) 'Time Series Analysis of Risk Frames in Media Communication of Agrobiotechnology,' in V. Sananiello and D. Zilberman (eds), *Market Developments and Genetically Modified Foods* (CAB International).

McCombs, M.E., and Shaw, D.L., and Weaver, D.L. (1997). *Communication and Democracy: Exploring the Intellectual Frontiers in Agenda-Setting Theory*. Mahwah, N.J. Lawrence Erlbaum.

Miller, M. and Parnell Riechert, N. (2000) 'Interest Groups Strategies and Journalistic Norms. News Media Framing of Environmental Issues,' in S. Allan, B. Adam and C. Carter (eds), *Environmental Risks and the Media* (Routledge).

Mullainathan, S. and Shleifer, A. (2002) *Media Bias*, National Bureau of Economic Research, Working Paper No. 9295.

Nisbet, M. and Lewenstein, B.V. (2001) 'A Comparison of U.S. Media Coverage of Biotechnology with Public Perceptions of Genetic Engineering 1995–1999,' paper presented to the *2001 International Public Communication of Science and Technology Conference*, Geneva, Switzerland, 1–3 February.

Palfreman, J. (2001) 'Sending Messages Nobody Wants to Hear: A Primer in Risk Communication,' *AgBioForum*, vol. 4, pp. 173–8.

Petts, J., Horlick-Jones, T., Hargreaves, D., McLachlan, S. and Löfstedt, R. (2001). *Social Amplification of Risk: The Media and the Public* (Health and Safety Executive [http://www.hse.gov.uk/research/crr_pdf/2001/crr01329.pdf, visited April 2004]).

Scherer, C.W. and Cho, H. (2003) 'A Social Network Contagion Theory of Risk Perception,' *Risk Analysis*, vol. 23, pp. 261–67.

Sheehy, N., Wylie, J. and McKeown, G. (2002) *Quantifying Risk Amplification Processes: A Multi-level Approach* (Health and Safety Executive. [www.hse.gov.uk/research/crr_pdf/2002/crr02367.pdf, visited April 2004]).

Siegrist, M. (2000) 'The Influence of Trust and Perceptions of Risks and Benefits on the Acceptance of Gene Technology,' *Risk Analysis*, vol. 20, no. 2, pp. 195–203.

Siegrist, M. and Cvetkovich, G. (2001) 'Better Negative than Positive? Evidence of a Bias for Negative Information about Possible Health Dangers,' *Risk Analysis*, vol. 21, pp. 199–206.

Slovic, P. (1987) 'Perception of Risk,' *Science*, New Series, vol. 236, pp. 280–85.

Splichal, S. and Wasko, J. (eds) (1993) *Communication and Democracy*, Ablex Publishing Corp.

Swartz, D.G. and Strand, I.E. (1981) 'Avoidance Costs Associated with Imperfect Information: the case of Keepone,' *Land Economics*, vol. 57, pp. 139–150.

United States National Research Council (1989) *Improving Risk Communication* (Washington, DC: National Academy Press Committee on Risk Perception and Communication).

Verbeke, W. and Ward, R.W. (2001) 'A Fresh Meat Almost Ideal Demand System Incorporating Negative TV Press and Advertising Impact,' *Agricultural Economics*, vol. 25, pp. 359–74.

Viscusi, K. (1998) *Rational Risk Policy* (Oxford: Clarendon Press).

Wahlberg, A. and Sjoberg, L. (2000) 'Risk Perception and the Media,' *Journal of Risk Research*, vol. 3, no. 1, pp. 31–50.

Wilson, K.M. (2000) 'Communicating Climate Change through the Media. Predictions, Politics and Perceptions of Risk,' S. Allan, B. Adam and C. Carter (eds.), *Environmental Risks and the Media* (Routledge).

Worcester, RM. (1999) 'Science and Democracy: Public Attitudes to Science and Scientists,' *World Conference on Science*, 28 June.

Wyse, R. and Krivi, G.G. (1987) 'Strategies for Educating the Public Concerning Biotechnology,' in L.R. Batra and W. Klasses (eds), *Public Perceptions of Biotechnology* (Agricultural Research Institute, Bethesda).

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 111-135

Chapter 5

COMPARING RISKY AND INTER-TEMPORAL DECISIONS: VIEWS FROM PSYCHOLOGY, ECOLOGY AND MICROECONOMICS

***Tobias Kalenscher*[*1] *and Philippe N. Tobler*[2]**
[1]Animal Physiology and Cognitive Neuroscience
Neurobiology Section of Swammerdam Institute for Life Sciences (SILS)
University of Amsterdam, Faculty of Science
Kruislaan 320, 1098 SM, Amsterdam, The Netherlands
[2]Department of Physiology, Development and Neuroscience
University of Cambridge, Downing Street,
Cambridge CB2 3DY, United Kingdom

ABSTRACT

When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints.

Risky and inter-temporal decisions have been captured theoretically primarily by Ecology and Microeconomics but findings from Behavioral Economics, Psychology and Neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive models has emerged to explain the findings. A subset of these models has stressed the similarities between risky and inter-temporal decisions. In this chapter we review both core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

* Ph: +31 (0)20 525 7658; Fax: +31 (0)20 525 7709: Email T.Kalenscher@uva.nl Web: http://home.medewerker.uva.nl/t.kalenscher/

1. INTRODUCTION

When we make decisions, the outcomes of our choices rarely occur with certainty, and often we have to wait some time for the consequences to happen. For example, investing time and money into a good education makes more likely, but doesn't guarantee, a successful professional career and a high income; in other words, the outcome of your investment decision is probabilistic. Likewise, when paying the high tuition fees for your education, you invest resources now for benefits that are yet to come, because you will only be able to harvest the fruits of your labor once you finish your education in a couple of years. Choices between probabilistic outcomes are called 'risky decisions' and choices between outcomes that will be realized at different instants in the future are called 'inter-temporal decisions'. Both types of decisions have been extensively discussed in several scientific disciplines, including biology, ecology, micro- and macroeconomics, psychology and cognitive neuroscience. In this chapter, we review some of the most influential theories on risky and inter-temporal decision making, and will outline the theoretical and empirical differences in the different approaches. We will then discuss to what degree attempts to unify the two research fields are and can be successful and identify key behavioral and neuroscientific experiments. We conclude with highlighting the importance of cooperation between the various disciplines in elucidating the effects of risk and time on choice.

2. DECISIONS WITHOUT RISK

Decisions between certain, immediate, but quantitatively different choice outcomes appear easy: You just compare which of the two outcomes results in the higher gain or the smaller loss, and choose accordingly. However, how do you compare two qualitatively different commodities, for example, apples and pears? In economics, this problem is solved by assuming that different commodities are translated into a common currency, the subjective value, or the utility of a prospect. Utility is a measure of relative satisfaction or gratification which allows to rank-order and therefore compare the different possible outcomes (Montague and Berns, 2002). Although frequently used in financial contexts (also in this chapter), utility does not exclusively refer to monetary gains (and losses), but also to more abstract benefits, such as obtaining pleasure from engaging in a favorite recreational activity, or enjoying one's favorite food, or the like. Although embracing essentially the same solution, behavioral ecology has given a biological twist to the common currency problem. It replaces utility with fitness which, depending on the model, may correspond to e.g. rate of energy gained per unit of time spent foraging (Charnov, 1976) or to reproductive success (Hamilton, 1964).

The utility of an outcome is not a linear function of its objective (e.g., monetary) value, but a function of the current level of wealth (Friedman and Savage, 1948; Bernoulli, 1954; Kahneman and Tversky, 1979; Tobler et al., 2007a). More precisely, it has been argued that each additional unit in the utility function, the so-called marginal utility (Mankiw, 2004), is smaller than the previous unit, resulting in a progressive decrease in marginal utility with increasing assets or energy reserves (Friedman and Savage, 1948; Bernoulli, 1954; Sibly and McFarland, 1976; Kahneman and Tversky, 1979; Kacelnik, 1997; Kacelnik and Bateson, 1997; Tobler et al., 2007a). As a consequence, the utility of a commodity is presumed to be a

decelerating concave function of this commodity (figure 1A). For example, winning $100 would be more valuable to you when you are poor than when you are a millionaire.

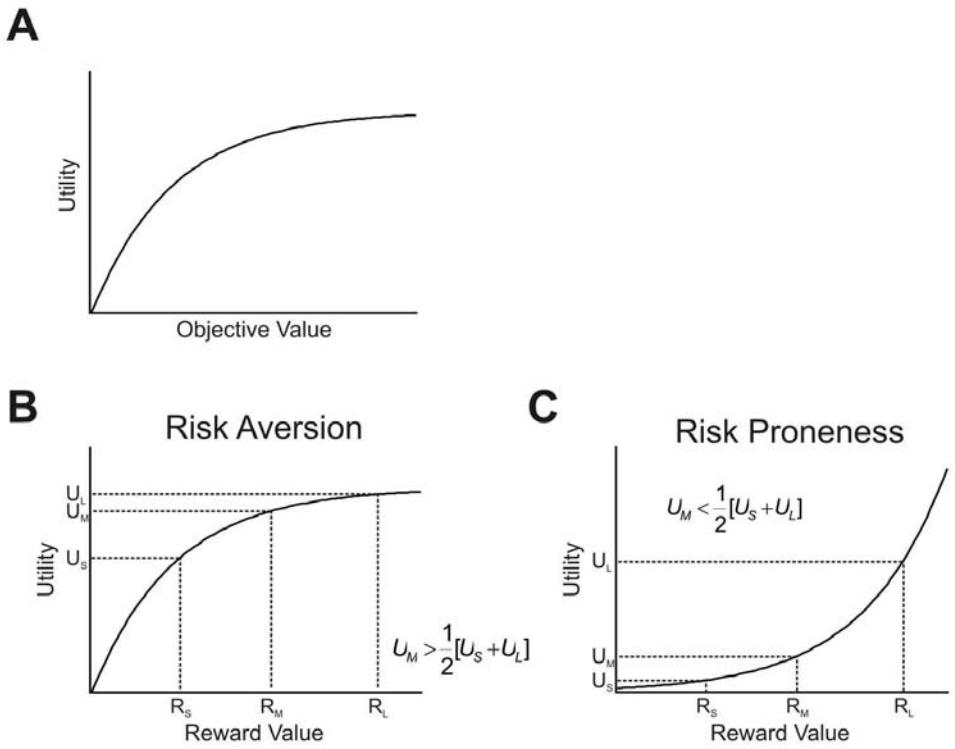


Figure 1. Utility functions. (A) Utility as a function of the objective value of a commodity. The utility curve is a concave function of the current level of wealth because the marginal utility, i.e., the utility increment with each additional unit, decreases with increasing level of wealth. (B) A concave utility function predicts risk aversion when choosing between a medium-sized, certain and a large and small risky reward. In a multi-choice situation, the average utility of the certain rewards exceeds the average utility of the risky rewards because the utility of the large reward is sublinearly larger compared to the utilities of the other rewards. (C) A convex function predicts risk proneness.

3. DECISIONS UNDER RISK

Risky decisions are decisions between probabilistic outcomes. The level of riskiness is equivalent to the spread from the mean (variance) of the risky outcomes. For example, a gamble that pays either $ 9 or 11 with 50% probability is less risky than a gamble that pays either $ 2 or 18 with 50 % probability (although both gambles have the same mean payoff of $ 10). This notion of risky decision making is important, as it differs from the common, folk-psychological conception of risky behavior, which frequently implies that a person engaging in risky choices is consciously willing to accept high losses (“the gambler who bets his house and family”). However, although the readiness to accept large losses may certainly play a role in biasing an individual’s risk attitude, academic research of risky decision making focuses on the formal impact of outcome variance on choice.

Neither humans nor animals are risk-neutral (Friedman and Savage, 1948; Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998): When deciding between a certain and a risky option, one is often chosen more frequently than the other, even when the expected values of the respective options (their probabilities multiplied with their objective values) are identical. Most animals, including humans, are risk-averse, but show occasional risk-proneness (Friedman and Savage, 1948; Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998; Glimcher, 2002; McCoy and Platt, 2005a, 2005b; Hayden and Platt, 2007; Tobler et al., 2007b). Importantly, this suggests subjective differences in the valuation of the same objective options. Utility theories have been put forward in order to capture subjective valuation with a particular focus on risk.

Economic Utility Theories

Utility theories assign numbers to preferences with both a descriptive and a normative purpose. Specially psychology is interested in explaining choice behavior and has thus focused on the descriptive aspect whereas economy and statistics, more so in the past, have put stronger weight on the normative aspect of characterizing consistent, coherent and optimal choice and ecology is interested in both descriptive and normative aspects. Utility theories all make basic assumptions (axioms) about the elements of the decision space and the preference relations of the decision maker with respect to these elements. From these axioms they deduce statements (theorems) for example about how the preference relations as observed from choices can be transformed into utility relations (numbers).

Von Neumann and Morgenstern (1944) axiomatized utility theory by requiring completeness, transitivity, continuity and (not explicitly stated but necessary) independence of preferences. Thereby they founded expected utility theory (EUT), which provides the most prominent normative framework for the analysis of decisions under risk. The completeness axiom requires that the decision maker has preferences across options, transitivity that preferences are in a basic hierarchical order, continuity that for each option there is a better and worse one and independence that preferences do not change by adding common outcomes to all options. If these axioms are fulfilled then a number u (x) can be assigned to each x so that:

$$x \leq y \text{ if and only if } u(x) <= u(y). \quad (1)$$

The preference relation $x \leq y$ may be read as “alternative x is not preferred to alternative y”, <= corresponds to the standard “smaller than or equal to” and $u(x)$ and $u(y)$ refer to the utilities of x and y. If this is true then utility function u preserves the ordering of $\leq$ and allows translating utilities to preferences and vice versa.

As long as the axioms described above hold EUT proposes that a decision whether to accept or reject a choice option should be made by multiplying the utility of all possible outcomes of the option with their probabilities, integrating across products, and choosing the option with the larger sum. Thus, the utility of a choice option with risky outcomes corresponds to its expected utility. For example, if you consider that it is relatively likely to receive a high income following a good education, and that with the expected income you

will be better than with your current situation, then you should decide to invest into a good education. In general, benefits represent positive utility values, costs negative values.

The shape of the utility curve can be related to people's risk attitudes. A risk neutral person has a linear utility curve, and for such a person the expected utility of a gamble is equivalent to the utility of the mean of the gamble. Convex utility curves correspond to risk-proneness, concave curves to risk aversion (figure 1B and 1C). To use a simple example, imagine a situation in which an agent chooses between a certain option, offering a medium-sized reward, and a risky option, offering a large and a small reward with a 50% chance each. The expected value of both options is identical. According to EUT, the expected utility of a given choice option is the sum of the utilities of each possible outcome multiplied with their probability:

$$E[U_{option}] = \sum p_{outcome} \bullet u_{outcome} \tag{2}$$

The expected utility of the certain option would accordingly be computed as:

$$E[U_{certain}] = 1 \bullet u_{medium-reward} \tag{3}$$

and the expected utility of the risky option would be computed as:

$$E[U_{risky}] = (0.5 \bullet u_{small-reward}) + (0.5 \bullet u_{large-reward}) \tag{4}$$

Due to the concavity of the utility function of risk averse agents, the utility of the large reward, $u_{large\text{-}reward}$, is sublinearly larger than the utilities of the medium or small rewards, $u_{medium\text{-}reward}$ and $u_{smaller\text{-}reward}$. Thus, EU_{risky} will be smaller than $EU_{certain}$ and agents will avoid the risky option (see figure 1B). Put in simpler words, an agent receives the large and small rewards with equal probability when choosing the risky option. Because the utility of the large reward is sublinearly smaller than the utilities of the other rewards, the mean utility of the large and small rewards (risky option) would be smaller than the utility of the gamble's expected value and thus the utility of the medium-sized certain option. As a consequence, the agent will avoid the risky option. Risk-proneness can be explained by assuming a convex, accelerating utility curve, in which the utility of the large reward is supralinearly larger than the utility of a medium or a small reward (figure 1C).

Violations of Preference Axioms

Empirical research showed violations of most of the normative axioms of EUT. Reports of violations of the independence axiom appeared relatively soon after von Neumann and Morgenstern's seminal work (Allais, 1953; Ellsberg, 1961). As an example, consider the following two decisions (Kahneman and Tversky, 1979):

Decision 1)
a) $ 2500, P = 0.33; $ 0, P = 0.67 (Read as: a 33% chance of winning $ 2500 and a 67% chance of winning nothing)
b) $ 2400, P = 0.34; $0, P = 0.66

Decision 2)
a) $ 2500, P = 0.33; $ 2400, P = 0.66; $ 0, P = 0.01
b) $ 2400, P = 1.0; $ 0, P = 0.0

Most people choose a) in decision 1) and b) in decision 2). However this pattern of preference reversals violates the independence axiom because 2a and 2b result from adding ($ 2400, P = 0.66) to 1a and 1b and therefore either a) or b) should be chosen in both cases.

Also the transitivity axiom can be systematically violated such that decision makers show cyclic preferences (A≥B, B≥C, C≥A; e.g. Loomes et al., 1991; Shafir, 1994; Waite, 2001). For example (Waite, 2001), blue jays prefer one raisin, 28 cm into a tube (option A) over two raisins, 42 cm into a tube (option B). They also prefer option B over three raisins 56 cm into a tube (option C) but when given the choice between options A and C, they do not prefer A. Humans also show systematic violations of transitivity in certain choice situations (Tversky, 1969), for example, when the choice options are composed of several features that vary along different dimensions. Models of context-dependent choice such as regret theory (Loomes and Sugden, 1982) suggest that violations of the transitivity axiom arise from changes in utility because decision makers evaluate options not in isolation but consider also the outcomes of unchosen alternatives.

Kahneman and Tversky (1979) have pointed out another problem of EUT in that it does not account for differences in how decision problems are described (framed). As an example, consider the following two decisions (Tversky and Kahneman, 1986):

Decision 3)
a) $ 240, P = 1.0
b) $ 1000, P = 0.25; $0, P = 0.75

Decision 4)
a) $ -750, P = 1.0
b) $ -1000, P = 0.75; $ 0, P = 0.25

Most people choose 3a and 4b. However, the combination of 3a and 4b is dominated by (has a lower expected value than) the combination of 3b and 4a.

Findings of axiom violations have provoked different reactions. Some theorists have relaxed one or more of the axioms (e.g. Machina, 1982; Fishburn, 1982) others have given up on the project of axiomatising utility theory and proposed purely descriptive models. The most famous of these latter approaches is prospect theory (Kahneman and Tversky, 1979; 1992). Prospect theory suggests that the subjective value function is concave for gains and convex for losses and steeper for losses than gains (figure 2). This reflects the finding that decision makers are usually risk seeking for losses, risk averse for gains and reluctant to accept a fair bet on the toss of a coin (in fact, potential gains have to exceed potential losses by a factor of about 2 in order to achieve indifference). The different steepness for gains and

losses introduces a "kink" in the value function which makes it difficult to treat mathematically and which has been termed "reference point". Although not formally defined, the reference point often corresponds to the status quo or the current wealth level. Moreover, decision weights modulate probabilities according to an inverted-S-shaped probability weighting function. This reflects the finding that many decision makers overweigh small and underweigh large probabilities, at least when making hypothetical decisions.

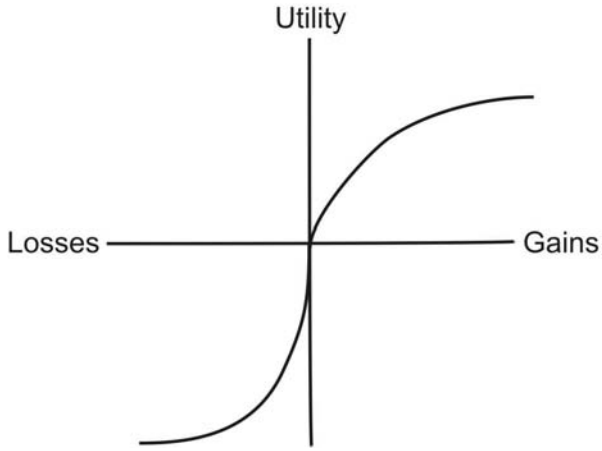


Figure 2. Utility function as proposed in Kahneman and Tversky's (1979) prospect theory. The utility function is concave in the domain of gains, and convex in the domain of losses, and steeper for losses than for gains. The crossing of the axes corresponds to the reference point against which the prospects are contrasted, for example, the current level of wealth.

Contribution from Behavioral Ecology Risk-Sensitive Foraging Theory

The importance of the utility function is to relate subjective to objective value. This is useful because subjective value is often not a linear function of objective value. However, what causes this non-linearity? Bernoulli (1954) suggested that wealth renders the utility function concave. Behavioral ecology suggests additional factors such as upper boundaries on how much energy can be stored and lower boundaries on how much energy is needed for survival can also introduce curvature or even "kinks" into the utility function. Consider for example birds foraging for the night. Because of their small body size and high metabolic rate they face the possibility of starving over night if they fail to accumulate enough resources during the day. As a consequence, a normally risk averse bird might become risk-prone towards the end of the day or at low temperatures, if its energy requirements for the night are not yet met (Caraco et al., 1990). Thus, in addition to varying over individuals, risk attitudes vary also over situations and time.

Energy requirements change not only over the course of a single day but also over the year, for example in the pre-migratory period of migratory birds (Moore and Simm, 1986). In that period birds must acquire sufficient reserves for the migration and they behave in a risk-prone manner until they reach maximal body size. Conversely, birds that are not in the pre-migratory period or pre-migratory birds that have reached maximal body size avoid risky foraging options.

Risk-sensitive foraging theory considers decision optimality given reserve constraints. Stephens (1981) showed that it is optimal for birds on a negative energy budget to choose options with higher variance. Conversely, they should avoid risky options when the less risky options provide the birds with a mean rate of intake that exceeds the starvation threshold, i.e. they are on a positive energy budget. For example, a bird almost starved to death (negative energy budget) should avoid the certain option if the certain food amount is not sufficient to guarantee its survival (certain death), and the only chance to survive would be to obtain the large, risky reward (possible survival). On the other hand, a bird on a higher energy budget should be risk-averse if the certain option is sufficient to guarantee survival (certain survival), and the small food amount of the risky option would be insufficient for survival (possible death). Current models of risk-sensitive foraging incorporate the possibility of sequential choices and the costs of foraging (e.g. McNamara and Houston, 1986).

In addition, life-history trade-offs may have favoured the evolution of risk-attitude in a similar fashion (Wolf et al., 2007). Animals differ in behaviours that affect their future fitness, for example, in their exploration effort during foraging. Animals that put more emphasis on future than immediate fitness returns have higher expectations regarding their future reproductive success than others. Investment into the future only pays out if the animal survives until it is able to realize the upcoming opportunities. Because survival is thus important for the strategy to work out, evolution may have favoured the development of risk-aversion towards predators and aggressive conspecifics in individuals with high fitness/reproductive expectations who have more to lose, whereas individuals with low expectations who have little to lose should be less risk-averse.

In summary, the minimal energy requirement (and by extension the maximal reserves) and/or the life-history of an animal may provide an inflection point ("kink"), where the curvature of its utility function changes.

A Normative Framework for Inter-Temporal Decisions

The second type of decisions making discussed in this chapter concerns decisions between outcomes that can only be realized at different instants in the future: Inter-temporal decisions. Humans and non-human animals prefer immediate over delayed rewards (so-called temporal discounting): Provided that the costs for both options are identical, the preference for an immediate or a temporally remote expected outcome is a function of the value of the respective outcomes and their delays, i.e., the time until the outcomes can be realized (McDiarmid and Rilling, 1965; Rachlin and Green, 1972; Ainslie, 1975; Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986; Logue, 1988; Benzion et al., 1989; Green et al., 1994, 1997; Evenden and Ryan, 1996; Evenden, 1999; Frederick et al., 2002; Reynolds et al., 2002; Kalenscher et al., 2005, 2006a).

Inter-temporal decisions have been extensively studied in psychology and ecology, but were and still are also of great interest in economic models of choice. As with decisions under risk, many of the normative models in economics are based on several theoretical assumptions and theorems, including preference monotonicity, stationarity, and maximization of utility rate.

A fundamental assumption in rational choice theories is that preference orders should be consistent across time. Preference monotonicity and stationarity are directly related to this

assumption. Monotonicity means that a prospect X1 that is preferred over another prospect X2 will also be assigned a higher utility than X2 as long as the utility function is monotonic. Monotonicity of time preference (Lancaster, 1963) holds that

$$X(t_1) \geq X(t_2), \text{ if, and only if, } t_2 \geq t_1 \quad (5)$$

This means that commodity X, available at timepoint t_1, will be preferred over X, available at timepoint t_2, only when t_2 occurs later than t_1.

Stationarity is related to the axiom of monotonicity of time preference and posits that

$$\text{If } X(t) \sim Y(t+\tau) \text{ then } X(s) \sim X(s+\tau) \quad (6)$$

This means that if an agent is indifferent (~) between commodity X, delivered at timepoint *t*, and commodity Y, delivered at timepoint $t+\tau$, he would still be indifferent when X was delivered at timepoint *s* and Y at timepoint $s+\tau$ (Strotz, 1955; Koopmans, 1960; Fishburn and Rubinstein, 1982). Indifference means that the utility of both options is identical, and thus the frequency of choosing A or B is about 50%. Thus, if both options were deferred by the same time interval, preference orders should be preserved. In other words, if you desire to receive $10 in 5 days as much as receiving $50 in 20 days, then you will still desire to receive $10 in 15 days as much as receiving $50 in 30 days, i.e., when both delays are prolonged by 10 days.

It was proposed that the discounting rate by which future commodities are delivered should be constant (Samuelson, 1937), for instance resulting in a linear or exponential discount function. Many theories, therefore, assumed exponential discounting (Lancaster, 1963; Fishburn and Rubinstein, 1982; Benzion et al., 1989; cf., Ainslie, 1975; Loewenstein, 1992; Fehr, 2002). Combining exponential discounting with stationarity yields (Lancaster, 1963):

$$(A,t) \sim Ae^{-k(t-t_0)} \quad (7)$$

This expression states that a reward with the amount *A*, delivered at timepoint *t*, is equally valuable (~) as a reward amount *A* at t_0 (i.e., now), exponentially discounted for the interval $t\text{-}t_0$, with t_0 referring to the present timepoint, and *k* being an individually different discount value. In other words, the utility of a future outcome can be expressed as an exponential function of the same outcome realized today.

Violation of Stationarity

As outlined above, stationarity predicts that the ranking of preferences between several future outcomes should be preserved when the choice outcomes are deferred by the same time interval. This has been investigated in an empirical study where human subjects chose between pairs of monetary rewards available after different delays (Green et al., 1994). Subjects preferred a small, short-delayed over a large, long-delayed reward, but their preference reversed away from the small towards the large reward when the delays to both

rewards were advanced by the same time interval. Notably, the prolongation of the delays resulted in a preference reversal even though the difference in the delays remained identical (Green et al., 1994). This finding therefore represents a violation of stationarity. Numerous other studies with human subjects (Ainslie, 1975; Logue, 1988; Benzion et al., 1989; Loewenstein, 1992; Kirby and Herrnstein, 1995; Green et al., 1997; Frederick, Loewenstein and O'Donoghue, 2002; McClure et al., 2004; Rohde, 2005), pigeons (Chung and Herrnstein, 1967; Rachlin and Green, 1972; Ainslie, 1974; Green et al., 1981) and rats (Ito and Asaki, 1982; Bennett, 2002) replicated and confirmed these results (cf., Kalenscher and Pennartz, *in preparation*). Thus, human and non-human animals systematically violate the crucial assumption of inter-temporal consistency of choice. Note that many studies in the animal literature do not defer both choice outcomes equally, but only one outcome is increasingly delayed, whereas the delay to the other outcome remains constant (see e.g. figure 3A). Preference reversals in those cases do not challenge stationarity, as changes in valuations would be predicted for the increasingly delayed outcome, but not the constant outcome.

The fact that human and animal subjects prefer the small, short-term reward over the large, delayed reward when the receipt to the small reward is near, but not when it is in the relatively far future, suggests that short-term rewards are discounted more steeply than long-term rewards. Such asymmetric discounting poses a strong challenge for the postulation of exponential discounting (Lancaster, 1963).

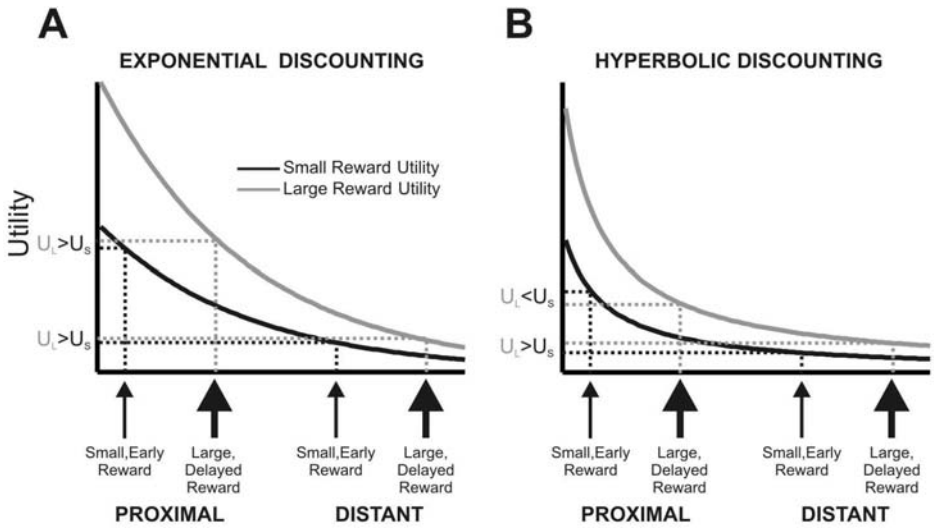


Figure 3. Exponential vs. hyperbolic discounting of future events. (A) Exponential utility curve of a large, delayed (grey line) and small, short-term reward (black line). With exponential discounting, stationarity holds because the utility of the large reward (UL) is always higher than the utility of the small reward (US). This is true when both rewards are temporally proximal, or when they are deferred by the same time interval into the future (distant rewards). (B) Hyperbolic discounting can explain preference reversals and the violation of stationarity. Due to the steeper decay for short delays, the utility of the small, short-term reward is higher than the large, delayed reward for temporally proximal rewards, but the utility order reverses when both rewards are deferred into the future.

Accordingly, as theoretically suggested by Ainslie (1975), and later empirically shown by Mazur (Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986) and others (Rachlin et al., 1991; Myerson and Green, 1995; Green and Myerson, 1996; Rohde, 2005; Jones and Rachlin,

2006; Laibson, 1997), discounting curves can be better approximated with hyperbolic than exponential or other constant discount functions, as outlined in figure 3. Why do humans and other animals systematically violate such crucial laws in economics?

Why Do We Discount the Future?

An assumption underlying most economic theories is utility maximization. This assumption is shared by many ecological theories of choice, namely the postulation that evolution favors choice mechanisms that maximize fitness levels, and minimize fitness losses. Applied to inter-temporal decision making, this means that the decision maker should act so as to maximize the utility rate, or in ecological terms, the energy intake rate per time unit (Stephens and Krebs, 1986). Rate maximization can explain why humans and animals sometimes prefer a less attractive, but temporally proximal outcome over a more attractive, but temporally remote outcome.

For example, we consider an inter-temporal choice task in which an animal has to choose between a small, always immediate reward and a large reward that is initially also delivered immediately, but that is delayed further as the experimental session progresses. Let's further assume that the large reward is 1.5 times as big as the small reward, and that the next choice opportunity always follows immediately after the animal has consumed its previous reward. The rate maximization hypothesis would predict that the animal should begin the session by preferring the large reward. However, the delay preceding the large reward gets longer and longer over the course of the session. So, at some point, the waiting time for the large reward, and thus the time until the next reward can be realized, will be more than twice as long as the delay preceding the small reward. Naturally, it would make sense now to prefer the small reward, as the hungry animal would be able to consume two small rewards in the same time that it would have to wait for only one large reward. Because two small rewards represent a larger food quantity than one large reward, the animal would maximize its energy intake per time unit by shifting its preference to the small reward once the delay preceding the large reward gets too long. In formal terms, optimal foraging theory assumes that organisms maximize, at least on the long run, the ratio of food intake and the time needed to obtain or consume the food, as described by the following quantity (Stephens and Krebs, 1986):

$$\max \frac{\sum_{i=1}^{\infty} G_i}{\sum_{i=1}^{\infty} t_i} \tag{8}$$

where G_i represents the net energy gain obtained from consuming the *i*th food item (here basically corresponding to its amount), and t_i represents the time between food item *i* and the previous food item *i-1*.

Animals Do Not (Always) Maximize Intake Rate

The above example implies that the next choice opportunity follows immediately after receipt of the reward. In many studies (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Evenden and Ryan, 1996; Cardinal et al., 2000; Isles et al., 2003, 2004; Winstanley et al., 2004, 2006; Kalenscher et al., 2005; Hwang et al., 2006; Louie and Glimcher, 2006), however, the inter-trial interval between reward and next choice was adjusted so that the total trial length was identical in all trials and independent of delay length and other factors. In such a scenario, the rate maximization hypothesis predicts that subjects should always choose the large reward, independent of the delay between response and reward, because only then would the animals maximize the total energy intake per trial, or per experimental session respectively. However, neither pigeons (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Kalenscher et al., 2005), nor rats (Evenden and Ryan, 1996; Cardinal et al., 2000; Winstanley et al., 2004, 2006; Roesch et al., 2006), mice (Isles et al., 2003, 2004), or monkeys (Hwang et al., 2006; Louie and Glimcher, 2006) show the predicted perseverance on the large reward alternative, but instead reverse their preference to the small, immediate reward once the large reward delay exceeds an individual threshold limit. This shows that the animals' choices depended on the waiting time preceding the rewards, but not on the ratios of reward amount and duration between the rewards, as would be predicted from rate maximization.

In fact, rate maximization models predict that amount and/or delay variations shouldn't play any role in the animals' decisions, because the choices should be only and exclusively directed towards maximizing the rate on the long-term. If, for example, an animal chooses between a fixed medium-term reward and or a variable-delay reward with either short or long delays (variable interval schedule), animals should always choose the option yielding the higher average reward rate. If the average reward rate is identical, animal should be indifferent between both options. However, contrary to this prediction, they usually prefer variable-interval over fixed schedules (Kacelnik and Bateson, 1996), indicating that delay variance does influence an animal's reward preference in addition to other factors, such as reward rate. This variance-proneness is interesting as animals are usually variance-averse if reward magnitude, and not delay, is variable, as explained above. Proneness to delay variance can be explained with hyperbolic discounting (see below).

In summary, animals do not make their choices according to the predictions of rate maximization models. They seem to employ rather short-sighted, waiting-time sensitive choice heuristics, and have a preference for delay variability.

Preference for Delay Variability

Hyperbolic discounting, as outlined in figure 3B, can explain the preference for variable over fixed interval schedules. Since, due to the hyperbolic decay, the utility of short-term rewards is disproportionally higher than the utility of medium-term or delayed rewards, but the difference in utility of medium-term and delayed rewards is negligible, the average expected utility of short-term and delayed rewards (variable interval schedules) will exceed the expected utility of fixed medium-term rewards. Hence, animals should prefer variable over fixed delays.

An alternative hypothesis, scalar expectancy theory (SET), can account for both variance aversion when reward magnitude is variable, and variance proneness when reward delay is variable. SET refers to the subjective time and magnitude representation which is normal around the actual means, but as a consequence of Weber's law, has a constant coefficient of variation (ratio of standard deviation to mean). Thus, the combination of an early and a late distribution results in a positively skewed integral, which explains preference for variation in delay (Reboreda and Kacelnik, 1991). Evidence for or against either SET or the hyperbolic discounting account is equivocal (Kacelnik an Bateson, 1997; Bateson and Kacelnik, 1998) and needs to be further tested in the future.

Ecological Models of Inter-Temporal Decisions: Ecological Rationality

In addition to the unclear support, neither of these accounts can explain why animals developed delay sensitivity in the first place: Why does evolution favor choice heuristics that produce suboptimal results in many cases by over-emphasising the delay to the next reward (e.g., through hyperbolic discounting), and ignoring the long-term relevance of time/amount sequences? Obviously, animals may equate delay with collection risk, as outlined in greater detail below. If delays are mentally treated as risks, a risk-averse animal will naturally avoid long delays. However, this doesn't provide an acceptable answer because the question remains why evolution has favoured suboptimal decision rules, be they related to risk avoidance or delay aversion. The first answer that comes to mind is that short-sighted rules have higher fitness values than long-sighted rules because the animals' constitutions do not allow them to tolerate too long waiting periods. For example, animals with a high metabolism cannot afford to wait too long for a large amount of food, or, in other words, what is the use of high quality, high amount of food if the animal has starved to death while waiting for it? Thus, short-sighted rules may have a certain evolutionary advantage over long-sighted rules.

This would certainly hold if the waiting times were close to the animals' starvation thresholds. However, mice, rats, pigeons, monkeys or other animals used in inter-temporal choice experiments shift their preference away from the economically more advantageous reward even when the waiting time to the larger reward exceeds less than a few seconds, and not hours or days (cf., McDiarmid and Rilling, 1969). Certainly, all those animals would be able to survive longer waiting periods than just a few seconds without food, but nevertheless, they prefer the short-term option over the long-delayed option, even if the long-delayed reward is a multiple of the short-term reward. Such extremely myopic decision patterns are difficult to explain with a fitness advantage of faster available food items. Why does evolution favor such extremely myopic choice heuristics?

Bounded rationality or ecological models, such as the ecological rationality hypothesis (Stephens et al., 2004) claim that choice heuristics that fail to produce maximum fitness in artificial experimental settings do, in fact, perform well in more ecologically valid contexts. For example, Stephens and colleagues (2004) argued that a more ecologically valid choice context entails decisions about limited food resources. A typical decision would consist of whether to entirely exploit all food resources in a given food patch, or leave the food patch early before having consumed all resources, and search for a new patch. The difference between the patch situation and the standard inter-temporal choice task is that, in a standard inter-temporal choice task, an animal has a binary choice between a large, delayed or a small,

immediate reward, whereas in the patch situation, it chooses whether to continue to stay in a given patch, or to leave and search for a new patch. Figure 4 illustrates an inter-temporal choice situation, often also referred to as a 'self-control' task (4A), and a patch situation (4B).

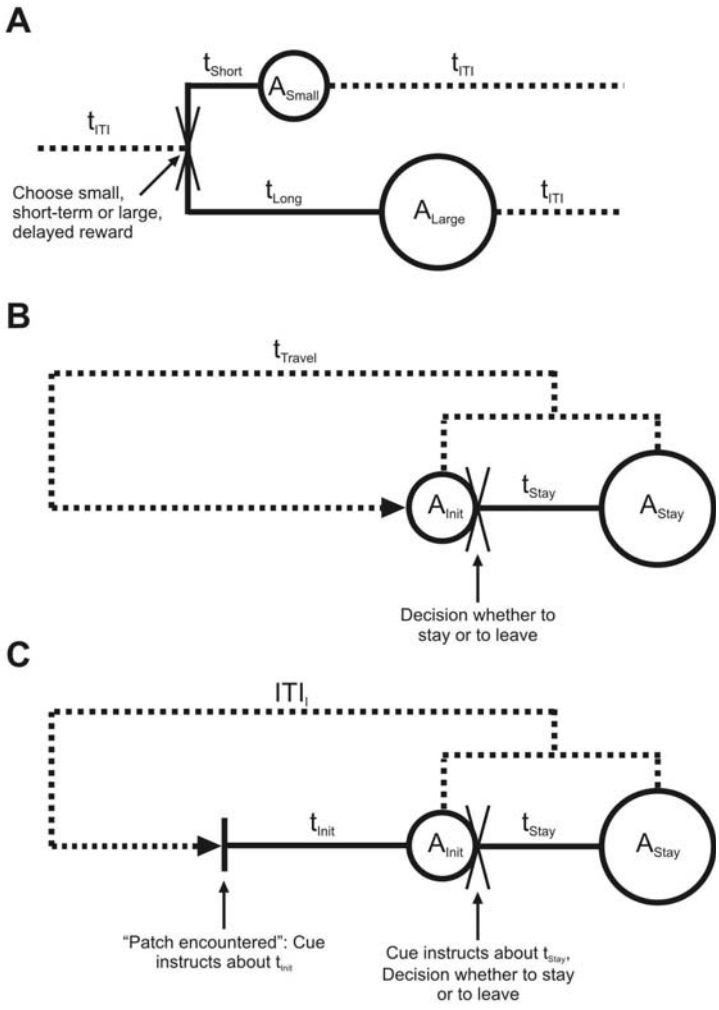


Figure 4. Inter-temporal choice and patch situation. (A) displays a schematic drawing of a standard inter-temporal choice task used in animal research, often also referred to as a 'self-control' task. Following an inter-trial interval (tITI), the animal makes a binary choice between a large and a small reward amount (ALarge and ASmall), delivered after a long or a short time delay (tLong and tShort). Reward consumption is followed by another ITI until the next choice opportunity. The length of the ITI, tITI, is adjusted to compensate for differences in delay length and choice, so that every trial is of identical duration. (B) Patch situation. An animal travels until it encounters a food patch, where it consumes an initial reward amount, AInit. After consumption, the animal has to decide whether to leave the patch and initiate travel time tTravel, until it finds the next patch, or whether to stay in the patch, wait time tStay for an additional food amount, AStay, and then leave the patch and initiate the travel time tTravel to the next food patch. (C) Patch situation as used in the experiment by Stephens and Anderson (2001). The situation is equivalent to figure 1B, but includes an additional initial waiting time, tInit, preceding the initial reward AInit. In all panels, the cross of thin lines indicates the timepoint where the animal makes its decision.

Put in more formalized terms, an animal consuming a reward in a food patch (the initial food amount, A_{Init}) has to decide whether to stay in the current patch and wait for further rewards until the patch is completely depleted, or whether to leave early and initiate a new travel time to the next patch. If it decides to leave, it has to travel for time t_{Travel} until the next patch is encountered where it receives a new initial reward amount A_{Init}. If it decides to stay, it obtains additional food rewards of amount A_{Stay} delivered after a certain waiting time t_{Stay} until the patch is depleted. It then has to leave the patch as well, and initiate a further travelling time t_{Travel} until it encounters a new patch and obtains reward amount A_{Init} in the new patch (see figure 4B).

Staying is more time-consuming ($t_{Stay}^{Total} = t_{Stay} + t_{Travel}$), but yields higher reward amounts ($A_{Stay}^{Total} = A_{Stay} + A_{Init}$), leaving is less time-consuming ($t_{Leave}^{Total} = t_{Travel}$) because the animal doesn't need to wait for the additional reward in the old patch, but it also misses out on that additional reward, and thus receives lower reward quantities ($A_{Leave}^{Total} = A_{Init}$). A far-sighted decision rule based on rate maximization would predict that the animal prefers to leave if it gained more rewards per time unit in the leave alternative than in the stay alternative, and it would stay in case it gained more rewards per time in the stay alternative than in the leave alternative. Assuming that travelling to a new patch takes always the same time, and that the yield in a new patch is always identical, an animal would leave if

$$\frac{A_{Init}}{t_{Travel}} > \frac{(A_{Add} + A_{Init})}{(t_{Stay} + t_{Travel})} \tag{9}$$

and it would predict to stay in the opposite case.

A short-sighted, waiting-time sensitive rule, as observed in most experimental settings on inter-temporal decision-making, would predict that an animal considers only the delay until the next reward in its decision. That is, the rule would predict that an animal prefers to leave if

$$\frac{A_{Init}}{t_{Travel}} > \frac{A_{Add}}{t_{Stay}} \tag{10}$$

and it would predict to stay in the opposite case. Because travel time, t_{Travel}, and initial reward amount, A_{Init}, are identical in the leave and the stay case, the long-term, rate-maximising rule in expression (9) is algebraically equivalent to the short-term, waiting-time-sensitive, impulsive rule in expression (10). Thus, in the above described patch situation, a short-sighted impulsive choice rule (expression 10) would approximate long-term rate maximization, consistent with predictions from optimal foraging theory.

If this theoretical line of argument was true, then the same short-sighted choice heuristic should produce rate maximization in the patch-situation, but not in the standard inter-temporal choice task. To test this prediction, Stephens and Anderson (2001) trained blue jays in a 'self-control' situation and a patch situation. The 'self-control' situation was essentially equivalent to the inter-temporal choice task sketched in figure 4A: Blue jays chose between a small, immediate or a large, delayed reward, A_{Small} and A_{Large}, delivered after t_{Short} and t_{Long}, by hopping on a perch on the left or right side in a training box. After reward delivery and consumption, they had to leave the perch, and an inter-trial interval (ITI, equivalent to travel time in figure 4B) was initiated, after which they could make their next choice. Instead of a binary choice between large and small rewards, the patch situation (figure 4C) consists of a sequence of choices between smaller rewards potentially summing up to a large reward: a cue instructed the animals about the delay t_{Init} to an initial reward of A_{Init} amount. By hopping on a randomly activated perch in the box, the initial reward was delivered after the indicated delay. Afterwards, the animals could choose whether to stay and wait for the additional reward by remaining on the perch (a second cue indicated the delay length t_{Stay} to the additional reward of amount A_{Stay}.), or whether to leave the perch, miss out on the additional reward, and initiate the ITI, and a new waiting time t_{Init} to the next small reward. Hence, the patch task resembled the situation illustrated in figure 4B, with the exception that there was an additional waiting time t_{Init} preceding the initial food reward A_{Init}. Moreover, in the patch situation, the travel time needed to leave the initial patch and move to the next one corresponds to the ITI of the standard inter-temporal choice task. Note that defined this way, the ITI becomes an integral part of the decision in the patch situation but not in the standard choice situation. This is because a short-term decision rule taking into account only the delay to the next reward would predict that, in the 'self-control' situation, animals consider the delays to the large and the small reward (t_{Short} versus t_{Long}), and that, in the patch situation, they consider the delays to the different rewards in the stay or leave cases (t_{Stay} vs. $t_{ITI} + t_{Init}$).

It is now possible to choose the task parameters so that the 'self-control' and the patch situation are economically equivalent: The sum of the two rewards in the patch situation equals the large reward amount in the 'self-control' situation, and the sum of the delays to the first and second rewards in the patch situation is equivalent to the delay preceding the large reward in the 'self-control' situation. This means that animals would receive the same amount of food within the same time when they chose to stay in the patch situation or when they chose the large, delayed reward in the 'self-control' situation

Because of this economical equivalence, the animals should always show identical preferences across both situations if their decision rule was purely economical, e.g., far-sighted. If, on the other hand, the animals indeed used a short-sighted decision rule, then they should show inconsistent choices under most circumstances, and should maximize their intake rate in the patch, but not in the 'self-control' situation. Such inconsistencies would arise because, first, the ITI is an integral part of the decision in the patch situation (remember that in the leave case, the total waiting time for the next reward would be $t_{Init}^{Leave} + ITI$), but not in the 'self-control' situation (the delay between choice and reward-delivery is not dependent on the ITI), and, second, because the large reward in the patch situation consists of a series of smaller rewards, and not of a one-shot delivery of one single large reward as in the 'self-control' situation.

In their study, Stephens and Anderson (2001) systematically varied the delays, amounts and ITIs, and tested the preference patterns of their blue jays. As predicted, they reported inconsistencies in choice between the two situations. In particular, they found that the blue jays were overall more ready to maximize their intake rate in the patch situation than in the 'self-control' situation. This supports the main notion of the ecological rationality hypothesis that one and the same short-sighted decision rule results in rate maximization in one, but not the other choice situation.

In summary, theoretical considerations and empirical evidence suggests that evolution may have favored the development of short-sighted choice heuristics because such rules produce long-term rate maximization in many natural situations, but not necessarily in artificial laboratory settings.

Ecological Models of Inter-Temporal Decisions: Feeding Ecology

Although all animals have in common that they discount the future, the rate by which future events are discounted differs dramatically between species. Mice, for example, seem to tolerate waiting times up to only a few seconds (Isles et al., 2003, 2004), capuchin monkeys wait for several minutes (Ramseyer et al., 2006), and humans can wait for months or even years for a relatively attractive reward (Green et al., 1994, 1997). The ecological rationality hypothesis can explain why evolution has favored the development of impulsive decision rules, but it cannot readily account for these large inter-individual and inter-species differences in delay-tolerance.

Another theory, the feeding ecology hypothesis, aims to explain those differences. It departs from the comparison of discount rates between very similar monkey species: Cotton-top tamarins and common marmosets. These two new-world monkeys are similar in terms of social behavior, mating system, life span, life history, home range size, parental care, body size and weight, brain size and weight, and other factors (cf., Stevens et al., 2005a). However, despite those similarities, the animals show very different choice behavior when tested in an adjusting delay procedure. In an adjusting delay procedure (Mazur, 1987), animals choose between a small, short-term reward and a large, delayed reward. After large reward choices, the delay to its receipt is increased in the next trials, after small reward choices, the delay preceding the large reward is decreased. This procedure allows the measurement of indifference points, i.e., the delay length at which the large, delayed reward has equal value to the small, short-term reward. Stevens and colleagues (2005a) found that the marmosets waited considerably longer for a large reward than tamarins. However, when tested in a spatial version of the adjusting feature task, in which travelling distance, but not delay, to a large reward was varied, the pattern reversed: Whereas the tamarins preferred the large reward independent of the travel distance to the large reward, the marmosets preference for the large reward continuously decreased with increasing travel distance (Stevens et al., 2005b). Taken together, space and time affected the monkeys' decisions differentially: compared to tamarins, marmosets were more patient when waiting for a delayed reward, but discounted spatially distant rewards steeper than tamarins.

How come that the two monkeys have evolved so different discounting patterns? Because of their striking similarity in many aspects, differences in metabolism, physical condition, starvation threshold or the like can be ruled out as possible explanations. Stevens et al.

(2005a, 2005b) point out that one of the main differences between the two New World monkeys is their diet: Although both species eat fruits, marmosets additionally feed on plant exudates, such as gum and sap, and tamarins feed on insects.

Feeding ecology plays a major role in shaping cognitive and neural functions (e.g., Clayton and Krebs, 1995; Basil et al., 1996; Emery and Clayton, 2001). Accordingly, the differences in foraging behavior between the tamarins and marmosets may prove to be the key evolutionary pressure for the differential development of the temporal and spatial discount rates: Marmosets feed on localized, immobile food sources that do not require far-distance travels. But feeding on gum and sap requires to scratch the bark of the tree and then wait for the sap to exude. For the marmosets, it is therefore essential to be patient (in time), but not necessarily mobile (in space) in order to get most of the slowly exudating sap. Conversely, for the insectivore tamarins, it is crucial to be constantly alert, and react quickly in order not to miss any passing-by insects. Moreover, they feed on dispersed food sources and have to cover rather large territories to find insects. Therefore, in contrast to marmosets, tamarins must be quick and impulsive, and willing to travel relatively far distances to find enough food to survive.

In summary, given the individual differences in foraging behavior, it may be more advantageous for the marmosets to be patient in time, but impulsive in space, and for the tamarins to be impulsive in time, but patient in space. The individual differences in foraging ecology may therefore explain the differential evolution of temporal and spatial discounting.

Commonalities and Differences Between Risky and Inter-Temporal Decisions

The preceding parts of this chapter have treated inter-temporal and risky decisions as separate. However, several authors argued that there might not be a real difference between delay and risk because each dimension can be expressed in terms of the other (Mischel, 1966; Stevenson, 1986; Rachlin et al., 1986, 1987, 1991; Mazur, 1989, 1995, 1997; Green and Myerson, 1996, 2004; Sozou, 1998; Hayden and Platt, 2007). A delayed reward might be less likely to occur (at least in natural situations) and therefore its expected value might be lower than that of earlier rewards. Moreover, as the state of the agent might change, the value of a later reward might also be more uncertain due to the unpredictability of the subject's own state, including its own survival. For example, our annual mortality risk is about 1% but was considerably higher in our evolutionary past. There is little use in waiting for a large, delayed reward if we will never experience the reward. Therefore a delayed reward may be equivalent to a risky reward and decision-makers may equate temporal distance with collection-risk (Kacelnik and Bateson, 1996; Sozou, 1998).

Theoretically, the proposal that delay and risk are processed similarly boils down to models that incorporate only either risk attitude or discounting as subjective weighting factors for utility but not both. For example, consider the following model (Mazur, 2007):

$$U = \Sigma P_i(A/1+KD_i) \tag{11}$$

Here, utility is a function of objective probability (*P*) but subjective time discounting (*K*). *A* denotes reward amount and *D* delay. Conversely, the following model comprises three subjective weighting factors determining utility (Kheramin et al., 2003, adapted):

$$U = \Sigma(1/(1 + Q/q_i) \times 1/(1 + Kd_i) \times 1/(1 + H\theta_i)) \quad (12)$$

where *q* denotes reward amount, *d* reward delay and θ odds against reward ($\theta=[1/P_i]-1$). The subjective weighting factors *Q*, *K* and *H* denote subjective sensitivity to reward magnitude, delay and probability, respectively. H>1 corresponds to risk aversion, H<1 to risk seeking, Note, that this model treats probability and delay similarly by using a hyperbolic form for both.

Conversely, instead of treating delayed rewards as uncertain, it has also been proposed that risky rewards may be treated as variably delayed rewards (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007). For example, take a gamble between two options with equal expected values, but where one option yields a medium-sized, certain reward and the other one yields either a large or a small reward with a 50% chance each. Animals are usually not risk-neutral in those types of tasks (Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998; McCoy and Platt, 2005b; Hayden and Platt, 2007). Macaques, for example, generally prefer the risky over the certain option (McCoy and Platt, 2005b; Hayden and Platt, 2007). To explain this risk-proneness, it has been argued that, if an animal consistently sticks with the risky option offering a 50% chance of a large pay-off, they will almost certainly receive the large reward eventually: If not now, then on a future trial. Therefore, the risky option gives a practically guaranteed, though potentially delayed large pay-off. Thus, because probabilistic rewards may be treated as large and delayed rewards, they may recruit similar cognitive mechanisms (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007).

The empirical evidence supporting commonalities of delay and risk sensitivity is equivocal (Mazur, 1989; Rachlin et al., 1986, 1991; Green and Myerson, 1996, 2004; Estle et al., 2006; Hayden and Platt, 2007). Commonalities are entertained by the occurrence of similar preference reversals when the delay or the probability of reward is increased for both options. Thus, violations of the independence axiom (Allais paradox) with probability are similar to violations of the stationarity axiom with delay (both described above). In other words, the utility of sooner and more probable rewards increases more than that of later and less probable rewards as reward immediacy and probability increase (reviewed in Green and Myerson, 2004). Accordingly, both probability and delay are amenable to hyperbolic discounting functions.

Conversely, there is considerable evidence that risk and delay are processed differentially. For example, humans discount smaller delayed rewards more steeply than larger delayed rewards but discount smaller probabilistic rewards less steeply than larger probabilistic rewards (Du et al., 2002; Estle et al., 2006). Inflation affects decisions involving delayed but not risky monetary rewards (Ostaszewski et al., 1998). Drug addiction appears to affect delay discounting more than risk processing (Reynolds et al., 2004) whereas problem gambling might have the opposite effect (Holt et al., 2003). Even culture appears to influence probability and delay differentially, with Japanese graduate students discounting probabilistic rewards more steeply and delayed rewards less steeply than Chinese students (Du et al., 2002). Further reinforcing the notion that delay differs from risk, earlier rewards may be

preferred for several reasons over later rewards over and above to the later reward being riskier or the future subjective state more uncertain (Kacelnik and Bateson, 1996):

- Earlier rewards can be put to use and earn compound interest (corresponding to offspring's offspring) before later rewards arrive.
- Waiting for a delayed reward may prevent an agent from pursuing other courses of action. This might diminish the value of delayed rewards.
- It might be easier to learn about action-reward and stimulus-reward contingencies with earlier rewards because at their arrival the mental representation of the causally relevant antecedents has decayed to a lesser degree compared with later rewards.
- With long delays to the later reward and fixed intertrial duration, choosing earlier rewards can maximize the energy intake per unit time even if the later reward is larger than the earlier reward (as explained above on the chapter on rate maximization).

CONCLUSION

In the preceding sub-chapters, we have outlined a selection of different attempts from different disciplines to explain decisions under risk and inter-temporal choices. Broadly speaking, normative approaches, such as EUT or optimal foraging (rate maximization), focus on how an animal should behave in order to meet formulated choice criteria, such as optimal decision-making, utility maximization or consistency of choice. Descriptive and empirical approaches, on the other hand, challenge many of the predictions and implications of normative models, and show that the normative analysis of decision making may not always be consistent with the empirical reality of choosing and acting. For example, humans do not always choose according to the predictions of EUT, and, contrary to the prediction of optimal foraging theory, animals frequently fail to maximize their intake rate. Ecological models, such as the budget rule or the ecological rationality hypothesis, deal with the question why evolution favored the development of choice patterns that often violate the predictions of the normative approaches. In particular, they provide an analysis within an ecologically valid framework of the sense and non-sense of the way animals (and humans) make their decisions. Last but not least, psychological approaches attempt to identify the choice-mediating cognitive mechanisms, such as whether animals employ far-sighted vs. myopic choice heuristics.

We hope to have illustrated that these different approaches have strong limitations when isolated from each other. In particular, we believe that normative models have high explanatory power, but are of questionable validity if not substantiated with empirical results. Empirical studies are of potentially higher validity per se, but it is difficult, if not impossible in many cases, to generalize single experimental results to a common framework of choice. They are therefore often of little use to answer the question how we generally make decisions when presented outside a theoretical context. Moreover, although empirical studies can be used to identify the short-comings of theoretical models, such de-construction is only useful when followed by the formulation of a better theoretical model. In conclusion, neither

normative, nor empirical, nor psychological approaches alone can produce useful results, but only the combination of all approaches yields fruitful outcomes.

ACKNOWLEDGEMENTS

We greatly appreciate the comments of Martin Kocher on an earlier version of the manuscript. TK is supported by a grant from SenterNovem (BSIK-03053).

REFERENCES

Ainslie G (1974) Impulse control in pigeons. *J. Exp. Anal. Behav.* 21:485-489.

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* 82:463-496.

Allais M (1953) Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica* 21:503-546.

Basil JA, Kamil AC, Balda RP, Fite KV (1996) Differences in hippocampal volume among food storing corvids. *Brain Behav. Evol.* 47:156-164.

Bateson M, Kacelnik A (1998) Risk-sensitive foraging: decision-making in variable environments. In: *Cognitive Ecology* (Dukas R, ed), pp 297-341. Chicago, Illinois: University of Chicago Press.

Bennett SM (2002) *Preference reversal and the estimation of indifference points using a fast-adjusting delay procedure with rats*. Dissertation, University of Florida.

Benzion U, Rapoport A, Yagil J (1989) Discount rates inferred from decisions: An experimental study. *Management Sci.* 35:270-284.

Bernoulli D (1954) Exposition of a new theory on the measurement of risk. *Econometrica* 22:23-36.

Caraco T, Blanckenhorn WU, Gregory GM, Newman JA, Recer GM, Zwicker SM (1990) Risk-sensitivity: Ambient temperature affects foraging choice. *Anim. Behav.* 39:338-345.

Cardinal RN, Robbins TW, Everitt BJ (2000) The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology* (Berl) 152:362-375.

Charnov EL (1976) Optimal foraging: The marginal value theorem. *Theor. Popul. Biol.* 9:129-136.

Chung SH, Herrnstein RJ (1967) Choice and delay of reinforcement. *J. Exp. Anal. Behav.* 10:67-74.

Clayton NS, Krebs JR (1995) Memory in food-storing birds: from behaviour to brain. *Curr. Opin. Neurobiol.* 5:149-154.

Du W, Green L, Myerson J (2002) Cross-cultural comparisons of discounting delayed and probabilistic rewards. *Psychol. Rec.* 52:479–492.

Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* 75:643-669.

Emery NJ, Clayton NS (2001) Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414:443-446.

Estle SJ, Green L, Myerson J, Holt DD (2006) Differential effects of amount on temporal and probability discounting of gains and losses. *Mem. Cognit.* 34:914-928.

Evenden JL (1999) Varieties of impulsivity. *Psychopharmacology* (Berl) 146:348-361.

Evenden JL, Ryan CN (1996) The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* (Berl) 128:161-170.

Fehr E (2002) The economics of impatience. *Nature* 415:269-272.

Fishburn PC (1982) Nontransitive measurable utility. *J. Math. Psychol.* 26:31-67.

Fishburn PC, Rubinstein A (1982) Time Preference. *Internat Econ. Rev.* 23:677-694.

Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: a critical review. *J. Econ. Lit.* 40:351-401.

Friedman M, Savage LJ (1948) The utility analysis of choices involving risk. *J. Polit. Econ.* 56:279-304.

Glimcher PW (2002) Decisions, decisions, decisions: choosing a biological science of choice. *Neuron* 36:323-332.

Glimcher PW, Rustichini A (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306:447-452.

Green L, Fisher EB Jr, Perlow S, Sherman L (1981) Preference reversal and self-control: Choice as a function of reward amount and delay. *Behav. Anal. Lett.* 1:43-51.

Green L, Fristoe N, Myerson J (1994) Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonom. Bull. Rev.* 1:383-389.

Green L, Myerson J (1996) Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *Amer. Zool.* 36:496-505.

Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130:769-792.

Green L, Myerson J, McFadden E (1997) Rate of temporal discounting decreases with amount of reward. *Mem. Cognit.* 25:715-723.

Grossbard CL, Mazur JE (1986) A comparison of delays and ratio requirements in self-control choice. *J. Exp. Anal. Behav.* 45:305-315.

Hamilton W D (1964) The genetical evolution of social behaviour. *J. Theor. Biol.* 7:1–52.

Hayden BY, Platt ML (2007) Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17:49-53.

Holt DD, Green L, Myerson J (2003) Is discounting impulsive? Evidence from temporal and probability discounting in gambling and nongambling college students. *Behav. Processes* 64:355–367.

Hwang J, Kim S, Lee D (2006) Neuronal signals related to delayed reward and its discounted value in the macaque dorsolateral prefrontal cortex. *Annual Meeting of the Society for Neuroscience*, Program No. 71.7.

Isles AR, Humby T, Walters E, Wilkinson LS (2004) Common genetic effects on variation in impulsivity and activity in mice. *J. Neurosci.* 24:6733-6740.

Isles AR, Humby T, Wilkinson LS (2003) Measuring impulsivity in mice using a novel operant delayed reinforcement task: effects of behavioural manipulations and d-amphetamine. *Psychopharmacology* (Berl) 170:376-382.

Ito M, Asaki K (1982) Choice behavior of rats in a concurrent-chains schedule: amount and delay of reinforcement. *J. Exp. Anal. Behav.* 37:383-392.

Jones, B. and Rachlin, H. (2006) Social discounting. *Psychol. Sci.* 17, 283-286.

Kacelnik A (1997) Normative and descriptive models of decision making: time discounting and risk sensitivity. *Ciba Found Symp* 208:51-67; discussion 67-70.

Kacelnik A, Bateson M (1996) Risky theories - the effects of variance on foraging decisions. *Amer. Zool.* 36:402-434.

Kacelnik A, Bateson M (1997) Risk-sensitivity: crossroads for theories of decision-making. *Trends Cogn. Sci*. 1:304-309.

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263-291.

Kalenscher T, Ohmann T, Gunturkun O (2006) The neuroscience of impulsive and self-controlled decisions. *Int. J. Psychophysiol.* 62:203-211.

Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Curr. Biol.* 15:594-602.

Kalenscher T, Pennartz CMA (2007) A bird in the hand is worth two in the future. *The neuroeconomics of inter-temporal decision making*. In preparation.

Kheramin S, Body S, Ho M, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2003) Role of the orbital prefrontal cortex in choice between delayed and uncertain reinforcers: a quantitative analysis. *Behav. Processes.* 64:239-250.

Kirby KN, Herrnstein RJ (1995) Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci*. 6:83-89.

Koopmans TC (1960) Stationary ordinal utility and impatience. *Econometrica* 28:287-309.

Laibson DI (1997) Golden eggs and hyperbolic discounting. *Q. J. Econ.* 42: 861-871.

Lancaster K (1963) An axiomatic theory of consumer time preference. *Internat. Econ. Rev.* 4:221-231.

Loewenstein G (1992) The fall and rise of psychological explanations in the economics of intertemporal choice. In: *Choice over time*, pp. 3-34. Eds G. Loewenstein, J. Elster. Russell Sage Foundation: New York.

Logue AW (1988) Research on self-control: An integrating framework. *Behav. Brain. Sci.* 11:665-709.

Loomes G, Starmer C, Sugden R (1991) Observing violations of transitivity by experimental methods. *Econometrica* 59:425-439.

Loomes G, Sugden R, (1982) Regret theory: An alternative theory of rational choice under uncertainty. *Econ. J.* 92:805-824.

Louie K, Glimcher PW (2006) Temporal discounting activity in monkey parietal neurons during intertemporal choice. *Annual Meeting of the Society for Neuroscience*, Program No. 605.5.

Machina MJ (1982) "Expected utility" analysis without the independence axiom. *Econometrica* 50:277-323.

Mankiw, NG (2002) *Macroeconomics*. 5th edition, Worth Publishers: New York.

Mazur JE (1984) Tests of an equivalence rule for fixed and variable reinforcer delays. *J. Exp. Psychol*: *Anim. Behav. Process* 10:426-436.

Mazur JE (1988) Estimation of indifference points with an adjusting-delay procedure. *J. Exp. Anal. Behav*. 49:37-47.

Mazur JE (1989) Theories of probabilistic reinforcement. *J. Exp. Anal. Behav*. 51:87-99.

Mazur JE (1995) Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *J. Exp. Anal. Behav.* 63:139–150.

Mazur JE (1997) Choice, delay, probability, and conditioned reinforcement. *Anim. Learn. Behav.* 25:131–147.

Mazur JE (2007) Species differences between rats and pigeons in choices with probabilistic and delayed reinforcers. *Behav. Processes* 75:220-224.

Mazur JE, Commons ML, Mazur JE, Nevin JA, Rachlin H (1987) *An adjusting procedure for studying delayed reinforcement*. In, pp 55-73. Hillsdale, NJ: Erlbaum.

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503-507.

McCoy AN, Platt ML (2005a) Expectations and outcomes: decision-making in the primate brain. *J. Comp. Physiol. A Neuroethol. Sens Neural. Behav. Physiol*. 191:201-211.

McCoy AN, Platt ML (2005b) Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci*. 8:1220-1227.

McDiarmid CG, Rilling ME (1965) Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychon. Sci*. 2:195-196.

McNamara JM, Houston AI (1986) The common currency for behavioral decisions. *Am. Nat.* 127:358-378.

Mischel W (1966). Theory and research on the antecedents of self-imposed delay of reward. In BA Maher (Ed.), *Progress in experimental personality research*: Vol. 3 (pp. 85–132). New York: Academic Press.

Moore FR, Simm PA (1986) Risk-sensitive foraging by a migratory bird (Dendroica coronata). *Experientia* 42:1054-1056.

Myerson J, Green L, (1995) Discounting of delayed rewards: Models of individual choice. *J. Exp Anal. Behav*. 64, 263-276.

Ostaszewski P, Green L, Myerson J (1998) Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychon. Bull. Rev.* 5:324–333.

Powell K (2003) *Economy of the mind*. PLoS Biol 1:E77.

Rachlin H, Castrogiovanni A, Cross D (1987) Probability and delay in commitment. *J. Exp. Anal. Behav*. 48:347–353.

Rachlin H, Green L (1972) Commitment, choice and self control. *J. Exp. Anal. Behav*. 17:15-22.

Rachlin H, Logue AW, Gibbon J, Frankel M (1986) Cognition and behavior in studies of choice. *Psychol. Rev.* 93:33-45.

Rachlin H, Raineri A, Cross D (1991) Subjective probability and delay. *J. Exp. Anal. Behav*. 55:233-244.

Ramseyer A, Pele M, Dufour V, Chauvin C, Thierry B (2006) Accepting loss: the temporal limits of reciprocity in brown capuchin monkeys. *Proc. Biol. Sci.* 273:179-184.

Reboreda JC, Kacelnik A (1991) Risk sensitivity in starlings: Variability in food amount and food delay. *Behav. Ecol.* 2:301-308.

Reynolds B, de Wit H, Richards J (2002) Delay of gratification and delay discounting in rats. *Behav. Processes* 59:157-168.

Reynolds B, Richards JB, Horn K, Karraker K (2004) Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav. Processes*, 65:35–42.

Rohde K (2005) The hyperbolic factor: a measure of decreasing impatience. Research memoranda 044. *Maastricht Research School of Economics of Technology and Organization: Maastricht*.

Samuelson PA (1937) A note on measurement of utility. *Rev. Econ. Stud.* 4:155-161.

Shafir S (1994) Intransitivity of preferences in honey bees: support for 'comparative' evaluation of foraging options. *Anim. Behav*. 48:55–67.

Sibly RM, McFarland DJ (1976) On the fitness of behaviour sequences. *Am. Nat.* 110:601-617.

Sozou PD (1998) On hyperbolic discounting and uncertain hazard rates. *Proc. R. Soc. Lond. B*. 265:2015-2020.

Stephens DW (1981) The logic of risk-sensitive foraging preferences. *Anim. Behav*. 29:628-629.

Stephens DW, Anderson D (2001) The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. *Behav. Ecol.* 12:330-339.

Stephens DW, Kerr B, Fernandez-Juricic E (2004) Impulsiveness without discounting: the ecological rationality hypothesis. *Proc. R. Soc. Lond B. Biol. Sci.* 271:2459-2465.

Stephens DW, Krebs JR (1986) Foraging theory. *Monographs in behavior and ecology*. Princeton. Princeton University Press: New Jersey.

Stevens JR, Hallinan EV, Hauser M (2004) The ecology and evolution of patience in two New World monkeys. *Biol. Lett.* 1:223-226.

Stevens JR, Rosati AG, Ross KR, Hauser MD (2005) Will travel for food: Spatial discounting in two new world monkeys. *Curr. Biol.* 15:1855-1860.

Stevenson MK (1986) A discounting model for decisions with delayed positive and negative outcomes. *J. Exp. Psychol: General* 115:131–154.

Strotz RH (1955) Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23:165-180.

Tobler PN, Fletcher PC, Bullmore ET, Schultz W (2007a) Learning-related human brain activations reflecting individual finances. *Neuron*. 54:167-175.

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2007b) Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97:1621-1632.

Tversky A (1969) Intransitivity of preferences. *Psych. Rev.* 76: 31-48.

Tversky A, Kahneman D (1986) Rational choice and the framing of decisions. *J. Business* 59:S251-S278.

Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk Uncert.* 5:297–323.

Von Neumann J, Morgenstern O (1944) *Theory of games and economic behavior*. Princeton. Princeton University Press: New Jersey.

Waite TA (2001) Intransitive preferences in hoarding gray jays (Perisoreus canadensis). *Behav. Ecol. Sociobiol.* 50:116-121.

Winstanley CA, Theobald DE, Cardinal RN, Robbins TW (2004) Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci*. 24:4718-4722.

Winstanley CA, Theobald DE, Dalley JW, Cardinal RN, Robbins TW (2006) Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb. Cortex* 16:106-114.

Wolf M, van Doorn GS, Leimar O, Weissing FJ (2007) Life-history trade-offs favour the evolution of animal personalities. *Nature* 447:581-584.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 137-161

Chapter 6

THE EFFECTS OF PERCEPTIONS ON MODE CHOICE

Maria Vredin Johansson
Department of Economics, Uppsala University
Box 513, SE-751 20 Uppsala, Sweden

ABSTRACT

While remaining within the traditional micro-economic framework of rational utility maximization, we enrich the standard and random parameters logit choice models with perceptions data. From the estimated models we derive a value of time and we also make a tentative attempt to derive a value of safety. Because we estimate the values simultaneously, we are able to explore whether values estimated in conjunction differ from values estimated in isolation. Survey data is used to measure the individual's perceptions of five modal attributes (time, cost, safety/risk, environmental friendliness and flexibility) and show how these perceptions affect the modal choice for work trips. The respondents' perceptions are elicited by a novel approach in which the names of two modes (car and bus) are used as attribute levels instead of objective levels. A difference between our survey and traditional ones is that we do not attempt to educate the respondents about, for example, the risks of travelling. Instead, we *record* the respondent's perceptions about the risk and the other the modal attributes.

Keywords: *Value of Time, Value of Statistical Life, Random Parameters Logit, Stated Preferences.*

1. INTRODUCTION

In the empirical literature on travel mode choice, most choice models use objective modal attributes and individual characteristics as explanatory variables (cf. Algers et al., 1995). Nonetheless, few people reject the thought of travel mode choice also depending on harder-to-measure, qualitative variables, such as comfort and convenience. It has even been

suggested that the choice of travel mode can depend on the perceptions[1] and images of the alternatives.

For instance, Koppelman and Pas (1980) argue that, as models with only demographic and objective attribute levels only provide a limited understanding of the behavioural processes underlying choice, incorporation of perceptions and feelings would improve the understanding of modal choices. Interestingly, Adamowicz et al. (1997) have in a study in the field of environmental economics found that a model based on perceptions slightly outperformed (i.e. had greater explanatory power) a model based on objective attribute measures.

Moreover, as it is recognised that objective information about a travel mode can be quite different from the perception of the information (Lichtenstein et al., 1978; Lee, 1981; Slovic et al.,1981; Mowen, 1990, ch. 2), the idea that objectively measured attributes are the only influential variables in travel choice models is open to question. If people respond to what they perceive, i.e. make choices on basis of their perceptions, the perceptions about the modal attributes are the relevant response variables to which people react, not the objective attribute levels. Thus, reconsideration when it comes to explaining travel choices seems both warranted and welcome.

In this paper, we use a mail survey to measure the individual's perceptions of five modal attributes (time, cost, safety, environmental friendliness and flexibility) and see how these perceptions affect modal choice for work trips. A difference between our survey and traditional ones, is that we do not attempt to inform, or educate, people about, for instance, the costs or risks of travelling. Instead, we *record* the respondents' perceptions about the modal attributes.

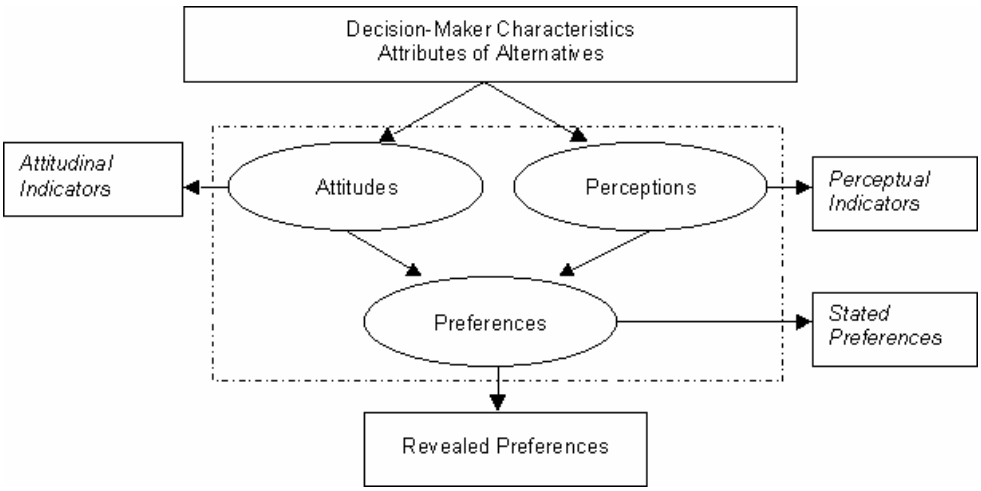


Figure 1. The consumer decision process (adapted from Morikawa et al., 1990, p. 3).

The work in this paper is based on the paradigm of individual choice behaviour depicted in Figure 1. In the figure latent variables are depicted by ovals and observable variables by boxes. Decision maker characteristics and objective attributes of alternatives are assumed to

[1] In this paper, perception is defined as an individual's subjectively formed opinions and ideas about a stimulus or a piece of information.

affect the *attitudes*, *perceptions* and *preferences* of the decision-maker.[2] These latent variables - all assumed to affect choice - constitute the ”black box” of the consumer. Because perceptions and attitudes are assumed to affect the individual’s preferences and, as the preferences are assumed to determine choice, it is possible that incorporation of perceptions in travel choice models will improve the models’ explanatory power.

While staying within the traditional microeconomic framework of rational utility maximisers (where preferences are expressed by utility functions), we enrich the standard and random parameters logit choice models by perceptions data. From the estimated models we derive values of time. We also make a tentative attempt to measure the monetary value of safety. Because the Swedish National Road Administration (SNRA) uses both a value of time savings and a value of safety (i.e. value of ”statistical life”) in their project evaluations, it is desirable to estimate both values in the same model. Presently the SNRA uses values derived from separate studies and, as far as we know, simultaneous estimation has never been performed in Sweden before. Because we estimate the values simultaneously, we are able to explore whether values estimated in conjunction differ from values estimated in isolation.

This paper has three objectives: (*i*) to elicit the individuals’ perceptions about the modal attributes time, cost, safety, environmental friendliness and flexibility; (*ii*) to use these perceptions to explain mode choice for work trips, and; (*iii*) to derive monetary values for time and safety in the same model estimation.

The remainder of the paper is organised as follows; the next section describes the data collection and the experimental design, section three gives a theoretical framework for stated preferences discrete choice models, section four presents the results from the empirical estimations and, finally, section five concludes and discusses the results.

2. THE SURVEY

In a mail survey, we focus on modal choice for work trips. There were several reasons for this limitation. First, work trips are easy to define for most people. Second, work trips are repeated journeys for which the nodes does not change unless the respondent has varying places of work. Third, work trips are frequently performed and people are, therefore, often well-informed about the alternative modes of travelling. Consequently the individuals’ preferences for different travel modes can be assumed to be fairly well known. Finally, according to Statistics Sweden (SCB, 1998) almost half of the trips in 1997 consisted of trips to and from work/school, a fact that makes work trips imperative to study.

In total, there were 22 questions in the questionnaire. Apart from socio-economic questions and questions regarding the respondent’s habitual mode of travelling to and from work (revealed preferences (RP) information), the questionnaire contained a question to elicit the respondent’s perceptions of modal attributes of two different travel modes (car and bus) and two stated preferences (SP) choice questions where the respondent was asked to choose the preferred of two hypothetical travel modes. Each travel mode was described by four *attributes*.

[2] *Attitudes* are defined as the decision-maker’s subjective importance of the attributes and *preferences* as the desirability of alternatives.

In December 1997, the questionnaire was mailed to 480 residents of Borlänge and Falu municipalities randomly selected[3] from the Swedish Official Register of Persons and Addresses (SPAR).[4] In the beginning of January 1998, a follow-up reminder was mailed to the non-respondents at that time. The overall response rate from the two mailings was 66.8 percent.

2.1. The Attributes

In the SP choice questions the hypothetical alternatives were described by five different modal attributes: time, cost, safety, environmental friendliness, and flexibility (see questions 20-21 in Appendix A).

Time and cost were included for obvious reasons. Safety was included for two reasons: first, there is evidence (Noland, 1995) that decreases in the perceptions of a travel mode's safety reduces the probability of that mode being chosen and, second, safety improvements are the second most important benefit (time savings are the most important) in the cost-benefit analyses of road investment projects made by the SNRA and, consequently, of considerable policy interest. The mode's environmental friendliness was added because we wished to analyse if the on-going debate on global warming has affected people's choice of travel mode.

Because we judged it unnecessary, and potentially too burdensome, for the respondents to evaluate five different attributes, we split the total sample in two so that each sub-sample consisted of 240 individuals, and let each sub-sample evaluate only four different attributes. Travel time, travel cost and safety were used in both sub-samples, while environmental friendliness and flexibility were used interchangeably. In this way we obtained an "environmental friendliness" sub-sample (I) and a "flexibility" sub-sample (II). Thus, the distinguishing feature between the two sub-samples is their fourth attribute (see Table 1).

Table 1. Attributes of the two sub-samples

Sub-sample I	Sub-sample II
Travel time (*t*)	Travel time (*t*)
Travel cost (*c*)	Travel cost (*c*)
Safety (*s*)	Safety (*s*)
Environmental friendliness (*e*)	Flexibility (*f*)
n_I=240	n_{II}=240

2.2. Attribute Levels

All the attributes vary in two levels, a "high" and a "low" level. For the factorial design, we construct a new, artificial, attribute, *ef*, which symbolises the environmental friendliness

[3] Because we wanted respondents who, with a larger probability compared to a completely random sample, were working, the respondents were required to be aged between 36 and 55 years.

(*e*) attribute in sub-sample I and the flexibility (*f*) attribute in sub-sample II. Having four, instead of five variables, reduces the complexity of the factorial design because the number of alternatives is reduced from 2^5 to 2^4.

The intrinsic natures of the attributes in this survey differ because the travel times and travel costs are quantitative variables for which objective (i.e. researcher defined) values are common, while safety, environmental friendliness and flexibility are, more or less, qualitative variables where objective values are rather uncommon. For these qualitative variables other measurement scales than objective must be used. For instance, a measure of the travel mode's objective risk level can be defined as the actual risk the individual faces during the specific trip expressed in historical number of annual fatalities. For the mode's environmental friendliness an objective measure would be the average emissions per passenger kilometre (Lenner, 1993). For the mode's flexibility it is harder to conceive of a good objective measure. Possible solutions are to use some kind of categorical measure, such as low, medium and high, or, to endogenise the measurement by defining the prevailing level as the one from which changes are made. If we were to use objective values on all attributes in the SP survey, the values of the qualitative attributes would have to be invented and communicated to the respondents.

We, however, circumvent the problem of constructing objective values for the qualitative variables in this survey by using the respondent's perceptions of these variables. As a matter of fact, in order to simplify and make the questionnaire consistent, we use the respondent's perceptions of *all* variables - even for those with objective values.[5] To elicit the respondents' perceptions of modal attributes, we propose a new approach, implemented by the four-step procedure described below. To the best of our knowledge, this approach has no previous application.

The first step is to use the *names* of two travel modes, car and bus, as levels for *all* attributes. Thus, an alternative could be described as having travel time "like car", travel cost "like bus", safety "like bus" and environmental friendliness or flexibility "like car". Consequently, one of the alternatives describes a travel mode with all attributes "like car", while another has all attributes "like bus". All other alternatives are mixtures of the "like car" and "like bus" attributes and, therefore, hypothetical constructs. To the best of our knowledge, this "like"-approach has only been used once before in an SP survey, by Louviere and Johnson (1991) who, in a marketing study, explored the retail images of different supermarkets.[6]

The second step is to decide whether car (bus) represents the high or the low level of the different attributes. Five assumptions about the attribute levels were used in the experimental design (see Table 2): (*i*) Going a specific distance by car is normally faster than going the same distance by bus, due to the stops the bus makes - consequently, the high level of travel time was designated bus and the low level car; (*ii*) The travel cost of car was considered

[4] Borlänge and Falu municipalities are situated approximately 200 kilometres north-east of Stockholm in the Dalecarlia region.

[5] An advantage of using perceptions is that we do not have to concentrate on a specific type of traveller for whom we know the modal attribute levels beforehand (it would be impossible to construct realistic alternatives unless we had information about the respondent's actual travel mode). By using perceptions we are able to survey respondents with a mail questionnaire.

[6] Louviere and Johnson (1991) used attribute levels in the "like" form (e.g. "like K-Mart") and measured the perceptions of the different supermarkets through ratings.

larger than the equivalent for bus, considering the additional costs of buying and running a car, rendering the car the high and the bus the low level; (*iii*) Based on historical statistics of traffic fatalities (SIKA, 1998; Kommunikationsdepartementet, 1997), bus safety was regarded superior to car safety and, therefore, safety on bus is high, while safety in car is low; (*iv*) Because the discharges per passenger kilometre from buses are lower than those from cars (Lenner, 1993) the environmental friendliness of bus is assumed to be higher than that of car; and (*v*) When assigning a high and a low level to the (admittedly ambiguous) notion of flexibility, we considered three dimensions of mode flexibility: departure times, walking distances and freedom in choice of route. Because cars, in addition to having a greater freedom in route choice, generally, has more flexible "departure times", and can often be parked in proximity of the travel destination, we regarded car as the more flexible travel mode (i.e. high level of flexibility).

The individual's preferences for the attributes are assumed to be monotone (increasing or decreasing). If the respondents perceive the attributes the way we do, lower travel cost and travel time are always preferred, as are higher safety, environmental friendliness and flexibility. Of course, the respondents may perceive the attributes differently, or make different assumptions about the attribute levels.[7] This should not be a weakness of the design, because the respondents make the choices based on their own preferences, not on *our* ideas about their perceptions.

The third step is measurement. Most desirable would be to measure modal perceptions on a quotient scale, i.e. an equidistantly graded scale with an absolute zero, so that comparisons could be made both within and between respondents.[8] To obtain quotient scaled data, we apply a method originating in measurement scales for comparative judgement in psychophysics (cf. Björkman and Ekman, 1957). In this method the respondents make pair-wise comparisons between a stimulus and a standard. The respondent's perception of the stimulus is frequently expressed as *a multiple* or *a fraction* of the standard, so that, for instance, the stimulus is either twice or half of the standard.

Table 2. Attribute levels

Attribute	High level (+)	Low level (-)
Travel time (*t*)	Bus	Car
Travel cost (*c*)	Car	Bus
Safety (*s*)	Bus	Car
Environmental friend. (*e*)	Bus	Car
Flexibility (*f*)	Car	Bus

[7] Gårder et al. (1994), for example, show that experts may judge on-street bicycle lanes as less safe (compared with no lanes at all) whereas the cyclists themselves perceive them as being safer.

[8] Ranking yields data with individual-specific scale units and origins (ordinal data), whereas rating yields equidistantly measured data with individual-specific origins (interval data). While ordinal data only describe the respondent's preferences (e.g. $X_1<X_2<X_3<X_4$), interval data also describe the magnitude of the difference in preferences (e.g. $X_4-X_2=2(X_2-X_1)$. Only data measured on a quotient scale can express the *intensity* of preferences (e.g. X_2 is twice as good as X_1, $X_2=2\ X_1$) (Björkman & Ekman, 1957).

In our survey we let the respondents compare bus (the stimulus) to car (the standard)[9] and express any difference perceived in percent. That is, while car is assumed to be ”100%” on each attribute, the respondents are asked to state how, compared to the car, they perceive the attributes of the bus (see question 19, Appendix A). For instance, if a respondent perceives the travel time of bus to be three times that of car, she responds ”300%” and, if she perceives the cost of bus to be a third of that of car, she responds ”33%”. Consequently, the perceptions obtained are either multiples or fractions of the car standard.

The fourth step is to infer real values, i.e. values in minutes and Swedish kronor (SEK)[10] , from the percentages, using the RP information provided by the respondent. For example, if the respondent travels to work by car, information about the actual time and cost by car can, in conjunction with the perceptions of the time and cost by bus, be used to give the ”like car” and the ”like bus” levels real values. A simple example: assume that the respondent travels to work by car and that it takes 10 minutes and costs SEK 20. The same respondent perceives the travel time of bus to be twice that of car and responds ”200%” to the relevant part of question 19. Similarly, the travel cost of bus is perceived to be half that of car and, therefore, she responds ”50%” to that part of the question. Using the RP information, i.e. $t_{auto} = 10$ and $c_{auto} = 20$, we can calculate the travel time of bus (t_{bus}) to 20 minutes and the travel cost (t_{bus}) to SEK 10.[11] Consequently, travel time ”like car” equals 10 minutes and travel time ”like bus” equals 20 minutes in the SP choice questions. Analogously, real values can be inferred for the cost attribute.

Equivalent transformations for the time and cost variables can easily be performed for respondents who ride a bus to work and, in fact, for all respondents who has either bus or car in their choice sets. However, both the actual mode’s time *and* cost must be known. If the respondent is unaware of an alternative’s cost or time, we judge the alternative unlikely for the respondent and do, therefore, not define it as a part of the respondent’s choice set.

For the qualitative variables, safety, environmental friendliness and flexibility, no such transformations can be made since we have no RP values for these variables. Consequently, both real values and ”pure” attribute perceptions are used to explain the SP choices in the estimation. The main advantage of inferring real values from perceptions is that we are able to calculate a value of time that is easy to interpret and compare with previous research findings.

2.3. Experimental Design

Because the survey was performed as a postal questionnaire, it was important to keep the questionnaire short and simple to obtain a high response rate. We, therefore, restricted the

[9] The choice of car as the standard is merely one of convenience, but because car is the most frequently used travel mode for work trips (SCB, 1998), we believe that most people find it easier to compare bus to car than the other way around.

[10] US$ 1 was, in 1997, on average equal to SEK 7.64. In 2006, US$ 1 was on average equal to SEK 7.38. Although there has been large variations in the average exchange rates over the years, the rate in 1997 was on level with the rate in 2006 (www.riksbank.se).

[11] General formulas for the calculations are: $t_{bus} = t_{auto} \cdot \tau / 100$ and $c_{bus} = c_{auto} \cdot \psi / 100$ where t_{bus} is travel time by bus, t_{car} travel time by car, c_{bus} travel cost by bus and c_{car} travel cost by car. τ and ψ are the perceptions of travel time and travel cost of bus compared to car expressed in percent.

size of the choice experiment, which is a function of the number of attributes to be varied. With n=16 we could construct as many as $n(n\text{-}1)/2$=120 different paired comparisons (full factorial).[12] As no respondent arguably could be required to evaluate 120 different pairs, we had to reduce the number of comparisons. By selecting the smallest orthogonal main effects plan from the full factorial through ”blocking” (see Appendix B), we were able to estimate non-confounded main (i.e. attribute specific) effects. In our survey, the smallest orthogonal main effects design consists of eight pairs. Because we believed that no respondent could be required to perform more than two paired comparisons, i.e. two different choice questions, the eight pairs were divided into four sets of two pairs each. This particular experimental design ensures non-confounded main attribute effects, but confounded (with the blocks) interaction effects. According to Batsell and Louviere (1991), interaction effects are rarely estimated in choice models and, unless the respondents in different blocks differ or behave differently, there will be no block effects.

With two paired comparisons in each questionnaire and a sub-sample size of 240, we got four ”sub-sub-samples”, A-D and E-H, respectively, consisting of 60 individuals each (see Table 3). Each individual was randomly assigned a treatment.

Because we kept the dominating and the dominated alternatives in the experiment, and as the pairs in the paired comparisons are *fold-over* pairs (which means that they are complementary as the signs of the attribute levels of the first alternative are reversed in the second, see Box et al. (1978) and Appendix B), we have one questionnaire in each sub-sample that contains the choice between the ”best” and the ”worst” alternatives.[13] In sub-sample I, questionnaire B contains the best-worst pair and in sub-sample II, it is questionnaire F. If these *designed* best-worst questions, in any way, are simpler to respond to, we expect greater response rates for these questionnaires. However, as can be seen in Table 3, this is obviously not the case since B, in fact, has the lowest response rate in sub-sample I and F has the second lowest response rate in sub-sample II.

Table 3. Response rates of the different sub-sub-samples

	Sub-sample I				Sub-sample II			
	A	B	C	D	E	F	G	H
Main sample	60	60	60	60	60	60	60	60
Responses	37	37	44	39	43	38	37	43
Response rate (%)	62.7	61.7	73.3	65.0	71.7	64.4	62.7	72.9

Note: The real sample is the main sample less the questionnaires that were.

[12] We do not compare an alternative with itself or consider the sequencing of alternatives.

[13] When used as a consistency check, approximately 22 percent in the environmental friendliness sub-sample and about 3 percent in the flexibility sub-sample chose the worst alternative in the best-worst question. However, what is right and what is wrong is based on the individual’s perceptions of the attributes. Therefore, we can not conclude that these respondents behaved inconsistently.

3. MODELLING CHOICES

The behavioural assumption underlying the model for travel choice used in this paper is maximisation of individual utility, i.e. the individual who, with certainty, knows the utility associated with every travel alternative, chooses the mode that maximises his/her utility. A prerequisite for this assumption is the existence of stable and well known preferences. Although the utility model has limitations when it comes to describing the human decision making process as it focuses on the final choice (Kahneman and Tversky, 1979), it makes empirical estimation possible. It, therefore, plays a major role in travel choice analysis as an approximation to real decision making.

In traditional microeconomic utility maximisation, the individual's utility is maximised with respect to a bundle of continuous goods, **G**, subject to a budget constraint. In a 1966 paper, Lancaster recognised that the primary source of utility is the qualitative attributes (e.g. comfort, flavour, softness) of the goods, and not the goods *per se*. Consequently, the individual's choice problem can be framed as a choice between bundles of attributes to achieve maximum utility.

There are several ways of formulating discrete mode choice models (cf. Becker, 1965; DeSerpa, 1971; Train and McFadden, 1978; Jara-Diaz and Videla, 1989). Our model is based on the fairly general disaggregate choice model by Jara-Diaz and Videla (1989).[14] In this model, time is treated as one attribute among others assumed to affect modal choice. There is no explicit time constraint and, in its simplest form (which we will employ), income does not affect choice.

A representative individual *n* (suppressing indexation) is assumed to choose both a discrete good (a mode) and continuous goods to maximise his/her utility:

$$\max_{\mathbf{G},i} U = U(\mathbf{G}, \mathbf{Q}_i) \tag{1}$$

subject to

$$Y = \mathbf{1}'\mathbf{G} + c_i \quad i \in M \tag{2}$$

where 1 is a $K\times1$ column vector of ones, $G=[g_1, g_2,\ldots, g_K]$ is a $K\times1$ column vector of consumed continuous goods, c_i is the travel cost of mode i, Y is income, $Q_i=[q_{i1}, q_{i2},\ldots, q_{iR}]$ is a $R\times1$ column vector of "quality" attributes associated with travel mode i (excluding cost, but including time) and M is equal to the choice set of travel modes. The utility function is assumed to be twice differentiable, quasi-concave and increasing in G. Equation (2) is the budget constraint, in which the travel cost and income are normalised by the price of G. Conditional on the choice of travel mode i, we can derive conditional demands for $g_{ik}= g_{ik}(Y-c_i, Q_i)$ Associated with the conditional demands is a conditional indirect utility function, which defines the maximum attainable utility level, conditional on the choice of travel mode i, $V_i= V_i(Y-c_i, Q_i)$.

[14] See also Jara-Diaz, 1998.

The travel mode chosen will be the alternative that renders the highest conditional indirect utility, i.e. mode *i* is chosen if $V_i \geq V_j$; $\forall j \neq i$. From the conditional indirect utility function it follows that the marginal utility of income (μ) is:

$$\mu = \partial V_i / \partial Y = -\partial V_i / \partial c_i .$$

The value of characteristic *r* is equal to the marginal rate of substitution between that characteristic and the (negative of the) marginal utility of income:

$$VOA_{ir} = \frac{\partial V_i / \partial q_{ir}}{\partial V_i / \partial c_i}.$$

If attribute *t* for mode *i* equals time, the value of time (VOT_i) is given by:

$$VOT_i = \frac{\partial V_i / \partial q_{it}}{\partial V_i / \partial c_i}.$$

We need to estimate a generic value of time, because our SP data do not consist of specific modes.

We assume that both the attribute vector associated with travel mode *i* ($\mathbf{Q}_i$) and the conditional indirect utility function are linear in the arguments. Moreover, we add an error term, ε_i, to account for measurement errors. Thus, the conditional indirect utility function equals:

$$V_i = \alpha_i + \mu(Y - c_i) + \boldsymbol{\gamma}' \mathbf{Q}_i + \varepsilon_i . \tag{3}$$

When comparing V_i and V_j, only variables that vary between modes affect choice. Therefore, the relevant part of the conditional indirect function is, $\overline{V_i}$, a truncated conditional indirect utility function, where income is not included

$$V_i = \alpha_i + \mu c_i + \boldsymbol{\gamma}' \mathbf{Q}_i + \varepsilon_i ; \ \forall i \in M .$$

In the estimation $\boldsymbol{\gamma} = [\gamma_t, \gamma_s, \gamma_{ef}]'$. If the respondent perceives the attributes the way we do, $\partial V_i / \partial c_i$ is expected to be negative (so that the negative of this parameter equals the marginal utility of income) because there is disutility from travel cost. By analogous reason, the time parameter, γ_t, is expected to be negative. Consequently, the *VOT* is the ratio of two negative numbers and, therefore, expected to be positive. All other parameters in the vector γ are expected to be positive because utility is assumed to increase from increases in the modal

safety, environmental friendliness and flexibility. The marginal rates of substitution between these attribute parameters and the cost parameter are therefore all expected to be negative.

Thus, the model used in the estimation is a linear, additive in the parameters, model. Linear additive models are *compensatory*, which means that the individual can trade-off a low value on one attribute for a high value on another and achieve the same utility. Empirically, compensatory choice models are found to work well in simple choice tasks such as in the choice between two alternatives (Payne, 1976).

4. ESTIMATION

4.1. Standard Logit

In the SP choice experiment respondent *n*'s choice set (*M*) consists of two alternatives, *A* and *B*. The respondent is asked to indicate which of the two alternatives that (s)he prefers, i.e. which alternative is "chosen".

The probability that the individual (suppressing individual indexation) will choose alternative *B* over *A* is given by:

$$\Pr(B) = \Pr(U_B > U_A) = \Pr(V_B + \varepsilon_B > V_A + \varepsilon_A) = \Pr(\Delta\varepsilon > -\Delta V),$$

where $\Delta\varepsilon = \varepsilon_B - \varepsilon_A$ and $\Delta V = V_B - V_A$. The probability that *A* will be chosen over *B* is logically Pr(*A*)=1-Pr(*B*). The dependent variable, $y_n \in \{0,1\}$ is an indicator variable taking unit value if alternative *B* is chosen and zero if alternative *A* is chosen. Different assumptions about the distribution of the random terms ε_A and ε_B or about their difference, $\Delta\varepsilon$, lead to different choice models. Assuming that $\Delta\varepsilon$ is normally distributed (ε_A and ε_B are normally distributed) results in the probit model and assuming that $\Delta\varepsilon$ is logistically distributed, i.e. that ε_A and ε_B are independently and identically distributed (IID) extreme value type I distributed, results in the logit model. Unless there are compelling reasons for assuming one distribution or the other, the choice between logit and probit is, in binary cases, only a matter of preference. Here, we assume $\Delta\varepsilon$ to be logistically distributed. The log-likelihood function is:

$$\ln \ell = \sum_{n=1}^{N} \left[(1 - y_n) \ln \frac{1}{1 + \exp(\Delta V)} + y_n \ln \frac{1}{1 + \exp(-\Delta V)} \right].$$

We estimate the model:

$$\Delta V = \alpha_B - \alpha_A + \mu(c_B - c_A) + \boldsymbol{\gamma}'(\mathbf{Q}_B - \mathbf{Q}_A) + \Delta\varepsilon.$$

Pooling the environmental friendliness and the flexibility sub-samples, and estimating the combined model with the use of dummy variables for the flexibility sub-sample result in

rejection of the null hypothesis of equal sub-samples.[15] Therefore, the two sub-samples are treated separately.

The results for the environmental friendliness sub-sample are given in Table 4.[16] Table 4 shows that the time and cost variables are significant at the five percent level with the expected signs (i.e. there is disutility from spending time and money on work trips). The safety and the environmental friendliness parameters have the expected signs, but are insignificant. The standard deviation of the environmental friendliness parameter is very low, indicating a good precision of the estimate. Nonetheless, the parameter is insignificant, meaning that considerations about the mode's environmental impacts are unimportant when making modal decisions for work trips. The number of correctly predicted choices in this model is 116.

Table 4. Environmental friendliness sub-sample (I): Results from standard logit estimations of stated choices

Attribute	Parameter	St dev	t-value
Constant	0.063	0.214	0.293
$TIME_I$	-0.037	0.008	-4.752
$COST_I$	-0.028	0.010	-2.804
$SAFETY_I$	0.000	0.001	0.289
ENV. FRIENDL	0.000	0.000	1.037
n	77		
Choices made	143		
LRI	0.277		
$\ln \ell$	-71.50		

Note: The likelihood ratio index $(LRI) = 1 - (\ln \ell / \ln \ell_0)$.

Table 5. Flexibility sub-sample (II): Results from standard logit estimations of stated choices

Attribute	Parameter	St dev	t-value
Constant	0.989	0.219	4.506
$TIME_{II}$	-0.047	0.010	-4.587
$COST_{II}$	-0.069	0.018	-3.891
$SAFETY_{II}$	0.004	0.002	1.899
FLEXIBILITY	0.001	0.001	0.873
n	85		
Choices made	148		
LRI	0.232		
$\ln \ell$	-71.07		

Note: The likelihood ratio index (*LRI*)=1-($\ln\ell/\ln\ell_0$).

[15] LR-test statistic: 11.02 with c^2 critical value: 9.49.
[16] The standard logit estimations were conducted in LIMDEP for both sub-samples.

Table 5 shows the results from the flexibility sub-sample. The time and cost parameters are, again, significant at the five percent level with the expected signs. The safety parameter is significant at the six percent level with the expected sign. The flexibility parameter is not significant, but has the expected sign. The relatively large variance of this parameter is likely to be a consequence of the lack of specificity in the definition of this attribute. Thus, the flexibility attribute seems to have been perceived differently by the respondents. The number of correctly predicted choices is in this model 114.

Table 6 gives the values of time for the two sub-samples, calculated as:

$$VOT = \frac{\partial V / \partial TIME}{\partial V / \partial COST} = -\frac{\beta_t}{\mu},$$

and the value of safety for the flexibility sub-sample, calculated as:

$$VOT = \frac{\partial V / \partial SAFETY_{II}}{\partial V / \partial COST_{II}} = -\frac{\beta_{s,II}}{\mu}.$$

The standard errors are calculated using the Delta method (cf. Greene, ch. 10, 1993).

Table 6. SP values of time and safety (1997 prices), standard deviations (St dev) and 95 percent confidence intervals (CI)

	Value	St dev	CI (95%)
VOT_I	79.50 SEK/hour	29.29	22.09 - 136.91
VOT_{II}	40.80 SEK/hour	8.85	23.45 - 58.15
$VOA_{s,II}$	0.07 SEK/percent	0.03	0.01 - 0.13

We refrain from calculating values of environmental friendliness and flexibility because these variables were all insignificant. On the same grounds, we refrain from calculating the value of safety in the environmental friendliness sub-sample.

The estimated values of time and safety are significant. The value of time in the environmental friendliness sub-sample (I) is almost twice the value time in the flexibility sub-sample (II). The difference is not significant, but the VOT_{II} has better precision than the VOT_I. A possible explanation for the lower value time in the flexibility sub-sample is the significance of the value of safety. That is, when the respondent really trades-off between attributes, the value of time may decrease.

Following Jones-Lee (1990), we express individual *n*'s marginal rate of substitution of wealth for risk by the ratio $WTP_{ns} = \partial\rho$, where $\partial\rho$ is the absolute change in risk level and WTP_{ns} defines the willingness to pay for a safety increase equal to $\partial\rho$. In a group of individuals affected by an increase in safety, the aggregated value of a statistical life equals the mean of these individual marginal rates of substitution. If the individuals affected by a risk reduction are relatively homogenous and the individual risk reductions are small, the value of statistical life is independent of the size of the group and the pattern of individual

risk reductions. Under such circumstances, the total value of statistical life equals $WTP_s/\partial\rho$ where WTP_s equals the population mean willingness to pay.

In our survey, the $VOA_{s,\,II}$ for every work trip equals SEK 0.07 per percentage reduction in risk. Given that most people have two work trips per day and, on average, 220 workdays per year, we get an annual value of safety, a willingness to pay for safety (WTP_s), equal to SEK 30,80 per percentage reduction in risk.

Because we used the respondents' perceptions of safety we are unable to calculate a *point estimate* for the value of statistical life.[17] Nonetheless, we can give a *function* for the value of statistical life at different levels of perceived risk. Assuming different initial levels of risk for work trips and reducing the risk by exactly 1/100,000 gives rise to the function in Figure 2. For example, in 1997, the *objective* risk for being *killed* or *seriously injured* in a car was approximately 29 per 100,000 of the average population (SIKA, 1998). If the average respondent perceives this objective risk as the initial risk for being killed in traffic, the value of statistical life is equal to SEK 10.6 million, based on the assumptions above.[18].

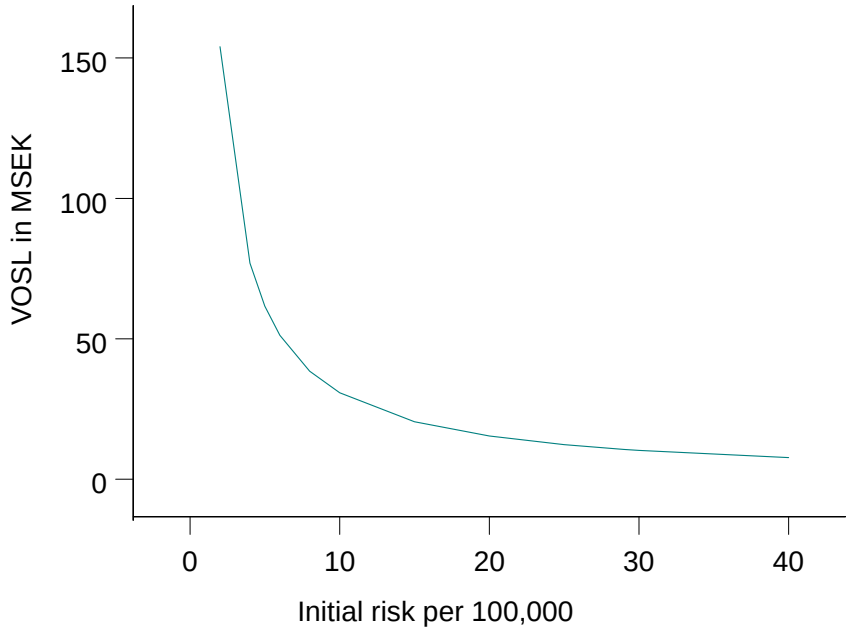


Figure 2. Value of statistical life (VOSL) at different initial risk levels.

[17] We assume, everywhere mentioned, that the initial risk is equal to the risk of being *killed* in road traffic. Therefore, the value of statistical life can be calculated from every assumed initial risk level. However, if the respondents perceive the initial risk (safety) differently, so that, for instance, the initial risk is perceived as the risk for being *killed* or *seriously injured* in traffic, an unambiguous value of statistical life can not be calculated.

[18] Reducing the risk from 29 to 28 in a hundred thousand equals a risk reduction of 3.45 percent. For all work trips in a year we have: $(-0.07*-3.45*2*220)*(1/100{,}000)^{-1}$ or, more generally, $VOSL = VOA_s * \partial\pi * n_{trips} * n_{days} * \partial\rho^{-1}$, where *VOSL* is short for value of statistical life and $\partial\pi$ is the risk reduction expressed in percent.

Similarly, the risk for being *killed* in car was 4 per 100,000 of the average population in 1997 (SIKA, 1998).[19] If the respondent correctly perceives this objective risk level as the initial risk level, the value of statistical life is equal SEK 77 million. Thus, the crucial assumption for the value of statistical life is the assumption about the perceived initial risk. A Swedish contingent valuation survey (Persson et al., 1998) has found that the mean perceived risk for being killed in a traffic accident during a year equal to 72/100,000 (median: 4/100,000). Employing their mean value gives a value of statistical life equal to SEK 4.3 million. This value could serve as a lower bound for the value of statistical life for work trips.

4.2. Random Parameters Logit

The standard logit model imposes several restrictions on the parameters of the model (McFadden, 1981; Ben-Akiva and Lerman, 1985). First, the parameters are assumed to be the same for all respondents, meaning that two identical respondents (with respect to the observed variables) must have the same parameter values, i.e. tastes. Second, following from the assumption of IID extreme value error term, the logit model suffers from the independence from irrelevant alternatives (IIA) (cf. Ben-Akiva and Lerman, 1985). With IIA, the logit model necessarily predicts that a change in one attribute of an alternative (or the introduction of a new alternative or the elimination of an existing alternative) changes the probability for the other alternatives proportionately, so that the ratios of probabilities remain unchanged (Brownstone and Train, 1999). In many circumstances, this is an unrealistic assumption, resulting in implausible substitution patterns. Third, in cases where there are repeated choices, either over time, as in revealed preferences situations, or over choices, as in stated preferences situations, the standard logit model assumes that unobserved variables affecting the choices are independent over time and choices so that there is no correlation in the unobserved utility over time and choices. In many settings, this is also an unrealistic assumption.

To relax the above restrictions, a less restrictive model, the ”random parameters” logit model[20] , was estimated. The random parameters logit model is a generalisation of the standard logit model, where the parameters of the attributes are allowed to vary randomly over respondents. The random parameters logit model does not exhibit the IIA property and may, therefore, represent any substitution pattern (Train, 1998). Furthermore, in the random parameters logit model the dependence of unobservable variables over choices is explicitly modelled. For instance, the individual’s value of time (the ratio between the time and cost parameters) may depend on the unobserved variable ”patience” which may vary among the respondents (Horowitz, 1981).

Random parameters logit models have been applied in various settings, e.g. in recreational demand models for the choice of fishing site (Train, 1998), in consumption demand models for households’ choice of appliance efficiency level (Revelt and Train, 1998) and households’ choice between vehicles with different fuel types (Brownstone and Train, 1999), and in travel choice models (Algers et al., 1998).

[19] In 1997, 348 persons were killed in car accidents (SIKA, 1998). It is important to remember that these are ”primary” fatalities (people being killed when travelling by car), whereas there are also ”secondary” fatalities (people not travelling by car who are killed by a car).

[20] Also known as mixed logit, random parameters logit and error components logit.

In the random parameters logit model, the indirect utility of alternative *i* in choice situation *t* for individual *n*, is specified as $V_{nit} = \boldsymbol{\beta}_n' \mathbf{x}_{nit} + \varepsilon_{nit}$, where x_{nit} is a vector of explanatory variables and β_n is a vector of unobserved parameters that varies in the population. The disturbance term, ε_{nit}, is still assumed to be IID extreme value distributed. This specification is analogous to the standard logit specification of V_i in Equation (3), except for the parameter vector β_n which varies over respondents and the sub-indexation *t* which accounts for the data's panel character (the repeated choices). Following Train (1998) the parameter vector β_n is expressed as the sum of the population means, b, and individual deviations, η_n, where η_n represents the respondent's taste relative to the population's. The utility may be re-expressed as:

$$V_{nit} = \mathbf{b}' \mathbf{x}_{nit} + \boldsymbol{\eta}'_n \mathbf{x}_{nit} + \varepsilon_{nit}. \tag{4}$$

The term $\boldsymbol{\eta}'_n \mathbf{x}_{nit} + \varepsilon_{nit}$ is unobserved and, as the researcher can not separate the individual deviation from the independent error term, the existence of η_n induces heteroscedasticity and correlation over choices. That is, because the respondent uses the same "tastes" when evaluating choices and, as the researcher is unable to completely observe these tastes, there is correlation in the unobserved part of the indirect utility function.

Generally, the unobserved parameter vector β_n is assumed to vary in the population with density $f(\boldsymbol{\beta}_n | \boldsymbol{\theta})$, where θ is a vector of the true parameters of the distribution (for example, the mean and standard deviation). β_n represents individual *n*'s "tastes", which are assumed to be constant over the repeated choices.[21] Our objective is to estimate θ, the true population parameters that describe the distribution of the individual parameters. Here, we assume $\boldsymbol{\eta}_n$ and $\boldsymbol{\beta}_n$ to be normally distributed, $\boldsymbol{\eta}_n \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\beta}_n \sim N(\mathbf{b}, \boldsymbol{\Sigma})$ and $\boldsymbol{\eta}_n \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.[22] Thus, b and Σ are the vectors of parameters we wish to estimate.

We estimate the these parameters in Equation (4) by maximisation of a simulated log-likelihood function (MSL) with *R*=1,000 repetitions.[23] All random parameters logit models were estimated with ten different starting values to ensure the finding of a global maximum.[24]

To begin with, the most obvious way to introduce individual heterogeneity would be to estimate models with random intercepts while keeping all other variables fixed. However, compared to the standard logit models, such models do not improve the (simulated) likelihood functions and are rejected by likelihood ratio tests (not reported). We, therefore, fix the intercept terms of both models, and let, instead, the explanatory variables be random.[25]

---

[21] Modification could be done so that the respondent's tastes vary over time (Train, 1998). However, in our case we assume that the time elapsed between the consecutive choices is too small to significantly affect the respondent's tastes.

[22] $\boldsymbol{\Sigma} = diag(\boldsymbol{\sigma}_p)$ where *p* equals the explanatory variables used in the standard logit model.

[23] See Train (1998) and Revelt and Train (1998) on the maximisation of simulated likelihoods. The estimations was conducted in a GAUSS program written by Kenneth Train, David Revelt and Paul Ruud.

[24] As starting values we use $\mathbf{b} = \hat{\boldsymbol{\lambda}} + \boldsymbol{\zeta}$, where $\hat{\boldsymbol{\lambda}}$ is a *p*×1 vector of standard logit estimates and $\boldsymbol{\zeta} \sim N(\mathbf{0}, \mathbf{0.1})$. For $\boldsymbol{\Sigma}$, we use draws from the distribution $\boldsymbol{\Sigma} \sim Uniform[0,1]$.

[25] Allowing all parameters to vary results in identification problems (Revelt and Train, 1998).

For the flexibility sub-sample, this model is rejected by a likelihood ratio test ($\ln\ell$=-69.84) but, for the environmental friendliness sub-sample, we can not reject this model ($\ln\ell$=-61.78).

The results from the environmental friendliness estimation are reported in Table 7. The mean of the time variable is significant at the five percent level and the mean and standard deviation of cost are significant at the ten percent level. Thus, there are significant variations in the cost parameter among the respondents in this sub-sample. The point estimate of the time parameter implies that six percent of the respondents have a positive time parameter, while the point estimate of the cost parameter implies that about 16 percent have a positive cost parameter. While it seems plausible that a few respondents may enjoy the time travelling to work, it seems less plausible that some respondents enjoy having more expensive work trips. An explanation for this unexpected value, also found by Algers et al. (1998), may be some respondents' eligibility to tax deductions for work trips.

Comparing the random parameters logit parameter estimates with those from the standard logit, we see that the estimated parameters are larger in absolute magnitude in the random parameters logit. Because the variance of the IID error term is greater in the standard logit than in the random parameters logit and, as the scale of utility is determined by the normalisation of the error term, this is what is to be expected (Brownstone and Train, 1999).

Table 7. Environmental friendliness sub-sample (I): Results from random parameters logit estimations of stated choices

Attribute	Parameter	St dev	t-value
Constant	0.119	0.426	0.280
Mean of time	-0.131	0.062	-2.107
St dev of time	0.086	0.062	1.382
Mean of cost	-0.190	0.113	-1.680
St dev of cost	0.191	0.111	1.715
Mean of safety	0.004	0.005	0.954
St dev of safety	0.007	0.009	0.773
Mean of env.friendl.	0.001	0.002	0.793
St dev of env.friendl.	0.000	0.002	0.176
N	77		
Choices made	143		
SLRI	0.376		
ln ℓ (simulated)	-61.78		

Note: The simulated likelihood ratio index (*SLRI*)=1-($\ln\ell_s/\ln\ell_0$).

The point estimate of VOT_I (using $\hat{b}_{\text{time}}/\hat{b}_{\text{cost}}$) is now SEK 41.37 per hour. Thus, it is lower than the equivalent from the standard logit model. Lower value of time, when employing a random parameters logit model with normal parameters, has also been found by Algers et al. (1998). Because each respondent now has an "own" value of time - which is the ratio between two normally distributed variables - the correct mean value of time would be given by integrating over this distribution. However, this considerably complicates the

calculation and also leads to some individuals having implausible values of time as the normal distribution allows both negative and positive individual parameters.

One may assume lognormal distributions for the time and cost parameters and, thereby, make sure that these variables get positive individual parameter values (when estimated on the negative of respective variable). However, for our dataset, this model failed to converge. Convergence failures in estimations with log-normally distributed variables are also recognised by Algers et al. (1998) and Brownstone and Train (1999).

5. CONCLUSION

Based on a stated choice experiment, we estimate standard and random parameters logit models of commuters' choice of travel mode. The choices are explained by the individual's perceptions of five modal attributes; time, cost, safety, environmental friendliness and flexibility. Time, cost and safety were evaluated by all respondents while environmental friendliness and flexibility were used interchangeably. Depending on the attribute evaluated by the respondent, the sample was divided into an environmental friendliness and a flexibility sub-sample. The respondents' preferences were elicited by a novel method in which the names of two modes (car and bus) were used as attribute levels instead of objective levels. Based on the respondents' revealed preferences, we infer "perceptions-based" real values for the time and cost attributes. In the analysis, the inferred values are used together with the perceptions of the other attributes. From the estimated model we derive a value of time for work trips and a monetary value for safety improvements (value of statistical life). Time savings and saved statistical lives are the major benefits in project evaluations of the Swedish National Road Administration (SNRA). Whereas the SNRA uses values derived from separate studies, we estimate them in the same model.

The survey produced several interesting findings. Overall, time and cost are found to be the most important attributes for work trip modal choice, while safety is non-negligible in one of the sub-samples. Whereas a model specification with normally distributed parameters (the random parameters logit) improved model fit for the environmental friendliness sub-sample, the standard logit model with fixed parameters could not be rejected for the flexibility sub-sample. For both sub-samples, the models with the best fit (highest log-likelihoods) resulted in values of time of about SEK 40 per hour. In a national survey from 1994 (Algers et al., 1995), the value of time for work trips was found to be SEK 34 - 54 per hour for bus, car and train. Adjusted for price changes, these figures are SEK 35 - 56 (1997 prices), which encompasses our estimate, even though a direct comparison is erroneous because the different values of time are estimated on different travel modes. Thus, in a "test-retest" sense, the point estimates of the value of time in our survey are reliable. In 2005, the value of time recommended to the Swedish National Road Administration (SNRA) for cost-benefit calculations is SEK 42 per hour for regional (<50 kilometres) trips (SIKA, 2005).

In this survey the work trip value of statistical life in road traffic depends on the perceived baseline, or initial, risk. We find a lower bound for the value of statistical life equal to SEK 4.3 million (1997 prices), based on a perceived initial risk of 72/100,000. As is recognised in psychological research, people tend to underestimate the risks of well known, voluntary, low probability events (Lichtenstein et al., 1978). Because most people probably

define driving to work as a low probability event, a much higher value of statistical life is implied. A Swedish contingent valuation survey found the value of statistical life to be SEK 19.4 million (Persson et al., 1998). In 2005, the value of statistical life recommended to the SNRA for cost-benefit calculations was SEK 14.2 million (SIKA, 2005).

Though many things could have been done differently, the use of perceptions data is a novelty to SP analysis.[26] And although our perceptions data do not dramatically change, or improve, the value of time estimate, the estimations show that perceptions can be useful for explaining travel mode choices. Nonetheless, the findings from this survey are encouraging and constitute a point of departure for future research. In order to improve model fit and predictive power of travel choice models, future research will focus on identifying latent variables that affect mode choice. For instance, purposefully collected survey data, modelled in linear structural equation models, where the latent variables are related to indicator variables (cf. Vredin Johansson et al., 2006; Morikawa et al., 1990; Bollen, 1989), will, by estimation in the LISREL program (Jöreskog and Sörbom, 1993), help to identify latent variables. Such estimations could provide fitted values that may be used as explanatory variables in discrete travel choice models.

ACKNOWLEDGEMENTS

The author would like to thank Per Johansson for generous and competent guidance throughout the work on this essay. The essay also benefited much from the comments and suggestions of Thomas Aronsson, Jörgen Hellström, Lars Hultkrantz and Karl-Gustaf Löfgren. Financial support from The Swedish Road Administration is gratefully acknowledged. The usual disclaimer applies.

APPENDIX A: THE PERCEPTIONS AND CHOICE QUESTIONS (QUESTIONNAIRE A, QUESTIONS 19-21)

In the question below we would like you to *compare* two different travel modes for a *journey to your workplace;* car and bus. We would like you to do the comparisons *even* if you do not have a car and *even* if it is impossible to go by bus to your workplace. Try to imagine what it would be like if both alternatives existed.

19. As you know different travel modes are associated with different attributes, for example different travel time, cost, safety and environmental friendliness. How do you *perceive* the different travel modes with respect to respective attribute?

Start out from the car's attributes and make comparisons. The attributes of the car are always 100%. *Example:* If you perceive the travel time with bus *equal to* the travel time with

car, you answer 100%. If you perceive the travel time with bus *three times* as long as the travel time with car, you answer 300%. If you perceive the travel time with bus *half of* the travel time with car, you answer 50%. You may, of course, use any percentages - the important thing is that you, as accurately as possible, describe *your own* perceptions.

a) Compared with *car*, I perceive the TRAVEL TIME with *bus* to be: _____%

b) Compared with *car*, I perceive the COST with *bus* to be: _____%

c) Compared with *car*, I perceive the SAFETY with *bus* to be: _____%

Compared with *car*, I perceive the ENVIRONMENTAL FRIEND-LINESS with *bus* to be: _____%

We ask the following two questions to find out how people value different attributes in a journey to work.

Suppose you are confronted with the choice between two hypothetical travel modes (A and B) for *a journey to your workplace*. These travel modes are characterised by *four different attributes* (travel time, cost, safety and environmental friendliness). There are no ”real” travel modes that are described, but travel modes characterised by combinations of different attributes. Now we ask you to *make a choice between A and B* on the basis of the attributes that characterise travel modes A and B. How you choose is up to you, but we ask you to think thoroughly through your opinions before answering question 20 and 21.

Remember: There is no right or wrong answer! You have the right to think and choose as you like - it is *your* opinion that are of interest to us.

20. Which alternative (A eller B) do you prefer?

Attributes	A	B
Travel time like:	Car	Bus
Cost like:	Bus	Car
Safety like:	Car	Bus
Environmental friendliness like:	Car	Bus

Answer: I prefer (mark with one cross):( ) Travel mode A
( ) Travel mode B
( ) The alternatives are equal

[26] For example, information gathering on exactly how the respondents perceive the attributes should be improved. Moreover, the survey should account better for situational constraints (such as bad weather and/or seasonal

21. Which alternative (A eller B) do you prefer?

Attributes	A	B
Travel time like:	Bus	Car
Travel cost like:	Bus	Car
Safety like:	Car	Bus
Environmental friendliness like:	Car	Bus

Answer: I prefer (mark with one cross):( ) Travel mode A
( ) Travel mode B
( ) The alternatives are equal

APPENDIX B: THE EXPERIMENTAL DESIGN

Following the notation of Box et al. (1978) the three variables and the combination variable are denoted as follows; 1=*t* (travel time), 2=*c* (travel cost), 3=*s* (safety), 4=*ef* (Environmental friendliness/Flexibility). All variables have two levels, a high level and a low level. As previously, we employ a plus sign (+) to indicate the high level and a minus sign (-) to indicate the low level. Furthermore, if the elements of any column are multiplied by themselves, we obtain a column of plus signs, I, i.e. I=11=22=33=44.

Table B1. Signs for a 2^4 factorial design

Alt.	Variables				Interaction effects										
	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
2	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-
3	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
5	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
6	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
8	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
9	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
10	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+
11	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
12	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
13	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
14	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-
15	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

differences) and, in order to maximise validity, be related to specific and recent choice situations.

In the following description of the experimental design, it is important to distinguish between the concepts of alternatives, blocks, pairs and sub-sub-samples (denoted sss in Table B2). When we have four variables varying at two different levels and desire two pair-wise comparisons in each sub-sub-sample, we need three ”block” variables to do the job. Therefore, let us introduce the block variables, B_1, B_2 and B_3, where B_1=12, B_2=13 and B_3=34 (see Table B2). Hence, B_1 is equal to the two-way interaction between *t* and *c*, B_2 is equal to the two-way interaction between *t* and *s* and B_3 is equal to the two-way interaction between *s* and *ef*.

Table B2. Block variables and experimental design

Alternative	B_1	B_2	B_3	Pair no	sss	Pair no	Alternatives	Questionnaire
1	+	+	+	8	1	8	1, 16	A and E
2	-	-	+	7				
3	-	+	+	6		7	2, 15	A and E
4	+	-	+	5				
5	+	-	-	4	2	6	3, 14	B and F
6	-	+	-	3				
7	-	-	-	2		5	4, 13	B and F
8	+	+	-	1				
9	+	+	-	1	3	4	5, 12	C and G
10	-	-	-	2				
11	-	+	-	3		3	6, 11	C and G
12	+	-	-	4				
13	+	-	+	5	4	2	7, 10	D and H
14	-	+	+	6				
15	-	-	+	7		1	8, 9	D and H
16	+	+	+	8				

Depending on the signs of the block variables, we assign ”pair” numbers to the different alternatives. Hence, the two alternatives with all block variables at the high level (+ + +) constitute a pair (pair no 8) while alternatives with block variables minus, minus, plus (- - +) constitute another (pair no 7), and so on. The result of using the block variables is eight different pairs. These pairs are divided into four sub-sub-samples consisting of two pairs each. For example, sub-sub-sample one consists of two pair-wise comparisons, pairs number eight (8) and seven (7), where pair eight equals the choice between alternative one and 16 and pair seven equals the choice between alternative two and 15. For the other three sub-sub-samples the alternatives are given in Table B2. The two alternatives in each pair are complementary in signs, as the signs of the attribute levels of the first alternative are reversed in the second. Therefore, each pair is said to be a *fold-over* pair (Box et al., 1978). For example, in pair no 8 there is a choice between alternative one and 16. Regarding Table B1, we see that while alternative one is low on all variables, alternative 16 is high on all variables.

This blocking arrangement is copied from Box et al. (1978, p. 347). Clearly, there are some, but not so many, other possible combinations of block variables. However, caution is called for, because the use of wrong block variables may lead to main attribute effects being

confounded with block effects. Using any other three *possible* (without confounding the block effects with the main attribute effects) combinations of block variables leads to the same result as the ones we use; all two-way interactions and the four-way interaction are confounded with the block effects and the pairs consist of the same alternatives.

If the we use the notation and the multiplicative properties from above, we have; B_1B_2=1213=23, B_1B_3=1234**,** B_2B_3=1334=14 and $B_1B_2B_3$=121334=24, which means that interaction effects 12 (time-cost), 13 (time-safety), 14 (time-environmental friendliness/flexibility), 23 (cost-safety), 24 (cost-environmental friendliness/flexibility), 34 (safety-environmental friendliness/flexibility) and 1234 (time-cost-safety-environmental friendliness/flexibility) are confounded with the block effects and, therefore, not possible to estimate separately. However, none of the main attribute effects are confounded.

REFERENCES

Adamowicz, W., Swait, J., Boxall, P., Louviere J., and Williams, M. (1997). Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *Journal of Environmental Economics and Management*, *32*, 65-84.

Algers, S., Bergström, P., Dahlberg, M., and Lindqvist Dillén, J. (1998). *Mixed Logit Estimation of the Value of Travel Time.* Working Paper 15, Department of Economics, Uppsala University.

Algers, S., Hugosson, B., and Lindqvist Dillén, J. (1995). 1994 års tidsvärdesstudie. Slutrapport. Del 1 Resultat. Transek, mimeo. (In Swedish with a summary in English).

Batsell, R. R., and Louviere, J. J. (1991). Experimental Analysis of Choice. *Marketing Letters*, *2(3)*, 199-214.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, *75*, 493-517.

Ben-Akiva, M., and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA, MIT Press.

Björkman, M., and Ekman, G. (1957). *Experimentalpsykologiska metoder*. Stockholm, Almqvist and Wiksell.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, Wiley.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. New York, John Wiley and Sons.

Brownstone, D., and Train, K. (1999). Forecasting New Product Penetrations with Flexible Substitution Patterns. *Journal of Econometrics*, *89*, 109-129.

DeSerpa, A. C. (1971). A Theory of the Economics of Time. *The Economic Journal*, *81*, 828-846.

Greene, W. H. (1993). *Econometric Analysis* (2nd edition). New York, Macmillan.

Gårder, P., Leden, L., and Thedéen, T. (1994). Safety Implications of Bicycle Paths at Signalized Intersections. *Accident Analysis and Prevention, 26(4)*, 429-439.

Horowitz, J. L. (1981). Identification and Diagnosis of Specification Errors in the Multinominal Logit Model. *Transportation Research, 15B*, 345-360.

Jara-Diaz, S. R. (1998). Time and Income in Travel Choice. Towards a Microeconomic Activity-Based Theoretical Framework. In T. Gärling, T. Laitila and K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (51-74). Oxford, Elsevier Pergamon.

Jara-Diaz, S. R., and Videla, J. (1989). Detection of Income Effect in Mode Choice: Theory and Application. *Transportation Research, 23B*, 393-400.

Jones-Lee, M. W. (1990). The Value of Transport Safety. *Oxford Review of Economic Policy, 6(2)*, 39-60.

Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language*. Chicago, SSI.

Kahneman, D., and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47(2)*, 263-291.

Kommunikationsdepartementet (1997): *På väg mot det trafiksäkra samhället*. Ds 1997:13, Stockholm.

Koppelman, F. S., and Pas, E. I. (1980). Travel-Choice Behavior: Models of Perceptions, Feelings, Preference and Choice. *Transportation Research Record, 765*, 26-33.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy, 14*, 132-157.

Lee, T. R. (1981). Perceptions of Risk. The Public's Perception of Risk and the Question of Irrationality. *Proceedings of the Royal Society of London, A376*, 5-16.

Lenner, M. (1993). *Energy Consumption and Exhaust Emissions Regarding Different Means and Modes of Transportation*. VTI Meddelande nr 718. (In Swedish with a summary in English).

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged Frequency of Lethal Events. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 551-578.

Louviere, J. J., and Johnson, R. D. (1991). Using Conjoint Analysis to Measure Retail Image. In A. Ghosh, and C. A. Ingene (Eds), *Research in Marketing. Spatial Analysis in Marketing: Theory, Methods and Applications*. Supplement 5, 137-156, Jai Press Inc.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. In C. Manski, and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications* (198-272), Cambridge, The MIT Press.

Morikawa, T., Ben-Akiva, M., and McFadden, D. (1990). Incorporating Psychometric Data in Econometric Choice Models. Draft version, July 1990.

Mowen, J. C. (1990). *Consumer Behaviour* (2nd edition). New York, McMillan.

Noland, R. B. (1995). Perceived Risk and Modal Choice: Risk Compensation in Transportation Systems. *Accident Analysis and Prevention, 27(4)*, 503-521.

Payne, J. W. (1976). Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. *Organizational Behavior and Human Performance, 16*, 366-387.

Persson, U., Nilsson, K., Hjalte, K., and Norinder, A. (1998). *Beräkning av Vägverkets riskvärden. En kombination av contingent valuation-skattningar och uppmätta hälsoförluster hos vägtrafikskadade personer behandlade vid fyra sjukhus*. The Swedish Institute of Health Economics (IHE), mimeo. (In Swedish).

Revelt, D., and Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics, 80(4)*, 647-657.

SCB (1998). *Svenskarnas resor. Riks-RVU resultatrapport för 1997*. Statistics Sweden. Stockholm. (In Swedish with a summary in English).

SIKA (1998). *Traffic Injuries 1997*. Official Statistics of Sweden, Örebro, SCB.

SIKA (2005). *Kalkylvärden och kalkylmetoder (ASEK). En sammanfattning av Verksgruppens rekommendationer 2005*. SIKA PM 2005:16. (In Swedish).

Slovic, P., Fischoff, B., and Lichtenstein, S. (1981). Perceived Risk: Psychological Factors and Social Implications. *Proceedings of the Royal Society of London, A376*, 17-34.

Train, K. and McFadden, D. (1978). The Goods/Leisure Tradeoff and Disaggregate Work Trip Mode Choice Models. *Transportation Research, 12(5)*, 349-353.

Train, K. E. (1998). Recreation Demand Models with Taste Differences over People. *Land Economics, 74(2)*, 230-239.

Vredin Johansson, M., Heldt, T., and Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A, 40*, 507-525.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 163-178

Chapter 7

THE PSYCHOLOGY OF DECISION-MAKING IN ECONOMICS: A REVIEW

Jeffrey A. Livingston and Michael A. Quinn
Bentley College

ABSTRACT

Economics has always focused on how individuals make decisions. Traditionally, the discipline has viewed individuals as rational agents maximizing their own utility. However, economists have recently begun to incorporate research from the field of psychology in creating a richer view of decision making. This push is the result of challenges to the neoclassical model made by theoretical advances such as Kahneman's Nobel-winning prospect theory model and from the field of experimental economics. This growing field has revealed many aspects of human behavior that cannot be explained by traditional economic models. Some of these aspects of behavior include loss aversion (as explored by Kahneman); relative deprivation (the theory that individuals consider their relative position as compared to others when making decisions), motivations of altruism, fairness, and reciprocity; and the endowment effect (individuals tend to value goods more highly if they are already in possession of them). These innovations have impacted economists' views of issues such as consumption, worker-firm relations, labor supply, equities and real estate.

This paper reviews the impact of psychology on economic models of decision making. The major trends will be discussed, along with implications that these changes have for both economics and public policy.

I. INTRODUCTION

The foundation of modern economics is the neoclassical model of consumption and production, which views individuals as rational agents who seek to maximize their own expected utility. This model has been a powerful tool in accurately predicting how economic agents behave in most circumstances. However, in recent years economists have begun to recognize the model's limitations. Faced with evidence that individuals sometimes behave in

ways that are inconsistent with neoclassical predictions, economists have developed a new field called behavioral economics that considers how psychological motivations impact the choices people make. Collaboration between psychologists and economists has resulted in the identification of many such behavioral "anomalies" that are now accepted as fact, and these studies have impacted a wide range of mainstream fields including macroeconomics, labor economics, financial markets and environmental economics.

The impact of psychology on economics tends to be viewed as a new phenomenon. However, many of the notable early writers in economics clearly believed in the importance of psychology as a basis for understanding individual motivations and behavior. For instance, in describing how individuals' human nature and beliefs influences their "rational decisions" David Hume writes "Reason is, and ought only to be, the slave of the passions, and can never pretend to any other office than to serve and obey them." (Hume, 1740). In 1919, Irving Fisher wrote of the interest rate "The basis of interest . . . lies in the preference for present over future goods: Neither the employers more than the employee likes to wait for the fruits of any enterprise . . . whoever does so is entitled to some reward . . . The essence of interest is impatience, the desire to obtain gratifications earlier than we can get them. It is a fundamental attribute of human nature." (Fisher, 1919).

In the 1950s and 1960s, economics moved towards the idea of a "rational" individual. This perspective proposes that individuals are "expected utility maximizers" and discounts emotional responses to situations. This move in economics was made by those trying to make economics into a "hard science". For instance, by the 1950s, economists such as Metzler describe the interest rate as "changes in the quantity of money accompanied by open-market operations do affect the rate of interest and thus the rate becomes a monetary phenomenon in the sense of being co-determined by the quantity of money outstanding and by changes in this quantity" (Metzler, 1951). This stands in stark contrast to Fisher's explanation of the same phenomenon a few decades earlier. It is this contradiction between those establishing the cold, rational individual that would move economics towards the physical sciences and the fact that economists study phenomena of human behavior that has caused problems. Behavior has been observed that simply cannot be explained by a model of "rational" individual behavior.

The resurgence of awareness of psychological motivations in economics in the 1990s is, however, quite different from that of previous economists. While early writers such as Fisher and Keynes saw a significant role for psychology in their theories, this role was based on their own insights and intuition. Recent work in behavioral economics is often based on experiments and data analysis. The methods of experimental economics employ rigorous methodologies to test for the different motivations behind individual behavior and use these insights to inform economic theory.

The next section of the paper discusses the methodologies employed in behavioral economics and how they have lead to insights into individuals' motivations. After the discussion of these new methodologies, there is an examination of the basic results of behavioral economics that are no longer considered controversial. This is followed by applications of these results in other mainstream economics fields.

II. THE METHODOLOGY OF BEHAVIORAL ECONOMICS

The primary tool of behavioral economists is experiments that are conducted in the laboratory. Most behavioral economics studies argue that one of the standard assumptions about individual behavior made in neoclassical theory is incorrect, leading to inaccurate predictions about how people behave in some particular situation. Experiments remain the most popular method because they allow the researcher to construct tests where different theories have clearly different predictions about how subjects will behave, so the experiment can cleanly test which theory more accurately predicts behavior.

"Ultimatum" game experiments are a nice example. In this well-known game, two subjects are endowed with some amount of money. One subject, frequently called the "proposer," is allowed to propose how this money should be shared with another subject. The second subject, frequently called the "responder," can accept or reject this offer. If the responder accepts the offer, then the subjects are paid according to the terms of the proposer's offer. If the responder rejects the offer, both subjects receive a payoff of 0. "Rational" responders would accept any offer that contains a positive payoff, since the alternative is a gain of 0, and "rational" proposers would accordingly offer the responder as little as possible, keeping the large majority of the funds for themselves. However, in many repetitions of this experiment, using many different designs, the results have been robust and they have not conformed to this prediction. Rather, subjects frequently reject proposals which they perceive to be unfair, showing a willingness to sacrifice their own monetary gains in order to punish people who treat them unfairly. Camerer and Thaler (1995) provide a nice discussion of how economists have dealt with these results, and Camerer and Loewenstein (2004) provide a clear discussion of why the experimental methodology is necessary to provide clean evidence of this aspect of human behavior:

> "Suppose we observed this phenomenon in the field, in the form of failures of legal cases to settle before trial, costly divorce proceedings, and labor strikes. It would be difficult to tell whether rejection of offers was the result of reputation-building in repeated games, agency problems (between clients and lawyers), confusion, or an expression of distaste for being treated unfairly. In ultimatum game experiments, the first three of these explanations are ruled out because the experiments are played once anonymously, have no agents, and are simple enough to rule out confusion. Thus, the experimental data clearly establishes that subjects are expressing concern for fairness." (Camerer and Loewenstein, 2004).

Behavioral economists have only recently caught on to this method of studying human behavior which psychologists have long been familiar with. However, there are four striking differences between the methods used by scholars in the two disciplines. Hertwig and Ortmann (2001) provide a full review of these differences and their relative advantages and disadvantages; for a more thorough discussion, readers are encouraged to consult their article. The first is how the instructions for the experiment are presented to the subjects. Economists, who are always (perhaps too) keenly aware of how people react to any incentive, always follow a pre-written script to the letter. Subjects are given precise descriptions of the players involved, the set of actions that each player can possibly take, and the payoffs resulting from each possible sequence of actions. Subjects are then each assigned a role (for example, as a proposer or a responder in an ultimatum game) and follow the script for that role.

Psychologists, meanwhile, frequently do not provide a script or assign roles, forcing participants to ad-lib. Second, economists use repeated trials, while psychologists typically instead conduct "snapshot" studies. Third, economists always use performance-based monetary incentives so that subjects make actual rather than hypothetical choices; psychologists often pay subjects a fee for showing up, but their payment is typically unaffected by how they behave in the experiment. Finally, economists consider it unacceptable to deceive subjects in any way, while psychologists frequently use deception in their experimental designs. Economists have elected to avoid deception on the grounds that it may pollute the behavior of subjects in future experiments. It would become impossible to rule out the possibility that subject behavior was guided by their suspicion of being mislead, rather than the monetary rewards that are assigned to each action.

The frequent use of experiments by behavioral economists leads to some confusion that behavioral economics and experimental economics are one in the same. This perception is often used as ammunition against the field – doubters present the important argument that behavioral anomalies found in experimental studies may not be prominent in real markets. The subjects of these experiments are typically university students who are very different from people who participate in actual markets. They are typically younger, poorer, and less experienced than participants in the business world, so it is easy for skeptics to dismiss laboratory studies on the grounds that once people gain more experience in the market, they will behave in ways that are more in concert with the neoclassical model of behavior.

In fact, experimental and behavioral economics are not synonyms: many experimental studies have been conducted that are not designed to challenge neoclassical behavioral assumptions, and behavioral studies are frequently conducted using non-experimental methods. Behavioral economists use all of the same tools used by economists in other fields – careful analysis of naturally occurring field data, theory, and most recently "field" experiments that are conducted in an actual market rather than in the laboratory, as well as traditional laboratory experiments. In fact, perhaps the most widely cited behavioral study is the theoretical paper by Kahneman and Tversky (1979), which introduced "prospect theory."

Field experiments in particular have quickly become very popular in recent years. The chief advantage of field experiments is of course the ability to study the behavior of actual market participants rather than students in the lab, while still maintaining some control of the experimental environment, making it possible to rule out many alternative explanations for results. The disadvantages are that control of the setting is no longer perfect, and field experiments tend to be enormously expensive to conduct. Examples of this work include Gneezy and Rustichini (2000), who studied the effect of introducing a small fine charged to parents who failed to pick up their children on time from a day-care center in Israel, and List (2003), who examined the "endowment" effect (the tendency for people to place a higher value on something if they are already in possession of it) among two sets of agents at a sports memorabilia show – agents who had little experience in this marketplace, and highly experienced dealers who were very familiar with the market.

III. Basic Themes and Results

Studies using all of these methodologies have over time come to a consensus on several different results, each of which has seen application in more mainstream fields. The most important of these results include:

Risk Preferences, Loss Aversion, the Endowment Effect and Framing Effects

Perhaps the seminal paper in all of behavioral economics is "Prospect Theory: An Analysis of Decision Under Risk" by Daniel Kahneman and Amos Tversky, published in *Econometrica* in 1979. Kahneman was awarded the Nobel Prize in Economics in 2002 "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty" largely on the basis of this paper. The theory is motivated by hypothetical experiments that showed several situations where people make decisions that are inconsistent with the assumptions of expected utility theory. For example, they surveyed 72 people and asked them to state what choice they would make when confronted with two problems:

Problem 1: Choose between

A: 2,500 with probability 0.33	B: 2,400 with certainty
2,400 with probability	0.66
0 with probability	0.01

Problem 2: Choose between

A: 2,500 with probability 0.33	B: 2,400 with probability 0.34
0 with probability 0.67	0 with probability 0.66

In problem 1, 82 percent of the respondents selected option B, while in problem 2, 83 percent of the respondents selected option A. In oversimplified terms, expected utility theory states that preferences over several possible outcomes $x_1,\ldots, x_n$ that occur with probabilities $p_1,\ldots, p_n$ respectively can be represented by the following function:

$U(x_1,p_1;\ldots;x_n,p_n) = p_1u(x_1) + \ldots + p_nu(x_n)$, where the utility function $u(x)$ is concave, i.e. people are risk averse. In this case, the responses to problem 1 suggest that for most people,

$$u(2400) > 0.33u(2500) + 0.66u(2400),$$

or

$$0.34u(2400) > 0.33u(2500).$$

The responses to problem 2, however, suggest that for most people,

$$0.33u(2500) > 0.34u(2400),$$

the opposite inequality. Clearly, expected utility theory is missing something. Kahneman and Tversky's theory of preferences explains why people might make these choices, as well as several other choices which violate expected utility theory.

The theory has a number of important implications. The first is *loss aversion*: people dislike losing more than they like winning. In more technical terms, they lose more utility from a loss of x than they gain utility from an increase of the same amount of x. A related concept involves how people evaluate outcomes: they compare outcomes to a reference point, often caring more about whether a choice results in a gain or a loss, not about the resulting level of wealth that they end up with. Somewhat paradoxically, however, people tend to be risk averse when choosing among possible gains, as in problem 1 above, but risk loving when choosing among possible losses – they will often choose a gamble that has a chance to result in a big loss, but also has a chance to avoid losses altogether, over a certain loss of intermediate value.

A corollary of loss aversion is the *endowment effect*: people tend to place more value on items that they already have in their possession (i.e., that are already part of their endowment). Loss aversion explains this effect: people overvalue objects that they possess because the utility they lose from giving up what they already have is greater than the utility they gain from obtaining something that they did not have previously. Experimental studies that show this effect exists include Knetsch (1992), Kahneman, Knetsch and Thaler (1990), and Bateman et. al. (1997). For example, Knetsch endowed half of his subjects with a coffee mug, and half of his subjects with a pen. All subjects were then allowed to exchange their good for the other option if they so desired. If there is no endowment effect, we would expect half of the subjects to agree to the exchange. However, only 22 percent chose to trade for the other good, showing that most subjects valued the good they were initially randomly endowed with more highly. Recent evidence from field experiments suggests that as people gain market experience, they are less prone to endowment effects. List (2003) and (2004) both show that inexperienced collectors at a sports memorabilia show display an endowment effect, but dealers who are highly experienced in the market show no endowment effect at all.

Prospect theory also explains many *framing effects*, where people may make a different choice depending on how the options are presented. The most famous example is presented in Tversky and Kahneman (1981). In this hypothetical problem, respondents are told that 600 people are threatened by a disease, and they are asked to choose between two options. One set offers the options in a "positive frame":

Option A: save 200 lives for sure
Option B: a one-third chance of saving all 600 lives, and a two-thirds chance of saving 0 lives

Another set offers the options in a "negative frame":

Option C: 400 people die for sure
Option D: a two-thirds chance of all 600 dying, and a one-third chance of nobody dying

Despite the fact that the options result in the same outcomes with the same probabilities, most people chose option A over option B, but most chose option D over option C. When framed as a gain, people tend to prefer the safe option, but when framed as a loss, people tend to prefer the risky option.

Social Preferences: Fairness, Reciprocity, Altruism and Trust

The standard neoclassical model of behavior assumes that people are self-interested, and only care about maximizing their own wealth. Behavioral economists have assembled a large amount of evidence that there are exceptions to this rule. One exception concerns fairness and reciprocity: people often have a strong preference for being treated fairly. This preference is revealed by their exhibition in experimental studies of *negative reciprocity*: a willingness to sacrifice their own wealth in order to punish people who they believe have treated them unfairly. As Camerer and Loewenstein (2004) point out, the ultimatum game discussed in the introduction has been studied in over 20 countries, and the results have been consistent. Responders lower their own earnings in order to punish proposers by rejecting offers of less than one fifth of the available funds about half the time, and a large majority of proposers anticipate this rejection, offering between one third and one half of the available funds to the responder.

Studies have also shown that people exhibit *positive reciprocity*: a willingness to sacrifice their own wealth in order to reward people who they believe have been helpful or treated them fairly. This behavior has been observed in many different contexts. Fehr et. al. (1993) find that subjects cast in the role of a firm on average pay higher costs to increase product quality after buyers offer to buy at a higher, more fair price. Andreoni et. al. (2003) find that in an ultimatum game where responders have the option to reward as well as punish, the size of rewards increases with the amount offered to the responder. Charness (2004) finds that subjects placed in the role of workers are willing to exert more effort when employers initially offer higher wages.

Positive reciprocity is also observed in "trust" games, which have been studied by work such as Berg, Dickhaut and McCabe (1995) and Glaeser et. al. (2000). In a typical version of this game, a first mover is given ten dollars, and can choose to send any portion of it to a second mover. Each dollar that is sent is tripled by the experimenter, so if the first mover sends ten dollars, the second mover receives 30 dollars. The second mover can then return any amount to the first mover. Rational second movers would keep everything they receive for themselves, and anticipating this, rational first movers would send nothing. However, the studies show that first movers frequently send at least some money to the second mover, and second movers frequently reward first movers for trusting them by returning a larger percentage of what they receive when the amount they receive increases.

Even when there is no chance of punishment or reward, people exhibit altruism in "dictator" games, such as the ones studied in Hoffman et. al. (1994) and Eckel and Grossman (1996). In these games, one player, the "dictator," is given some amount of money, and is allowed to send some of it to a second player if they so choose, but the game ends there so there are no possible financial consequences to sending the second player nothing. Despite this fact, subjects frequently send substantial amounts to the second player. In the experiments of Hoffman et. al. (1994) the dictator was given $10 and knew the second player's identity. Out of 48 dictators, 14 sent $3, 5 sent $4, and 3 sent $5. The rest sent less than $3.

The motivation behind these behaviors is not entirely clear. Charness (2004) shows, for example, that the positive reciprocity he observes is due both to a desire to reward employers for fair treatment, as well as a desire to keep payoffs equal regardless of the actions of others. Several theories have been offered as explanations. Rabin (1993) presents a model where

people gain psychological payoffs from kind treatment. Fehr and Schmidt (1999) suggest a model of inequality aversion where people find all inequality objectionable, but their distaste for inequality is stronger if they are the ones at a disadvantage. Finally, Charness and Rabin (2002) present a "quasi-maximin" model where people prefer the payoffs of the worst-off individual to be as high as possible, but also are concerned with the size of total payoffs for all individuals.

IV. APPLICATIONS OF PSYCHOLOGY IN ECONOMICS

The recent advances in behavioral economics stemming from increased use of psychological motivations have many applications to economic decision making. This section will discuss work in the areas of money illusion, fair wages and unemployment, intertemporal substitution, rising consumption profiles over time, equity premiums, real estate markets and environmental economics.

Money Illusion

Money illusion refers to individuals' decision making being influenced by nominal (rather than real) changes in variables. Nominal values of a variable are those actually observed, while "real" values take into account changing prices. Economic theory holds that individuals should only consider "real" values when making decisions, such as those regarding salaries and contracts. However, even early on, economists observed that individuals did not always base their decisions on real values. The idea that nominal accounting methods (i.e. inflation causing a change in nominal but not real values) can impact decision making has been cited by economists throughout the 20th century (Fisher, 1928; Patinkin, 1965; Fischer and Modigliani, 1986).

Early cases of money illusion observed by economists included downward rigidities in nominal wages. Individuals resist decreases in nominal wages more than they resist the same change in their real wage being caused by inflation. In 1936, Keynes noted

> "Since there is imperfect mobility of labour, and wages do not tend to an exact equality of net advantage in different occupations, any individual or group of individuals, who consent to a reduction of money-wages relatively to others, will suffer a relative reduction in real wages, which is a sufficient justification for them to resist it. On the other hand it would be impracticable to resist every reduction in real wages, due to a change in the purchasing-power of money which affects all workers alike; and in fact reductions of real wages arising in this way are not, as a rule, resisted unless they proceed to an extreme degree." (Keynes, 1936).

While Keynes and others noted this phenomenon, they did not have the psychological insights necessary to examine the motivations behind this behavior. While even economists in the early 20th century observed money illusion, they did not have the means to satisfactorily explain why it occurred. The psychological insight necessary for explaining instances of money illusion comes from framing effects.

Framing effects refers to individuals having different responses to the same situation if it is presented in a different manner to them (Tversky and Kahneman, 1981; Kahneman and Tversky, 1984). It has been observed that nominal wages rarely fall, even in times of recession (Bernanke and Carey, 1996; Bewley, 1998; Kahn, 1997). Solow's theory of efficiency wages proposes that individuals react to changes in their nominal wage, not just their real wage (Solow, 1979). If people observe their nominal wage falling then they may put forth less effort, lowering productivity and reducing any gain which the firm made in lowering wages. Solow proposes that an individual faced with a real wage falling by 2% will react better to a scenario in which nominal wages rise by 3% in a period of 5% than they will if nominal wages fall by 1% in a period of 1% inflation. The fall in real wages is 2% in either case but the individual perceives the situation differently and so has a different response regarding work effort. This is explained by framing effect but not by the "rational" economic view of the individual.[1]

A second observation of money illusion comes in the form of contracts. In particular, if individuals were concerned with "real" values then contracts for wages and loans should be indexed to inflation (so that the real value does not change). However, except under periods of high inflation, it is rare to see indexed contracts. The one exception to this tends to be union contracts which are often indexed. However, only 12% of the American labor force is unionized (CPS, 2007).

The Fair Wage-Effort Hypothesis and Unemployment

Psychological insights concerning with reference groups and relative deprivation has lead to advances in economic theories of worker productivity and unemployment. Akerlof and Yellen's (1990) work specifically deals with how worker effort is impacted by their feelings of "fair" treatment. This model stood in contrast to the prevailing orthodoxy in economics which held that worker productivity was impacted only by variables such as education, capital and technology. Traditional economics had no role for factors such as workers' emotional responses to the perception of their relative treatment. It should be noted that Akerlof and Yellen proposed their model as an expansion of current models. They also incorporated traditional economic variables into their model.

The basis of the fair wage-effort model is relative deprivation theory. Relative deprivation proposes that individuals view their situation as compared to a reference group (Runciman, 1966). In this application, individuals view their wages in reference to wages of their peers. Worker effort becomes a function of how people view their relative wages. If people believe they are relatively deprived with regards to wages then they will seek to punish the firm by reducing their work effort. Individuals will match their effort to their perceived treatment. Also, individuals who believe they are being paid relatively well may put forth additional effort.

This theory can help to explain why firms might pay above market clearing levels. On an economy-wide basis, this can result in involuntary unemployment. If hiring workers at market clearing levels would result in them being paid lower than their peers either at the firm or in

[1] For a more comprehensive and detailed survey of research on money illusion, readers are directed to Shafir, Diamond and Tversky (1997).

the industry then these workers may put forth less effort. This lower effort increases the marginal cost of production to the point at which the firm actually could have lowered its costs by paying its workers more.

Intertemporal Substitution

One of the commonly held views in labor economics has been that people will choose to work more hours if they are offered a higher wage (Lucas and Rapping, 1969). The higher wage increases the opportunity cost of leisure time so people decide to work more hours (and consequently earn and consume more). Despite the popularity in mainstream economics of this proposition, the results from empirically testing it have been mixed (Altonji, 1986; Browning, Deaton and Irish, 1985; Laisney, Pohlmeier and Staat, 1992; Mankiw, Rotemberg and Summers, 1985; Mulligan, 1995). Thus, this view has recently been challenged by evidence suggesting that people do not make labor supply decisions in such a manner.

An alternative view is that people work to achieve income/consumption targets and will work whatever hours are necessary to achieve these goals. Camerer et al., (1997) tracked the hours worked by cab drivers on days with more fares (resulting in higher hourly wages) and those with less fares to see how their choice of hours worked was altered (cab drivers in New York have flexibility in choosing the number of hours worked).[2] The result which they found is that people did not work more on days with a higher per hour wage; rather, they actually worked more hours on days with a lower hourly wage. This is the opposite prediction of traditional economic theory. However, if drivers were targeting a daily income number then it is completely logical that they would work less on more lucrative days. Drivers also seemed to exhibit loss aversion as they would work considerably longer hours to avoid missing their target but would not work many more hours to increase gains above the target on good days. Daily targeting is also consistent with the idea that individuals are concerned with a possible lack of self-control (Shefrin and Thaler, 1988). Without daily targets, individuals may work less hours assuming that they would then make up the income on another day but then fail to do so.

Seniority and Rising Consumption Profiles

Economists have sought to explain why wages rise faster with seniority than does worker productivity (Lazear, 1979). Explanations that involve firm-specific training and productivity changes with experience have failed to account for the differences seen between junior and senior employees at firms. While traditional economics cannot explain this, it is completely consistent with a view of the individual that incorporates “demonstration effects”.

Demonstration effects propose that individuals attempt to match the consumption of their peers. This results partially from a lack of self-control by individuals. Showing the influence of psychology on economics early on, this idea goes back in economics to the work of Duesenberry and his relative income hypothesis (Duesenberry, 1949). According to this idea,

[2] Similar types of studies have been done with farmers (Berg, 1961; Orde-Brown, 1946) and with self-employed proprietors (Wales, 1973).

individuals attempt to match the consumption of their peers, regardless of whether or not they can afford this consumption. Duesenberry popularized the phrase "keeping up with the Joneses" in explaining this phenomenon.

If individuals react in such a manner then this is problematic for lifetime savings if wages are flat or decrease over time. In such cases, individuals would overconsume when young and would not have sufficient income later to save for retirement. Therefore, workers prefer a consumption profile that rises over time with seniority. Rising earnings profiles can be seen as a form of forced savings (Frank and Hutchens, 1993; Lowenstein and Sicherman, 1991; Neumark, 1995).[3] This is preferred even though productivity may not be rising as much over time, therefore workers are willing to accept lower wages when younger for the promise of higher wages later in life.

The Equity Premium Puzzle

Economists have long wondered why people buy bonds. This is an apt question as the long-run return to bonds is much lower than that of stocks. There would be some difference expected in returns because of the additional risk of stocks (individuals are risk averse). However, this difference in returns is far higher than the difference predicted by theories of risk aversion. The unexplained difference in returns between stocks and bonds is called the equity premium. This puzzle was first coined by Mehra and Prescott (1985). This phenomenon was found in further studies using earlier stock and bond data (Siegel, 1992a,b; Weil, 1989) as well as in studies using international data (Campbell, 1996; Gielen, 1994; Hirose and Tso, 1995).

Recent work has sought to explain this puzzle using prospect theory (Benartzi and Thaler, 1995; Siegel and Thaler, 1997). As discussed earlier, prospect theory is based on loss aversion and mental accounting (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). These properties lead individuals to prefer returns that exhibit low variability, even in the short-run. Individuals want low variability because they are loss averse (Kahneman, Knetsch and Thaler, 1990; Tversky and Kahneman, 1991). The impact of this loss aversion on its own is not enough to explain the entire equity premium. However, the mental accounting aspect of prospect theory considers what the length of evaluation period individuals have (Kahneman and Tversky, 1984; Thaler, 1985). In particular, if individuals have a short length of evaluation then they will periodically review their portfolios and at any given point stocks may not be outperforming bonds. In fact, it is not uncommon for stocks to have losses in the short-term. So, the recurring short evaluation periods combined with loss aversion results in individuals still purchasing significant amounts of bonds despite the much lower returns.

[3] Other types of forced savings mechanisms such as Christmas clubs are discussed by Loewenstein and Thaler (1989), Loewenstein and Prelec (1992), Stigler (1966) and Thaler and Shefrin (1981).

Other Asset Markets: Real Estate

Given the scale of the real estate market, the behavior of buyers and sellers is of significant interest to economists. Economists have observed a negative relationship between selling prices and volume that could not be explained by "rational" economic behavior (Genesove and Mayer, 2001; Ortalo-Magne and Rady, 1998; Stein, 1995). However, prospect theory has offered a viable explanation for this result.

It has been observed that when prices fall in the housing market, many individuals will refuse to sell their home. The inventory of unsold properties increases as individuals become unwilling to sell their home "at a loss". Individuals will continue to offer their homes at selling prices well above the market rather than sell at market prices.

Prospect theory can explain this behavior through reference points and loss aversion. When individuals are selling their homes, they view the selling price relative to a reference point. This is often, but not necessarily, their original purchase price. This reference could also be a previous high point in the market. Individuals feel that if they sell the property for less than this reference point then they are selling "for a loss". Prospect theory also holds that people exhibit loss aversion so they feel losses more than they do gains (Tversky and Kahneman, 1991). Therefore, individuals become averse to sell their property for what they perceive to be even a small "loss."

Environmental Economics

A cornerstone of environmental economics is the contingent valuation (CV) method, which is used to estimate the value of all kinds of environmental services and benefits. These estimates are often used in an analysis of the costs and benefits of a policy that may cause environmental damage. The method involves conducting a survey to assess the monetary value people place on the environmental damage that would result from the policy. One major criticism of this method is related to the endowment effect. Depending on how the question is phrased, the reported valuation of the environmental benefits of a policy can be very different. Questions that ask how much you would be *willing to pay* (WTP) for a particular benefit typically elicit lower values than questions that ask how much you would be *willing to accept* (WTA) to give up the same benefit. Hammack and Brown (1974) found in a survey of duck hunters that they would be willing to accept the destruction of a wetland needed to support the duck population at its current level in return for $1,044 on average, but they would be willing to pay only $247 on average to maintain the wetland. In an experimental study, Knetsch and Sinden (1984) estimate that the difference between the amount subjects were willing to pay for a lottery ticket was $1.28, but the price that subjects had to be paid before they were willing to give up the same ticket when already endowed with it was $5.18. This wide divergence in value measurement based on the phrasing of the question has highlighted the importance of identifying whether the appropriate measure in each particular context is willingness to pay or willingness to accept, and in general has called the validity of the CV method into question.

V. CONCLUSION

The neoclassical economic view of decision making has encountered several challenges in recent years. These challenges have been motivated by empirical observations of phenomenon that neoclassical models of decision making cannot explain. In particular, there has been an increasing use of psychological motivations in explaining individual behavior that does not fit the predictions of the neoclassical "rational agent".

These advances in behavioral economics have crossed many fields in economics. For example, economists' views on topics such as unemployment, wages, work hours, stock prices, real estate values and environmental valuations have all been impacted by the increased use of psychological motivations for human behavior. While these may seem to be disparate threads of research there are some common themes. Ideas such as prospect theory, fairness and reciprocity, and relative deprivation/consumption can be seen as running themes which impact several fields in economics. These themes represent an alternative view of decision making and therefore have a broad impact in economics, as individual decision making is at the core of economics. Therefore, advances in decision making such as prospect theory will no doubt continue to have a wide-ranging impact on the discipline of economics in the future.

REFERENCES

Akerlof, George and Janet Yellen. 1990. "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, 105, 255-283.

Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Microdata," *Journal of Political Economy*, 94(3), S176-S215.

Andreoni, James, William Harbaugh and Lise Vesterlund. 2003. "The Carrot or the Stick: Rewards, Punishments, and Cooperation," *The American Economic Review*, 93(3), 893-902.

Bateman, Ian, Alistair Munro, Bruce Rhodes, Chris Stamer and Robert Sugden. 1997. "A Theory of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 112(2), 479-505.

Benartzi, Shlomo and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1),.75-92.

Berg, Elliot J. 1961. "Backward-Sloping Labor Supply Functions in Dual Economies-The Africa Case," *Quarterly Journal of Economics*, 75, 468-492.

Berg, Joyce, John Dickhaut and Kevin McCabe. 1995. "Trust, Reciprocity and Social History," *Games and Economic Behavior*, 10, 122-142.

Bernanke, Ben S. and Kevin Carey. 1996. "Nominal Wage Stickiness and Aggregate Supply in the Great Depression," *Quarterly Journal of Economics*, 111(3), 853-883.

Bewley, Truman F. 1998. "Why Not Cut Pay?" *European Economic Review*, 42(3-5), 459-490.

Browning, Martin., Angus Deaton and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle," *Econometrica*, 53, 503-543.

Camerer, Colin F., Linda Babcock, George Loewenstein and Richard H. Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 112(2), 407-441.

Camerer, Colin F. and George Loewenstein. 2004. "Behavioral Economics: Past, Present and Future," in Camerer, Loewenstein and Rabin, eds., *Advances in Behavioral Economics* (New York: Russell Sage Foundation).

Camerer, Colin F. and Richard H. Thaler. 1995. "Anomalies: Dictators, Ultimatums, and Manners," *Journal of Economic Perspectives*, 9, 209-219.

Campbell, John. 1996. "Consumption and the Stock Market: Interpreting International Experience," NBER Working Paper No.5619.

Charness, Gary. 2004. "Attribution and Reciprocity in an Experimental Labor Market," *Journal of Labor Economics*, 22, 665-688.

Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817-869.

Current Population Survey. 2007. Percentage of Workforce Unionized. Accessed online at http://www.bls.gov/cps/cpsaat40.pdf.

Duesenberry, J.S. 1949. *Income, Savings and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.

Eckel, Catherine C. and Philip J. Grossman. 1996. "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 16, 181-191.

Fehr, Ernst, Georg Kirchsteiger and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics*, 108(2), 437-459.

Fehr, Ernst and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817-867.

Fisher, Irving, 1919. *Elementary Principles of Economics*. New York: Macmillan.

Fischer, Stanley and Franco Modigliani. 1986. "Towards and Understanding of the Real Effects and Costs of Inflation," *Weltwirtschaftliches Archiv*, 114, 810-833.

Frank, Robert H., and Robert M. Hutchens. 1993. "Wages, Seniority and the Demand for Rising Consumption Profiles," *Journal of Economic Behavior and Organization*, 21, 251-276.

Genesove, David. and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116 (4), 1233-1260.

Gielen, Gregor. 1994. *Konen Aktienkurse noch Steigen?* Wiesbaden, Germany: Gabler.

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman and Christine L. Soutter. 2000. "Measuring Trust," *Quarterly Journal of Economics*, 115(3), 811-846.

Gneezy, Uri and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All," *Quarterly Journal of Economics*, 115(4), 791-810.

Hammack, Judd and Gardner Mallard Brown. 1974. *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Baltimore: Johns Hopkins Press for Resources for the Future.

Hertwig, Ralph and Andreas Ortmann. 2001. "Experimental Practices in Economics: A Methodological Challenge for Psychologists," *Behavioral and Brain Sciences*, 24(4), 383-403.

Hirose, Hiroaki., and Yi Tso. 1995. "Japanese Market Returns," mimeograph, The Wharton School of the University Of Pennsylvania.

Hoffman, Elizabeth, Kevin McCabe, Keith Shachat and Vernon Smith. 1994. "Preferences, Property Rights and Anonymity in Bargaining Games," *Games and Economic Behavior*, 7, 346-380.

Hume, David. 1740. *Treatise on Human Nature*. New York: Oxford University Press, 2000.

Kahn, Shulamit. 1997. "Evidence of Nominal Wage Stickiness from Microdata," *American Economic Review*, 87(5), 993-1008.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98, 1325-1348.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.

Kahneman, Daniel and Amos Tversky. 1984. "Choices, Values and Frames," *American Psychologist*, 39, 341-350.

Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan; New York. Harcourt Brace.

Knetsch, Jack L. 1992. "Preferences and Nonreversibility of Indifference Curves," *Journal of Economic Behavior and Organization*, 17(1), 131-139.

Knetsch, Jack 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(3), 507-521.

Laisney, Francois., Winfried Pohlmeier and Matthias Staat. 1992. "Estimation of Labor Supply Functions Using Panel Data: A Survey," in Matyas and Sevestre, eds., *The Economics of Panel Data: Handbook of Theory and Applications* (Dordrecht, the Netherlands: Kluwer, 1992).

Lazear, Edward. 1979. "Why is There Mandatory Retirement?" *Journal of Political Economy*, 87, 1261-1284.

List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118, 41-71.

List, John A. 2004. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2), 615-625.

Loewenstein, George. 1988. "Frames of Mind in Intertemporal Choice," *Management Science*, 34, 200-214.

Loewenstein, George and Nachum Sicherman. 1991. "Do Workers Prefer Increasing Wage Profiles," *Journal of Labor Economics*, 9, 67-84.

Loewenstein, George and Drazen Prelec. 1992, "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107(2), 573-597.

Loewenstein, George and Richard H. Thaler. 1989. "Anomalies: Intertemporal Choice," *Journal of Economic Perspectives*, 3, 181-193.

Lucas, Robert E., and Leonard A. Rapping.1969. "Real Wages, Employment and Inflation," *Journal of Political Economy*, 77, 721-754.

Mankiw, N.Gregory., Julio J. Rotemberg and Lawrence H. Summers. 1985. "Intertemporal Substitution in Macroeconomics," *Quarterly Journal of Economics*, 100, 225-251.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-162.

Metzler, Lloyd A. 1951. "Wealth, Saving, and the Rate of Interest," *Journal of Political Economy*, 59, 93-117.

Mulligan, Casey B. 1995. "The Intertemporal Substitution of Work – What Does the Evidence Say?" University of Chicago, Population Research Center Discussion Paper Series #95-11.

Neumark, David. 1995. "Are Rising Earnings Profiles a Forced-Savings Mechanism?" *Economic Journal*, 105(428), 95-106.

Orde-Brown, G. 1946. *Labour Conditions in East Africa*. London: Colonial Office, H.M.S.O.

Ortalo-Magne, Francois and Sven Rady. 1998. "Housing market fluctuations in a life cycle economy with credit constraints," Stanford University GSB Research paper 1501.

Patinkin, Don. 1965. *Money, Interest and Prices*. New York: Harper and Row.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83(5), 1281-1302.

Runciman, Walter G. 1966. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. University of California Press.

Shafir, Eldar., Peter Diamond and Amos Tversky. 1997. "Money Illusion," *Quarterly Journal of Economics*, 112, 341-374.

Shefrin, Hersh M. and Richard H. Thaler. 1988. "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 26, 609-643.

Siegel, Jeremy J. 1992a. "The Equity Premium: Stock and Bond Returns Since 1802," *Financial Analyst Journal*, 48, 28-38.

Siegel, Jeremy J. 1992b. "The Real Rate of Interest from 1800-1990: A Study of the U.S. and U.K.," *Journal of Monetary Economics*, 29, 227-252.

Siegel, Jeremy J. and Richard H. Thaler. 1997. "Anomalies: The Equity Premium Puzzle" *Journal of Economic Perspectives*, 11(1), 191-200.

Solow, Robert M. 1979. "Another Possible Source of Wage Stickiness," *Journal of Macroeconomics*, 1, 9-82.

Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Constraints," *Quarterly Journal of Economics*, 110, 379-406.

Stigler, George J. 1966. *The Theory of Price*. New York: Macmillan.

Thaler, Richard H. 1985. "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199-214.

Thaler, Richard H. and H.M. Shefrin. 1981. "An Economic Theory of Self-Control," *Journal of Political Economy*, 89(2), 392-406.

Tversky, Amos and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453-458.

Tversky, Amos and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics*, 106(4), 1039-1061.

Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.

Wales, Terence J. 1973. "Estimation of a Labor Supply Curve for Self-Employed Business Proprietors," *International Economic Review*, 14, 69-80.

Weil, Philippe. 1989. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401-421.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 179-193

Chapter 8

WHY DO PEOPLE CHANGE THEIR MINDS? EVIDENCE FROM THE PURCHASE OF LONG-TERM CARE INSURANCE

Anne Theisen Cramer*[1] and Gail A. Jensen[2]
[1]Davenport University, 27650 Dequindre Road, Warren, MI, 48092
Institute of Gerontology, Wayne State University, Detroit, MI, 48202
[2]Institute of Gerontology, Wayne State University,
87 E. Ferry Street, Detroit, MI, 48202

ABSTRACT

Today more than ever, people try to anticipate financial needs and to plan wisely for a lifetime of financial security. Information about financial options is plentiful, and financing for health and long-term care (LTC) is no exception. With all of the information and advice that is available, under what circumstances would a person decide that his/her decision was no longer the best option?

We address this question by looking at the market for LTC insurance, and estimate logistic regressions to model consumer decisions to drop or renew an existing LTC insurance policy. We explore events that occurred after the policy was last purchased and before the current policy was dropped or renewed. The price and benefit design of each policy is not directly observable so several proxy measures of the price of a policy are explored.

Data is obtained from the publicly available Health and Retirement Survey (HRS). Data from 2002 is used to identify those who have a LTC policy and to establish baseline financial circumstances. Data from 2004 is used to determine whether the policy was renewed, and to identify potentially influential events that occurred since 2002.

The study sample includes 1,375 individuals who reported an existing, private LTC insurance policy in 2002, and were therefore eligible to renew the existing policy before 2004. Proxy prices were calculated and assigned using publicly available price schedules.

Preliminary findings suggest that price was an influential factor in the decision to drop an existing policy, even though the price of the policy did not increase as a result of

* Email: anne.cramer@davenport.edu. Or Email: atcramer@wayne.edu.

age. Those with newer policies were less likely to allow a policy to lapse. Those with low levels of assets (less than $200,000) were more likely to allow a policy to lapse, as were those with more than $1.5 million in assets.

Our results suggest that financial considerations are important, and a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC.

INTRODUCTION

One of the biggest challenges facing Americans today is that of planning their retirement. Life expectancy is longer, retirement age is changing, and medical advances are providing opportunities that did not exist a few years ago. In 2004, life expectancy at birth was 77.8 years (National Center for Health Statistics, 2006), a full seven years higher than in 1970, and the average duration of retirement was 17.4 years, four years longer than in 1970 (Gendel, 1998). In order to adequately prepare for these retirement years, financial advisors are recommending earlier planning, and are recommending that the planning specifically address potential health care needs (Vanderhei, 2006; Sahadi, 2006), including long-term care (LTC).

In 1960, the majority of LTC spending was for nursing home care, with paid long-term care (LTC) accounting for only 3.1% of national health expenditures (National Center for Health Statistics, 2006). Today, LTC can be provided in a variety of settings including the home, and LTC spending accounts for 8.4% of national health expenditures (National Center for Health Statistics, 2006). LTC insurance policies have been marketed since the 1980s, but relatively few adults purchase them, and of those who purchase policies, many allow their policies to lapse before services are used. The industry association AHIP (America's Health Insurance Plans, 2004) reports that only about seven out of every ten purchased policies remains in force. In addition, industry underwriters cite lapse rates that have been as high as 7% per year, and approximate 2% per year in recent years (Morisato, 2004; O'Brien, 2004).

The potential for lapsed policies is one of great concern. Advisors typically warn seniors against purchasing LTC policies that they cannot afford. CNN and Money Magazine editors advise readers "Unless you're confident that you can afford the premiums for the long haul, don't sign on. You could waste thousands on a policy that lapses before you need it." (Feldman, et. al., 2000). If a policy is allowed to lapse, not only will the individual be without insurance coverage for LTC, but he or she will also be without the money paid as premiums. So why might an individual purchase a LTC policy, just to let it lapse at some future date? What factors might cause an individual to reverse an initial decision to purchase and decide that LTC insurance is no longer the best option?

BACKGROUND

LTC is assistance with basic care and can include nursing care, skilled care such as physical therapy and occupational therapy, personal care, custodial care, and household services. LTC insurance typically covers care when the insured becomes cognitively impaired or can no longer perform two or more standard "activities of daily living" (ADL's) such as dressing, eating, or bathing. The price of most LTC insurance is based on the age of the

purchaser at the time of purchase and does not increase as a result of aging, providing a strong incentive to continue to renew the policy once it is in force.

Studies regarding long-term care insurance have typically focused on the initial decision to purchase LTC insurance, identifying influential factors such as information and knowledge, family circumstances, and availability of informal care. Several of these prior works are used as bases for empirical tests in this paper. The basic health insurance purchase decision is modeled by McKenna (1986). The purchase of long-term care insurance is described theoretically in works by Pauly (1990) and by Zweifel and Strüwe (1998). Gupta and Li (2004) model the LTC insurance purchase decision as an optimization problem across two periods (pre- and post-retirement).

In addition to studies that focus on the initial purchase decision, two recent studies have focused specifically on lapsed LTC policies using early HRS data. Finkelstein and McGarry (2005) acknowledge the high lapse rate – over 25% over five years in their sample – and suggest that information that becomes available after the initial purchase might induce subscribers to drop their policies. McNamara and Lee (2004) look at HRS data, and find high lapse rates, with only 23.2 percent of subscribers in their sample keeping their coverage over the five-year period of the study. They suggest that one explanation is a lack of accurate information about LTC insurance. Another is inaccurate information about the risk of needing LTC.

In the study presented here, we again consider long-term care insurance lapses. We focus on whether or not the consumer allows his or her LTC insurance policy to lapse, given that a policy is already in force. We include price as a determinant of the decision to allow an existing LTC policy to lapse, along with other potential determinants of this decision. We also use the newly introduced validation question in the Health and Retirement Survey to verify that respondents are not mistakenly describing Medicare as a long-term care policy.

METHODS

Data

Data for the study are drawn from the Health and Retirement Study (HRS), an on-going, nationally representative survey of older adults in the U.S. (Health and Retirement Study, 2004). In 2004, the survey included 20,147 respondents. Each respondent was born in 1953 or earlier (age 51 or older in 2004), or was married to a respondent who was born in 1953 or earlier.

In this study, we focus on the decision to allow an existing LTC insurance policy to lapse. We include in our sample those who initially purchased a policy prior to 2002, and report that the policy is still in force in 2002. By including these individuals, we focus on policies that are comparable to those offered today, and we focus on potential repeat purchasers. We eliminate from the sample persons for whom essential data was missing, yielding a final analytic sample of 1,375 individuals who were eligible to renew an existing LTC insurance policy in 2004.

Theoretical Framework

We use a simple two period model of the decision to purchase long-term care insurance that takes into account the intertemporal nature of such policies. Consider an individual who is deciding whether to renew an existing insurance policy to cover the risk of a loss that could occur in the current period, the future, or both. The individual is risk averse with preferences defined by a von Neumannn-Morgenstern utility function u for consumption in each period. Her risk of a loss, L, increases over time. A loss occurs in the current period with probability p_1, and it occurs in the future with probability p_2, where p_1 is less than p_2. Assume this individual is endowed with W_1 units of wealth today, and W_2 units in the future. Absent insurance, her ex-ante expected utility is given by:

$$EU_0 = [p_1u(W_1–L)+(1–p_1)u(W_1)] + b[p_2u(W_2–L)+(1–p_2)u(W_2)] \quad (1)$$

where b reflects her valuation of tomorrow's consumption compared to today's.

Suppose that a market for LTC insurance contracts exists in which there are willing suppliers of policies that pay an indemnity Y should a loss occur. Policies are renewable and there is a fixed premium-per-period, h, at which such contracts are traded. As long as h is paid to the seller each period, full coverage remains in force. One distinguishing feature of LTC insurance policies is that h remains constant over time. This model reflects this feature.

If the individual purchases a contract, her ex-ante expected utility will be

$$EU_1 = [p_1u(W_1–h–L+Y)+(1–p_1)u(W_1–h)] + b[p_2u(W_2–h–L+Y)+(1–p_2)u(W_2–h)] \quad (2)$$

Clearly, she will purchase a policy today if and only if $EU_1 > EU_0$. From this simple model we see that a number of factors are likely to influence the decision to renew the policy. The price, h, and the characteristics of coverage, Y, are clearly relevant. The decision also depends on an individual's resources (W_1 and W_2), her rate of time preference (b), the likelihood of incurring a loss and the magnitude of that loss (p_1, p_2, and L), as well as the individual's attitudes towards risk, which are reflected in $u( )$. These considerations guide the selection of variables in our empirical work. Factors that influence the individual's attitudes toward risk and perceived likelihood of incurring a loss are relevant to the model, as discussed in prior studies (Pauly, 1990; Zweifel and Strüwe, 1996, 1998; Sloan and Norton, 1997; McCall et al., 1998; Finkelstein and McGarry, 2004; McNamara and Lee, 2004). We also recognize that information may have become available and events may have occurred after the initial decision to purchase and may therefore prompt the individual to change her decision. In our model, the consumer will allow LTC insurance to lapse if her expected utility in the insured state (EU_1) falls below expected utility in the uninsured state (EU_0). More formally, we model the decision to allow LTC insurance to lapse by estimating variations of the equation:

$$D_i^* = \beta_0 + \beta_1 P_i + \beta_2 X_i + \beta_3 W_i + \beta_4 E_i + \varepsilon_i \quad (3)$$

where i indexes the individual. If $D_i^* > 0$, the individual chooses to let their LTC policy lapse, whereas if $D_i^* < 0$ the individual renews their policy. P_i is the annual premium to be incurred,

X_i is a vector of individual characteristics, W_i is a measure of financial variables including wealth and income, E_i is a vector of events that have occurred since the policy was last reported in force, the β's are vectors of coefficients, and ε_i is a random disturbance. We do not observe D_i^* directly. Rather, we observe whether each LTC policy holder has actually let their coverage lapse. Define $D_i = 1$ if the policy has lapsed, and $D_i = 0$ if it has not. D_i is the variable we observe. We assume ε_i follows a logistic distribution and estimate the parameters of (3) by maximum likelihood.

In 2002 and 2004, HRS participants were asked about basic health insurance (both private and government). They were then asked, "Not including government programs, do you now have any long term care insurance which specifically covers nursing home care for a year or more or any part of personal or medical care in your home?" If the individual answered, "Yes," then they were asked, "Is that one of the plans you have already described, or a different plan?" For our study, we include the individual in our sample if he replied "Yes" to the first question in 2002, indicating a LTC policy, and then replied "a different plan" to the second question, indicating that the LTC policy under discussion was different from his Medicare, Medicaid, or basic health insurance policy. We then use responses to the same questions in 2004 to determine whether the policy was renewed or allowed to lapse in 2004.

Overall, about 15% of people in our sample allowed a private LTC insurance policy to lapse in 2004. Of the 1375 observations in our sample, 1171 individuals decided to renew LTC insurance, and 204 decided to allow the policy to lapse. Based on data reported by many individual insurers, AHIP (America's Health Insurance Plans) estimates that 30% of all policies purchased are no longer in force. In our sample, only 15% of policies have lapsed, probably for two reasons. First, the AHIP rate includes policies that have lapsed due to death, while our sample includes only those who are living and choose not to renew. Second, the AHIP rate is cumulative, while we examine lapses over only the two-year period from 2002 to 2004.

Explanatory Variables

Explanatory variables fall into four general categories: price and characteristics of the policy, individual characteristics, financial variables, and recent events. A list of explanatory variables and descriptive statistics is presented in Table 1.

For this analysis the price of insurance, P, is measured by the annual premium the individual would have faced when the LTC policy was first purchased. For each respondent in the sample, the premium was calculated and assigned on the basis of the individual's age and health status at the time of initial purchase, using publicly available price schedules for LTC coverage (described below). The year of initial purchase was determined using self-reported duration of the contract. Consistent with standard insurance company practices, price discounts were given for non-smokers who did not report an identified pre-existing condition. Using this as our measure allows us to assign the price of a standard policy to all potential purchasers. This methodology has been used in past studies regarding health insurance (Leibowitz and Chernew, 1992) and term life insurance (Pauly et al. 2003, Brown and Goolsbee 2003), where prices were assigned based on industry practices.

Table 1. Descriptive Characteristics of Potential Lapses (N=1375)

Variable		Mean	Standard Deviation	Expected Sign
Dependent Variable				
	Lapse: Equal to "1" if LTC insurance is Lapsed (not renewed) in 2004	0.15	0.36	
LTC Insurance				
	G E Capital Price	2155.98	1357.28	+
	AHIP Price	2404.00	1439.05	+
	TIAA Price	3244.53	1512.11	+
	Years Contract Was In Force	7.60	5.10	+
Demographic Variables				
	Male	0.41	0.49	+
	White	0.96	0.19	?
	Black	0.03	0.17	?
	Age	71.56	8.58	–
	Married or Partnered	0.73	0.45	?
	Number of Children	2.89	1.85	?
	Education in Years (Max = 17, graduate level)	13.97	2.51	–
Geographic Variables				
	Urban/Suburban = 1; Rural = 0	0.71	0.45	–
	New England (ME, NH, VT, MA, RI, CT)	0.04	0.20	
	Mid Atlantic (NY, NJ, PA)	0.09	0.29	
	East North Central (OH, IN, IL, MI, WI)	0.16	0.37	
	West North Central (MN, IA, MO, ND, SD, NE, KS)	0.18	0.38	
	South Atlantic (DE, MD, DC, VA, WV, NC, SC, GA, FL)	0.23	0.42	
	East South Central (KY, TN, AL, MS)	0.05	0.22	
	West South Central (AR, LA, OK, TX)	0.06	0.24	
	Mountain (MT, ID, WY, CO, NM, AZ, UT, NV)	0.05	0.21	
	Pacific (WA, OR, CA, AK, HI)	0.13	0.34	

Table 1. Descriptive Characteristics of Potential Lapses (N=1375) (Continued)

Variable		Mean	Standard Deviation	Expected Sign
Health				
	Health Status: Equal to "1" if self-reported status is at least "good."	0.84	0.37	?
	ADL Count: Number of ADL's reported as needing help (maximum = 10).	1.95	2.28	–
Finances				
	High Asset Level: Equals "1" if household assets > $1.5 million.	0.06	0.25	+
	Low Asset Level: Equals "1" if household assets < $200,000.	0.38	0.49	+
	Real Estate Equity ($000's)	90	81	?
	Income ($000's). Household income divided by 2 if married; divided by 1 if not married.	51	114	–
Recent Events				
	Recently Moved Residence	0.14	0.34	+
	Spouse Recently Died	0.03	0.18	+
	Parent Recently Died	0.03	0.18	–
	Recently Married	0.01	0.08	–
	Health Improved	0.73	0.44	+
Other				
	Uses email and internet	0.45	0.50	–
	Satisfied: Satisfaction with health insurance. (Equals "1" if very satisfied, "0" if not satisfied.)	0.97	0.16	–
	Has a Will	0.85	0.36	–

In this study, we consider price (premium) variables based on three different industry models. The first variable, “GE Capital Price” is based on rate schedules filed by GE Capital, a major insurer in the LTC insurance market (State of Michigan, 1996). The second, called “AHIP Price,” is based on a table of average rates nationwide for LTC insurance published by America’s Health Insurance Plans (AHIP, 2004). The third, "TIAA Price," uses a mathematical formula developed by Gupta and Li using premium data from TIAA- CREF (Gupta, 2004) and assuming a standard set of benefits. It is expected that price will be positively correlated with the probability of allowing a LTC insurance policy to lapse. We also control for the age of the policy.

In addition to price and the policy’s age, explanatory variables also include individual characteristics that are unlikely to have changed since the original purchase decision was

made, individual characteristics that may change, and events that may result in a decision to stop renewing a LTC policy.

Characteristics that are unlikely to have changed since a LTC policy was first purchased include gender, race, and level of education. Gender is included as a dummy variable. It is expected that men will be more likely to allow a LTC policy to lapse because of a lower life expectancy, implying that they may need LTC for less time. Race is included as a set of dummy variables in order to capture cultural differences that might impact the decision to renew LTC insurance. According to a national survey conducted for AARP (Belden, 2001), white individuals are less likely to provide informal care if LTC is needed, relying instead on paid formal care. In addition, prior studies (Pandya, 2005; Wallace, 1998) have shown that non-white cultures have traditionally relied more on family members to provide informal LTC. If this is the case, white individuals might be less likely to allow a LTC policy to lapse.

Education is expected to be correlated with a lower probability of lapsing LTC insurance for a number of reasons. First, education can potentially improve efficiency of health production (Kenkel, 1991), improving the individual's ability to make a lasting decision about LTC insurance when the policy is first purchased. Second, education can improve the individual's knowledge about health (Kenkel, 1991) and the need for LTC in the first place, making it less likely that new information will cause the individual to change his LTC decision.

Age is also expected to influence the decision to allow a LTC policy to lapse. First, because most policies are rated according to age at the time of purchase, an older policyholder who purchased the policy at an early age is paying less than the market price. Second, as a policyholder ages, there are fewer payments remaining. For these reasons, policies become relatively more valuable at older ages, and a policyholder may be less likely to allow the policy to lapse as he or she ages.

Demographic variables that may have changed include marital status, number of children, geographic variables, health status, and financial circumstances. Marital status and the number of children are expected to influence the LTC insurance decision in two ways. First, married individuals and those with children might expect informal care from their spouse or children to serve as a substitute for formal paid care. If this is the case, the presence of children or a spouse might diminish the need for LTC insurance (Pauly, 1990; Zweifel and Strüwe, 1998), making it more likely that the policy will be allowed to lapse. Alternatively, the desire to provide for a spouse or children is a possible incentive to retain a policy, since assets that would be used to pay for LTC in absence of insurance could be preserved for family members if the LTC policy is renewed. In our model, we include marital status, the number of children, and a dummy variable that indicates whether or not the individual has a legal will.

Geographic location of the respondent is measured by a set of regional dummy variables reflective of the nine Census regions in the US. In addition, a dummy variable reflecting an urban or suburban setting is included to approximate the likelihood of living near a nursing home or intermediate care facility. It is expected that proximity to a facility may decrease the probability of allowing a LTC policy to lapse. In our sample, 71% of respondents live in an urban or suburban setting.

It is expected that people who require help with Activities of Daily Living (ADL's) such as walking and rising from a chair will be less likely to allow a policy to lapse, since they may have a more immediate need for long-term care. Therefore, we include the variable "ADL

Count" which indicates how many ADL's each person requires help with. "ADL Count" is expected to enter negatively, since those with a high ADL need are not expected to allow the policy to lapse.

Financial circumstances are expected to influence the probability of renewing LTC insurance in several ways. With regard to assets, we include two dummy variables that indicate if assets fall in the “high” range (over $1.5 million) or “low” range (less than $200,000). The reference category is mid-range assets ($200,000 to $1.5 million), and is based on a recommendation by Consumer Reports (2003) that those with assets in this mid-range are most likely to benefit from LTC insurance. For consumers with assets in the “low” range, Medicaid would likely cover LTC costs if care were needed. For those with assets in the “high” range, self-insurance (relying on personal funds) is a viable option if LTC is needed. Therefore, we expect LTC insurance lapse to be positively correlated with high and low asset levels.

In addition to asset level indicators, we include real estate equity, including the main home. Because LTC insurance is perceived by many as a way to protect one’s home, equity may be negatively correlated with LTC insurance lapse. Alternatively, individuals can draw on the equity in their homes to pay for LTC, so equity may serve as a substitute for LTC insurance and be positively correlated with a lapse.

Income is expected to influence LTC insurance renewal because individuals with higher incomes are better able to continue to afford the LTC insurance premiums. In the model presented here, we include household income divided by 2 if married and divided by 1 if single. We expect that income will be negatively correlated with a lapse.

Finally, we include a number of dummy variables that reflect changes occurring over the last two years, since the policy was last reported “in force.” These variables could influence the individual's financial circumstances, perception of risk, perception of the need for long-term care, and family circumstances. Any of these changes could result in a re-evaluation of the long-term care policy, and could possibly cause the policy to be dropped. The self-reported health improvement variable is of particular interest. A perception of improving health can reflect a lower probability of needing care, thereby increasing the probability of allowing a policy to lapse. Alternatively, it can reflect a higher life expectancy and therefore a higher probability of living long enough to require care, and a lower likelihood of allowing the policy to lapse.

Finally, we include two additional items of interest. We attempt to capture awareness of current information by including a dummy variable for whether the person uses email. We also include a dummy variable indicating whether the individual is satisfied with his or her overall health insurance package. If a policy is dropped, this satisfaction variable may indicate whether it was due to dissatisfaction with administration of the policy as opposed to inability to afford the premium or perception that long-term care insurance will not be needed.

RESULTS

Regression results are shown in Table 2. In this model, which includes only potential renewals (individuals who already had LTC coverage in the prior period), the coefficient on

price is positive and statistically significant. This suggests that price is a critical determining factor in the decision to renew, consistent with anecdotal evidence that individuals sometimes allow their policies to lapse because the policies become unaffordable. Marginal effects of price are small but significant, with a $1000 increase in price corresponding to an increase of approximately 4 percentage points in the probability of lapse (Table 3).

Table 2. Logit Model of Decision to Allow LTC Insurance Policy to Lapse in 2004

	G E Capital Pricing		AHIP Pricing		TIAA Pricing	
	β	S. E.	β	S. E.	β	S. E.
Price	0.0004**	0.0001	0.0004**	0.0001	0.0004**	0.0001
Years Contract Was In Force	0.0902**	0.0206	0.0991**	0.0229	0.0909**	0.0205
Age	-0.0887**	0.0203	-0.0973**	0.0234	-0.0902**	0.0202
Male	0.3869*	0.1721	0.3935*	0.1716	0.3869*	0.1724
White	0.2107	0.6694	0.2086	0.6716	0.2412	0.6704
Black	1.3899	0.7642	1.4104	0.7647	1.4355*	0.7648
Education	-0.0478	0.0374	-0.0468	0.0371	-0.0458	0.0375
Married	-0.3193	0.2109	-0.3319	0.2094	-0.3330	0.2111
Children	0.0218	0.0513	0.0219	0.0511	0.0226	0.0517
Urban	-0.0403	0.2220	-0.0415	0.2215	-0.0432	0.2221
Mid Atlantic	-0.7236	0.4846	-0.7215	0.4854	-0.7380	0.4846
EN Central	-0.9631*	0.4864	-0.9581*	0.4881	-0.9731*	0.4867
WN Central	-2.0051**	0.5490	-2.0011**	0.5490	-2.0105**	0.5484
South Atlantic	-0.9176	0.4752	-0.9202	0.4766	-0.9326*	0.4752
ES Central	-0.3364	0.5347	-0.3483	0.5334	-0.3419	0.5342
WS Central	-0.7330	0.5845	-0.7408	0.5865	-0.7475	0.5865
Mountain	-1.0721	0.5905	-1.0437	0.5896	-1.0842	0.5892
Pacific	-1.1563*	0.5173	-1.1566*	0.5183	-1.1692*	0.5174
High Asset Level	0.9448*	0.4394	0.9540*	0.4407	0.9556*	0.4381
Low Asset Level	0.8174**	0.2022	0.8059**	0.2021	0.8159**	0.2022
Real Estate Equity ($000's)	-0.0001	0.0002	-0.0001	0.0002	-0.0001	0.0002
Income ($000's)	-0.0003	0.0006	-0.0003	0.0006	-0.0003	0.0006
ADL Count	0.1234**	0.0363	0.1242**	0.0362	0.1233**	0.0364
Health Status Improved	0.2628	0.1953	0.2774	0.1945	0.2625	0.1952
Changed Residence	0.2751	0.2410	0.2710	0.2426	0.2673	0.2415
Spouse Recently Died	-0.6467	0.6106	-0.6819	0.6115	-0.6627	0.6143
Parent Recently Died	-0.4314	0.4915	-0.4138	0.4925	-0.4329	0.4935

Table 2. (Continued)

	G E Capital Pricing		AHIP Pricing		TIAA Pricing	
	β	S. E.	β	S. E.	β	S. E.
Recently Married or Partnered	0.6735	0.7772	0.6973	0.7730	0.6632	0.7842
Has Will	-0.9617**	0.2160	-0.9712**	0.2173	-0.9698**	0.2167
Uses Email	-0.3797	0.2245	-0.3788	0.2234	-0.3803	0.2250
Satisfied With Health Insurance	-0.6545	0.4337	-0.6854	0.4301	-0.6547	0.4350
Constant	5.0474**	1.6253	5.4606**	1.7220	5.1508**	1.6167
Number of Observations	1375		1375		1375	
McFadden's adjusted R^2	0.126		0.124		0.126	
Log Likelihood	-472.73		-473.60		-472.38	
LR (31)	209.16		207.41		209.86	

* Statistically significant at the α = 0.05 level.
**Statistically significant at the α = 0.01 level.

Table 3. Marginal Effects of Price Variables

		Marginal Effect
Price Proxy ($000's)	GE Capital Price	0.0372**
	AHIP Price	0.0400**
	TIAA Price	0.0344**

**Statistically significant at the α = 0.01 level.

Having an older policy is positively and significantly correlated with LTC insurance lapse. This is expected, because current insurance policies are typically more comprehensive than older policies, and information about LTC financing is more prevalent today than it was 25 years ago. In addition, public coverage is more comprehensive than it was prior to 1989 (Norton and Kumar, 2000). Therefore, those with older policies may have based their original purchase decisions on the paucity of substitutes available at the time, and may be less in need of a private policy today.

While the coefficient on race suggests that black respondents are more likely to allow a policy to lapse, the finding was not significant, with only a 94% confidence level. It should be noted, however, that non-white individuals represented only 4% of the sample.

Age is negatively correlated with LTC insurance lapse, as expected. This may reflect the increased probability of needing care as one ages, and also the increasing value of a renewed policy that was priced at a younger age, resulting in a tendency to continue renewing the policy as one ages. Other demographic factors, including education, marriage, and children are not significant. It is possible that these factors influence the initial decision to purchase LTC insurance, but are not factors in the decision to renew or drop a policy once it is in force.

Whether an individual lived in an urban setting does not significantly influence the decision to keep a policy. At the same time, the results show that lapses in coverage are more common in some areas than others, the most notable example being the "West North Central" region (MN, IA, MO, ND, SD, NE, KS), which is negatively correlated with policy lapses. It is interesting to note that in 2002, only five states in the U.S. had a LTC market penetration of over 15% (AHIP, 2004). Four of these states are in the “West North Central” census region.

As expected, an individual’s current financial circumstances are correlated with the probability of lapse. Those with household assets in the lowest range of “less than $200,000” are more likely to allow a policy to lapse. It is possible that individuals with assets below this level expect to eventually qualify for Medicaid, and therefore do not perceive a need to renew a private LTC policy. Those with household assets above $1.5 million are also more likely to allow a policy to lapse, possibly choosing to rely on personal funds if LTC becomes necessary. This finding is consistent with recommendations by financial advisors that those with assets in the middle range are most likely to benefit from LTC insurance. Real estate equity and income are not significantly correlated with a lapse.

Those who have a will are significantly less likely to allow a policy to lapse. This is as expected, because one motive for purchasing a LTC insurance policy is to preserve assets for children or a spouse.

Surprisingly, individuals who have difficulty with ADL's are more likely to allow a LTC insurance policy to lapse. LTC insurance is marketed as way to provide flexibility if care is needed, so one might expect those who are in a position to use LTC insurance benefits to continue to renew their policies. The finding here suggests the opposite.

Why might this occur? There are several possibilities. First, some individuals may have found that in the absence of a policy, they received some coverage of LTC services under Medicare or Medicaid, causing them to re-evaluate the necessity of continuing with a private policy. In early years of the HRS, respondents were asked if they had ever allowed a LTC insurance policy to lapse, and if so, why (HRS 1996-2002). Our finding here would be consistent with the second most frequently cited reason for allowing a policy to lapse: the private insurance was not needed.

Second, it is quite possible that those who have tried to obtain benefits under an existing policy had trouble receiving benefits they were entitled to (Duhigg, 2007). This would be consistent with another frequently cited reason for a policy lapse in the HRS survey: “general dissatisfaction” with the policy.

Finally, events that occurred over the past year, including deaths of family members and changes in residence and marital status do not significantly affect the decision to allow the policy to lapse.

CONCLUSION

In this empirical analysis of the decision to allow a LTC insurance policy to lapse, we find that financial considerations have a significant influence. Price was a significant factor in deciding whether to renew, even though most policies become a better deal as the person ages. Individuals with less than $200,000 in household assets were more likely to allow a policy to lapse, possibly reflecting the potential to qualify for Medicaid. Those with assets of

over $1.5 million were also likely to allow the policy to lapse, possibly relying on personal funds for LTC if care becomes necessary. Finally, those who had a will were less likely to drop the policy, probably with the intent of preserving assets for their heirs.

In this study, those who had trouble with ADL's were more likely to allow a policy to lapse. This may suggest a perception that the policy is not needed after all, or it may reflect dissatisfaction with administration of the policy. Older individuals were less likely to drop a policy.

Our findings suggest several approaches for the design and regulation of long-term care policies. First, the statistical significance of the price supports the claim that when an individual does allow an existing policy to lapse, it may be because of affordability. Second, the lower probability of renewal for people in the low-asset range suggests that Medicaid may, in fact, be crowding out some private purchase. These findings suggest that a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC. The importance of careful assessment has been a focus for senior advisors, since premiums paid into a policy are lost if the policy eventually lapses before care is needed.

There has been much discussion about people's understanding of LTC insurance and, in particular, concern that they believe they have LTC insurance coverage when in fact they have only a standard Medicare or basic health policy (AARP, 2001). Through the follow-up question "Is this a policy you've already described?" we are able to validate responses, identifying and recoding those who describe Medicare, supplemental, or basic health insurance as a LTC insurance policy. While this verification has greatly improved reporting accuracy, there is still a possibility that some people may report LTC insurance when in fact they have Medicare. Since people do not drop Medicare, this would tend to understate the lapse rate.

ACKNOWLEDGEMENTS

This research was supported by a Ruth L. Kirschstein National Research Service Award (Grant No. 1-F32AG023438-01) from the National Institute of Health/National Institute on Aging. The conclusions presented here are the responsibility of the authors and do not necessarily reflect the views nor the endorsement of the research sponsor.

REFERENCES

America's Health Insurance Plans (AHIP). 2004. *Long-Term Care Insurance in 2002* [accessed on February 11, 2005]. Available at: http://www.ahipresearch.org/pdfs/18_LTC2002.pdf

American Association of Retired Persons (AARP). 2001. "The Costs of Long-Term Care: Public Perceptions versus Reality." [accessed on April 4, 2007.] Available at http://assets.aarp.org/rgcenter/health/ltc_costs.pdf.

Belden, Russonello and Stewart. 2001. "In the Middle: A Report on Multicultural Boomers Coping With Family and Aging Issues." National Survey conducted for AARP by

Research/Strategy/Management [accessed on August 26, 2005]. Available at: http://assets.aarp.org/rgcenter/il/in_the_middle.pdf.

Brown, J. and A.Goolsbee. 2002. "Does the Internet Make Markets More Competitive?" *Journal of Political Economy* 110: 481-507.

Consumer Reports. 2003. "Long-term Care Insurance: Should You Buy?" accessed on February 11, 2005]. Available at: http://www.consumerreports.org/main/content/display_report.jsp?FOLDER%3C%3Efolder_id=344787&ASSORTMENT%3C%3East_id=333147&bmUID=1108757501241.

Duhigg, Charles (2007, March 26). "Aged, Frail and Denied Care by Their Insurers." *New York Times*, p. A1.

Finkelstein, Amy, Kathleen McGarry, and Amir Sufi. 2005. "Dynamic Inefficiencies in Insurance Markets: Evidence From Long-Term Care Insurance." *National Bureau of Economic Research, Inc.* NBER Working Paper 11039. [accessed on August 30, 2005]. Available at: http://www.nber.org/papers/w11039.

Feldman, Amy, Leslie Haggen, Laura Lallos, Teresa Tritch, Walter Updegrave, and Penelope Wang. 2000. *CNNmoney* 29 (7) [accessed on August 30, 2005]. Available at: http://premium.money.cnn.com/pr/subs/magazine_archive/2000/07/QUEB.html.

Gendel, Murray. 1998. "Trends in Retirement Age in Four Countries, 1965 – 95." *Monthly Labor Review* 121(8): 20 – 30. [accessed on April 4, 2007]. Available at: http://www.bls.gov/opub/mlr/1998/08/art2full.pdf.

Gupta, Aparna and Lepeng Li. 2004. "A Modeling Framework for Optimal Long-Term Care Insurance Purchase Decisions in Retirement Planning." *Health Care Management Science* 7: 105-17.

Health and Retirement Study. 1992-2004. Public use dataset produced and distributed by the University of Michigan with funding from the National Institute on Aging. Ann Arbor, MI [accessed on August 30, 2005]. Available at: http://hrsonline.isr.umich.edu.

Kenkel, Donald S. 1991. "Health Behavior, Health Knowledge, and Schooling." *Journal of Political Economy* 99: 287 – 305.

Leibowitz, A. and M. Chernew. 1992. "The Firm's Demand for Health Insurance." *Health Benefits and the Workforce*. U.S. Department of Labor. Washington, D.C.: U.S. Government Printing Office.

McCall, Nelda, Steven Mangle, Ellen Bauer, and James Knickman. 1998. "Factors Important in the Purchase of Partnership Long-Term Care Insurance." *Health Services Research* 33: 187-204.

McKenna, C. J. 1986. *The Economics of Uncertainty*. Brighton, Sussex: Wheatsheaf.

McNamara, Paul E. and Nayoung Lee. 2004. "Long-term Care Insurance Policy Dropping in the U.S. from 1996 to 2000: Evidence and Implications for Long-term Care Financing." *The Geneva Papers on Risk and Insurance* 29 (4): 640-651.

Morisato, Susan. 2004. "Growing Pains: Industry at the Crossroads" panel discussion. 17th Private Long Term Care Insurance Conference, June 23-25, 2004.

National Center for Health Statistics. 2006. *Health, United States, 2006 With Chartbook on Trends in the Health of Americans.* Hyattsville, MD: 176, 378. [accessed on 4/3/2007.] Available at http://www.cdc.gov/nchs/data/hus/hus06.pdf#027.

Norton, Edward C. and Virendar Kumar. 2000. "The Long-Run Effect of the Medicare Catastrophic Coverage Act." *Inquiry – Blue Cross and Blue Shield Association* 37 (2): 173-87.

O'Brien, Dennis. 2004. "Pricing Issues in Individual LTC Insurance" panel discussion. 17th Private Long Term Care Insurance Conference, June 23-25, 2004.

Pandya, Sheel. 2005. "Racial and Ethnic Differences Among Older Adults in Long-Term Care Service Use." *AARP Public Policy Institute Research Report.* [accessed on August 26, 2005]. Available at: http://www.aarp.org/research/longtermcare/trends/fs119_ltc.html.

Pauly, Mark V. 1990. "The Rational Nonpurchase of Long-Term-Care Insurance." *Journal of Political Economy* 98: 153-168.

Pauly, Mark V., Kate Withers, H. Krupa Subramanian-Viswanathan, Jean Lemaire, John C. Hershey, Katrina Armstrong, and David A. Asch. 2003. "Price Elasticity of Demand for Term Life Insurance and Adverse Selection." *National Bureau of Economic Research, Inc.* NBER Working Paper 9925. [accessed on February 11, 2005]. Available at: http://www.nber.org/papers/w9925.

Sahadi, Jeanne. 2006. "Retirement: You'll Need Far More Than You Think." *CNNMoney.com.* [accessed April 4, 2007.] Available at http://money.cnn.com/2006/09/21/pf/retirement/new_calcs_savings/index.htm?postversion=2006092214.

Sloan, Frank A. and Edward C. Norton. 1997. "Adverse Selection, Bequests, Crowding Out, and Private Demand for Insurance: Evidence from the Long-term Care Insurance Market," *Journal of Risk and Uncertainty* 15 (3): 201-19.

State of Michigan, Office of Financial and Insurance Services. 1996 – 1998. Rate filings. Lansing, MI.

VanDerhei, Jack. 2006. "Measuring Retirement Income Adequacy: Calculating Realistic Income Replacement Rates." *Employee Benefit Research Institute Issue Brief* 297. [accessed April 4, 2007.] Available at http://www.ebri.org/pdf/briefspdf/EBRI_IB_09-20061.pdf.

Wallace, S.P., L. Levy-Storms, R.S. Kingston, and R.A. Anderson. 1998. "The Persistence of Race and Ethnicity in the Use of Long-Term Care. *Journal of Gerontology B* 53: S104-12.

Zweifel, Peter and Wolfram Strüwe. 1996. "Long-Term Care Insurance and Bequests as Instruments for Shaping Intergenerational Relationships." *Journal of Risk and Uncertainty* 12 (1): 65-76.

Zweifel, Peter and Wolfram Strüwe. 1998. "Long-Term Care Insurance in a Two-Generation Mode." *Journal of Risk and Insurance* 65 (1): 33-56.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 195-206

Chapter 9

LEARNING, ACCEPTANCE AND THE PERCEPTIONS OF RISKS AND BENEFITS OF NEW TECHNOLOGY DEVELOPMENTS IN SPAIN

***Montse Costa*[a] *and Joan Costa-Font*[a,b]**
[a]London School of Economics, UK.
[b]Departament de Teoria Economica,
Universitat de Barcelona, Spain

ABSTRACT

The acceptance of risks associated with new technologies is a key issue that is likely to limit the extent of innovation in a 'risk society'. However, given the limited comprehensible information available to the public of new technologies, it is likely that risk information provision will have a heterogeneous effect on public perceptions. In order to examine this issue, we empirically examine the determinants of risk perceptions, benefit perceptions and risks acceptance of new technology developments in Spain. Our findings indicate that risk and benefits perceptions are not independent but affected by common information sources. Furthermore, by taking into account this effect individual's knowledge of science heterogeneously increases both risks and benefits perceptions.

Keywords: *risks perceptions, benefits perceptions, bivariate probit, Spain.*

1. INTRODUCTION

The diffusion of the potential risks and benefits of new technology developments products (e,g., products resulting from genetic manipulation of crops and animals), stands in forefront of the current policy disputes in Europe. Important skepticism in the acceptance of

* Contact address: Dr Joan Costa Font, LSE Health and Social Care, Cowdray House, London School of Economics, Houghton Street, WC2A 2AE , London. E-mail: j.costa-font@lse.ac.uk.

some new technology applications has centered the debate on science communication in the European Union that tackles an increasing media audience. However still small is known on *how individuals learn on new science developments* as well as on the understanding of how individuals perceive (or mentally weight) information signals as being beneficial and /or prejudicial for their health, the environment or society as a whole. The demand for new methods of risk communication challenges how innovation should be disseminated and applied with a 'knowledge economy'. Indeed, new products penetrate the supply chain by improving some qualities of existing ones however pose some concerns regarding potential side effects to human health and the environment although often unproven. In addition to those risks, new technologies can potentially changes people's lifestyles (e.g., mobile phones, etc). Whilst some people might welcome such changes and envisage them as benefits, others might perceive them as harmful and risky.

Individuals both as societal stakeholders as well as consumers often lack sufficient information to make decisions on new technologies and/or information available is often provided by untrustworthy channels e.g., though corporate marketing campaigns. Often not even publicly provided information is trusted as a result of recent food scandals in Europe. According to Sheehy et al (1998) even highly educated individuals exhibit small knowledge in some areas of science and technology. Partial (dis) information on the benefits and risks of new products often casts an emotional response that leads individuals to exercise their 'exit voice' by the (dis) approving the commercialisation of new technologies. However, an appropriate understanding of how attitudes are form in new technology developments is important both for private and public bodies to design communication strategies.

Some studies have shown that although consumer has small information they view genetic engineering as a risky process (Wohl, 1998). People perceive environmental as well as safety risks. Furthermore, often risks rather than being objectively measurable are qualitative categorization resulting from lack of sufficient information and knowledge, and thus stand as unknown risks. Therefore, the decision on whether to consume certain goods produced though new technology developments results from benefits out weighting the risks.

In a world of perfect information, real and perceived risks arguably would perfectly comparable. However, even when information is available, normally individuals information update comes at some cost and, still education places and important barrier to entry for some individuals to understand how certain technologies improves individuals well being. Therefore, human beings are not always open to update their knowledge and even when they do so they might be selective in which information they collect and thus information will rely on the importance they confer to certain information channels. Therefore, role of information channels and especially the media place important effects in explaining the paths that guide individual's attitudes and perceptions towards new science applications. Given individuals lack of information (which exhibit known risks) consumption decisions are likely to be influenced by external aspects of individual perception. Indeed, it has been widely demonstrated that individuals exhibit some aversion to "the unknown" (Ellsberg, 1961). If this stands in the biotechnology area, the we might expect that people would value higher prospects involving known risks even when potential benefits are larger in alternative protects involving unknown risks. Therefore, knowledge is expected to play a key role in determining the extent to which individuals perceive the risks and benefits of certain technologies, and consequently on how they react to their exposure.

Public resistance to new technologies might harm growth and trade and thus, society will be excluded of potential benefits. It has been suggested that in order to increase acceptance by improving the benefits and reducing potential risks (Franks, 1999). Industry and the government need to have a better understanding of consumer's acceptance of biotechnology as far as " the full benefits of biotechnology will only realize if consumers and the food industry accept the use of these new technologies as safe and beneficial"(Hoban 1988). However, risks and benefits, which are supposed to guide attitudes for new biotechnology, cannot often be regarded as independent. It might be the case that certain risks are perceived as more intense due to the lack of benefits, or the other way around, lack of benefits results from small benefits and high risks. Kanheman and Tversky (1986) associate perception of risks and benefits by suggesting that people are risks averse when the outcome of a decision is perceived as a benefit rather than reduction of a loss. Therefore, a second important research question refers to testing whether risks and benefits perceptions of new technologies are impendent, and whether the determinants of risks and benefits differ when risks and benefits are examined simultaneously. An interesting country, which has been relatively less examined, is that of Spain. Spain is a southern European country that has progressively adapted to European core values although still high religiosity is highly influential (Inglehardt and Baker, 2000). However, the extent of modernization and dissemination of new technologies has taken place must faster than in center European countries (Atienza and Lujan, 1997).

People perceptions related with technological developments are likely to depend on the type and level of information individuals handle, which determine the extent which individuals are susceptible to alarms and exert some outrage reaction (Standman, 1992)". Prior research indicates that people tend to overestimate unfamiliar, less known, unfair, involuntary and artificial risks, as well as the one they have no control, the ones extremely publicized and the ones where enhancing moral underlying concerns.

This paper aims at providing some understanding of the risks and benefits perception, and therefore the acceptance of new technologies. We undertake and empirical analysis of survey data on the Spanish perceptions to new technology developments. First, we argue that due to the significant lack of information associated with new technology developments lead to risks and benefits perceptions to be influenced by similar underlying effects. Second, knowledge exerts important role as increasing the benefits perceptions and reducing the risks of technology developments. Finally, we examine the determinants of risks and benefits perceptions, and most notably the role of information acquisition and socioeconomic determinants.

The structure of the paper is the following. Next we provide a theoretical discussion on the independence of risks and benefits perceptions and issues concerning the empirical specification of the model on the determinants of risk and benefits perceptions. Next section describes the data and the preliminary evidence. Section four deal with results and section five concludes.

2. TECHNOLOGY ACCEPTANCE AND THE INDEPENDENCE OF RISK AND BENEFITS PERCEPTIONS

2.1. The Theoretical Background

Yet, under imperfect information, individuals are likely to gather messages from several sources ($\eta_1,....\eta_n$) that convey the following information of risks in terms of a probability of a hazard $(\pi_1,.....,\pi_2)$, being the potential loss $(L_1,.....,L_n)$ and the potential utility improvement $(B_1,....., B_n)$. Individuals in deciding on new technologies consumption, they are assume to take into account all this information which is summarized in the net perceived utility of each product *(P_i)* subject to the existing public and private information *(I_i)*[1] . Yet, in the event of potential harmful effects in the future, consumer's current utility can be summarized in two main arguments: *H* (that refers to health), and *X* (consumption of other goods) such as $U_i = U_i(X,H)$, such that $\frac{\partial U}{\partial X} \geq 0$ and $\frac{\partial U}{\partial H} \geq 0$. On the other hand, future (expected) future utility can be conceptualised as $EU_i = \pi_i U(X,H) + (1-\pi_i)U(X,H-L_i)$, whereby consumption results if the net expected tulity is postive, or in other words if the potentials benenfits - both for health and consumption - assoicated with the consumption of a certain product overome the potential risks.

However, as far as consumers handle a limited amount of information on the possible consequences of certain events, potential information provision might exert an impact on individual's evaluation of the net expected utility of consumption. This effect takes place through a process of weighting new information, either public of private by sources γ_j determining the perception of risks ($\pi_i = \gamma_0 I_0 + ... + \gamma_n I_n$) and perceptions of benefits ($B_i = \gamma_0 I_0 + ... + \gamma_n I_n$) . This process of information updating resembles the one of the Bayesian learning models (Viscusi, 1992), because information sources might affect both decisions at a time, one might argue that perceptions of risks and benefits are liable to be associated, and thus are nor independent as one might theoretically conceive.

2.2. The Emprirical Specification

The decision to consume a certain new product following section 2.1 results from individuals' risks acceptance, or in other words, from perception of benefits overcoming risks. Let us specific the perception of risks as follows:

$$\begin{aligned} RISK^*_i &= X_i\beta_i + u_i \\ RISK_i &= 1 \; if \; RISK^*_i > 0 \\ RISK_i &= 0 \; otherwise \end{aligned} \qquad (1)$$

[1] Availability of information available can be conceptualized as "signals" that might lead to some individuals to subjectively become aware of some possible "individual or societal benefits/looses".

where RISK* is not observable but only where the individual responds to a survey question on whether he perceives risks on science innovation. Similarly, as benefit perceptions (BP*) are not observable but we only observe whether a dichotomous variable based on individuals responses as follows:

$$\begin{aligned} BP^*_i &= Z_i\delta_i + \varepsilon_i \\ BP_i &= 1 \; if \; BP_i^* > 0 \\ BP_i &= 0 \; otherwise \end{aligned} \tag{2}$$

As is common practice for a probit model we assume that the errors are distributed N(0,1) and the two models' errors are independent of one another, so that Cov(u_i, ε_i) = 0. However, it might well be that $u_i = \eta_i + \nu_i$ and $\varepsilon_i = \eta_i + \omega_i$, so that the errors in each model consist of a part ν_i, ω_i that is unique to that model, and a second part η_i that is common to both. If this is the case, the error terms are likely to be dependent. Thus, we're interested in the joint probability we use a bivariate normal distribution (u_i, ε_i) ~ BiN(0,0,1,1,),: whereby ρ is a correlation parameter denoting the extent to which the two error term covary. Finally, a decision that stands as key information for risk policymaking is that of risk acceptance. Again, because net utility is unobservable (NU*), we only observe whether individuals perceive that the benefits of science innovation overcome the risks as follows:

$$\begin{aligned} NU^*_i &= H_i\omega_i + \vartheta_i \\ NU_i &= 1 \; if \; NU_i^* > 0 \\ NU_i &= 0 \; otherwise \end{aligned} \tag{3}$$

3. THE DATA AND PRELIMINARY EVIDENCE

The data that we employ in this study is gathered from public sources, and in particular it refers to a survey commissioned by the Spanish Centre for Sociological Research in 1996 on 'Attitudes towards Scientific and Technology Innovation'. This is a representative survey of the Spanish population between 18 and 64 years of age. The sample initially was made of 2552 respondents; it was personal interviews to individuals from 91 municipalities and 43 provinces. The variable definition and descriptive statistics are explain are in Table 1. As expected the vast majority of the Spanish population beliefs on the 'expert opinions', only 12% is techno-sceptic and knowledge of science. About 45 trusts the state ad doing a good job, half of the same is made up of male and women, mean age is 41 years and 57% are married. Only 23% regard themselves as right wing, 2% practices religion and 37% is a family head.

Perceptions of risks (benefits) were measured from the response to the following question: "*Do you expect in the next 20 years the technology developments to bring many/some/few/ no risks (benefits) to society?*" . This question provides in formation on the

two hypothetical dimensions guiding individual acceptance of new science technologies. On the other hand, the survey includes another question, which goes like this: "Do you think that benefits of new technologies overcome the risks" (Yes/No).

Table 1. Variables definition

Variable	Definition		Mean	s.e
Experts	Do not Belief on experts	D	0.79	0.01
Tecno	Belief that technology will not improve way of life	O	0.12	0.01
Know	Knowledge level[2]	C	6.96	0.05
estate	Government Trust =1	D	0.45	0.01
gender	Female==1	D	0.50	0.01
Age	Age in Years	C	41.1	0.03
Married	Married ==1	D	0.57	0.01
Politic	Leftist=1	C	0.23	0.74
Practice	Follows a religion	D	0.02	0.01
headfam	Head of household	D	0.37	0.01

Table 2 measures the perception of risk and benefit of the technological and scientific advances. Interestingly, the vast majority of the population perceived both risks and benefits associated with technology innovation, although overall the share of those that perceive benefits overcomes that of risks. However, a sizeable share of the population perceives some risks which might arguably result from some resistance to new technologies, or an alternatively from the ignorance on the effects of new technologies, which are noted might lead to the so called 'risk ignorance or ambiguity aversion'. Furthermore, on the question of whether benefits of new technologies overcome the risks, 57% agreed this the assertion, indicating that although some individuals perceive technology related risks, the potential benefits seem to be larger.

Table 2. Perception of risks and benefits in the next 20 years (%), as a result of technological and scientific advances

Perception	Risks	Benefits
Many	16.69	21.69
Some	40.24	52.24
Few	28.21	13.06
None	4.11	2.98
n.s.	10.70	9.57
n.c.	0.04	0.47

In order to find some explanation to this evidence we examine further data. Interestingly, only the 57% of the population trusts on new technologies and 13% argue that new

[2] Derived from 12 answers to questions on science and technology (information provided upon request).

technologies cannot solve the problems of the oldest ones, 36% respond that life without technologies would be better, which could be regarded as 'techno skeptics'. 65% believe that the decisions related to technology cannot be exclusively based on consumer's knowledge but that need the intervention of experts. However, 92% agree that science and technology will on the whole improve the population quality of life. This, there is certainly some ambivalence in individuals responses to 'positive' and 'negative' dimensions of science developments. The clearest explanation relies in that science and technology might produce heterogeneous effects, both positive and negative, and that question on risks and benefits might stress specific effects. Accordingly, on the basis of our results, individuals in thinking on new technologies might suffer from a 'perceived aggregation effect'. Another explanation might lie in that individuals might be sensitive to the provision of information and might exhibit specific framing effects. However, given the general nature of the question and the fact that other survey questions referred to a vast array of technologies such as computers, cloning, space exploration, solar energy etc there is no reason to suggest that responses would be biased in some specific way.

4. RESULTS

This section reports the results of the different specification of the determinants of risks and benefits perceptions of new scientific innovation in Spain.

4.1. Risk and Benefit Perception

Let us begin with the determinants of risks perceptions by examining the significance of the estimated coefficients. As expected, individuals that are less likely to trust expects would perceive larger risks and the same applied for those individuals that exhibit a pessimistic belief on science and technology – those that perceive that technology will not improve way of life-. Thus, consistently with some previous literature risks perceptions convey relevant information on individuals trust although trust in the government did not place an effect on risks perceptions (Table3)[3]. Knowledge of science was significantly associated with lower risks perceptions, thus providing some support of 'ignorance aversion', whereby the lower the individuals knowledge the more likely she will be to perceive high science-related risks.

Interestingly, age does not exert a significant effect that indicates individuals of different ages, which are potentially exposed to different intensity of information acquisition, do not perceive risks systematically differently. On the other hand, Table 3 indicates that women are more likely to perceive risks than men, which is consistent with previous literature. Finally, whilst political orientation and religious are not significant predictors of technology risks perceptions, being married does exert a negative effect whilst family heads ado not lead to higher risks perceptions. The finding that married people would perceive fewer risks on technology developments might have to do with the fact that they are likely to have children that might benefit from future technology developments.

[3] We have tested the existence of multi-colinearity amongst regressions by performing cross-correlation, but all of them were lower than 0.30.

Table 3. Risk perceptions (ordered probit), Benenfit perceptions (ordered probit) and Risk Acceptance (probit model)

	Risk Perception		Benefit Perception		Risk Acceptance	
	Coeff.	s.e	Coeff.	s.e	Coeff.	s.e
NoExperts	0.262**	0.057	-0.189**	0.059	-0.284**	0.067
Tecno	-0.702**	0.072	0.547**	0.082	0.547**	0.083
Know	-0.019*	0.01	0.076*	0.012	0.077**	0.013
estate	-0.031	0.046	0.022	0.047	0.012*	0.055
gender	0.123**	0.052	-0.112	0.053	-0.132*	0.063
Age	0.003	0.022	0.069*	0.023	0.054*	0.026
Married	-0.117*	0.054	0.028	0.056	-0.105	0.066
Politic	0.001	0.001	0.051	0.001	-0.001	0.001
Practice	0.122	0.191	-0.267	0.224	-0.118*	0.060
headfam	-0.071	0.057	0.106**	0.057	0.026	0.069
Intercept					0.393	0.136
RV Chi (2,11)	140.59		164.62		137.41	
Likelihood Ratio Test	-2603.45		-2313		-1442	
Pseudo R^2	0.08		0.07		0.1	
% Corr	67%		71%		89%	

Overall, Table 3 reveals that benefits perceptions are influenced by similar variables as that of risks perceptions but they display opposite coefficients. This is the case of lack of trust in experts that is found to reduce the probability of perceiving benefits from science developments. As expected an optimistic belief on the effects of science in improving the quality of life s well as scientific knowledge increase the benefits perceptions. However, now although age does exert some positive effects, which indicate that possibly, older cohorts that have envisaged significant innovation changes tend to perceive higher benefits than younger cohorts. Gender exert just the opposite effect as in the risks perceptions case, female are less likely to perceive positive effects of new science developments. Interestingly, some political affiliation – those that regard themselves as left wing- is positively associated with larger benefits perceptions whilst those practicing a religious would perceive lesser benefits.

Figure 1 reports the predicted probabilities of each response conditioned on the knowledge effect. Interestingly, knowledge does exert some effect on reducing the probability of risk perception response and increasing the probability of a benefit perception response. Therefore, if knowledge captures the capacity of agents of updating information on new technologies, our results indicate that pro-active information policies would have a strong impact on risks acceptance of new technologies.

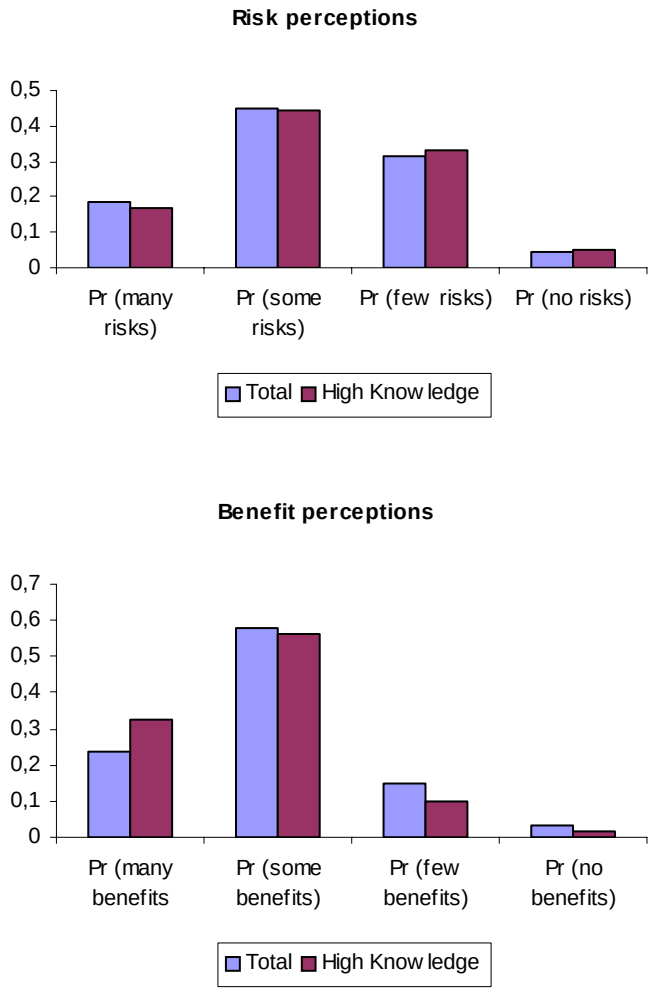


Figure 1. Risks and benefit perception of new technologies.

4.2. Risk Acceptance

According to theoretical model of previous sections, individuals in making decision on issues that convey some risks; they have to balance potential risks with accruing benefits. In Table 3, we examine using a probit model the determinants of individuals perceiving larger benefits that risks of science developments. Interestingly, trust in experts, beliefs on the potential improvements of quality of life, gender and age stand as the key risk acceptance drivers. Older individuals and especially men are more likely to accept technology related risks. Other relevant variables are that of religion, individuals practicing religion tend to be less likely to accept technology related risks. An interesting issue, in risks acceptance is the significance of the intercept term, which according to the prospect reference theory conveys information on prior assessment or risks and benefits, which displays a positive effect. Figure

2, indicates that predicted risks acceptance increases with individuals knowledge, indicating the sensitivity of individuals attitudinal reactions to knowledge of science.

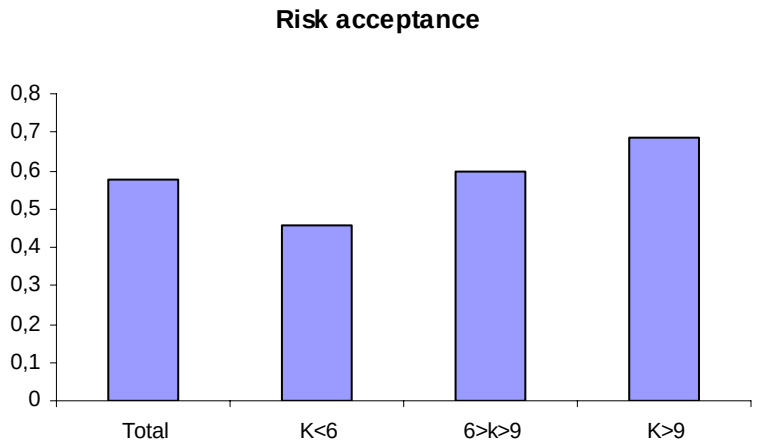


Figure 2. Risk Acceptance and knowledge.

Table 4. Joint estimation of risks and benenfit perceptions (bivariate probit)

	Risk perceptions			Benenfit perceptions		
	Coef.	Std. Err.	z	Coef.	Std. Err.	z
NoExperts	0.181	0.063	-2.878	-0.287	0.067	4.282
Tecno	-0.733	0.087	-8.416	0.539	0.080	6.717
Know	0.032	0.011	2.886	0.126	0.012	10.652
estate	0.091	0.051	1.786	0.030	0.056	0.535
gender	0.131	0.058	2.242	-0.071	0.065	1.087
Age	-0.025	0.024	-1.051	0.079	0.026	3.031
Married	0.133	0.060	2.204	-0.087	0.066	-1.317
Politic	0.000	0.001	-0.410	0.001	0.001	1.958
Practice	-0.238	0.205	-1.161	0.046	0.219	0.212
headfam	0.082	0.064	1.290	-0.018	0.071	-0.256
Intercept	0.066	0.122	0.540	-0.094	0.130	-0.722
ρ	-0.20	0.01				
FV	-2.999.77					
Wald (2,22)	345.31					
RV (ρ =0)	32.45					

4.3. The Independence of Risk and Benefits Perceptions

The fist results indicates that as expected both decision are jointly formed, as far as the correlation coefficient of the erros terns is significantly different from zero. Optimistic beliefs on technology would produce higher benenfit perceptions and lower isks perception, howewer the effect of knowledge of science is shcnages. Interestingly, a higher kwnoweldge leads to higher risks and benenfits perceptions. Furthemore, the gender effect is only prevalent on risks perceptions, again incresing risks perception but not on benenfit perceptions. The religion specific effect disper whilst the political affiliation effect remain for benenfit perceptions.Overall, the results suggest that the process of risk and benenfit formatio is not independent and that certain information channels incresing risks perception might display an effect on benenfit perceptions.

5. CONCLUSION

The acceptance of new technologies conveying benefits (and potential risks) to the population is determined by information sources, and primarily the individuals trust in experts as well as their knowledge to update prior information. However, in the light of our results, risks and benefits perceptions are not independent and that taking into account the potential dependence of similar information channels might affect the risk learning determinants. However, there were some information channels that that affect only benefits and risks separately. This was the case of age, religion practice and political affiliation which affected only benefits perceptions whilst being married that affected only risks perceptions while was not significantly associated with risks perceptions.

This study provides some issues for discussion in the light of risks communication of new technology developments. On the one hand, we have raised the point that benefits of new technologies although perceived are largely dependent on individuals knowledge. Second, we have shown that new information signals conveying risks information are likely to enhance lower benefits and the other way around and that when taking into account this feature, the effect of knowledge of science does determine the increase of both risk and benefits perceptions. However, if individuals are ambiguity or ignorance averse, they prefer known risks and thus they might be skeptical about the acceptance of new technologies until sufficient information is disseminated and their knowledge achieves the desired levels, which arguably are culturally determined.

REFERENCES

Atienza, J and Lujan, JL (1997). Opiniones y Actitudes, N 14. *La imagen social de las nuevas tecnologías en España*, CIS, Madrid.

Ellsberg, D. (1961), Risk, Ambiguity and the Savage Axioms, *Quarterly Journal of Economics* 75, 643-669.

Franks, J (1999). The status and prospects for genetically modified crops in Europe. *Food Policy*, 24: 565-584.

Inglehardt, R and Baker, W.E (2000). Modernization, Cultural Change and the Persistence of Traditional Values. *American Sociological Review*, February, 2000.

Kanheman, D and Tversky, A (1986). Prospect Theory and Evaluation. *Journal of Cognitive Psychology*, 73: 34-48.

Sheery, H Legault, Ireland, D (1998). Consumer biotechnology: a synopsis of survey and focus group research. *Journal of Consumer Policy*, 21: 359-386.

Stadman, PM (1992). Hazard vs Outrage. IOMA Broadcaster, 6.

Wohl, J.B (1998). Consumers' decision-making and risk perceptions regarding foods produced with biotechnology. *Journal of Consumer Policy*, 21: 387-404.

In: Psychology of Decision Making in Economics… ISBN: 978-1-60021-917-7
Editor: K. P. Hofmann, pp. 207-215

Chapter 10

THE CONFLICT BETWEEN MONEY AND SELF-ESTEEM IN INVESTMENT

Liqing Zhang*
Peking University

ABSTRACT

The present theory proposes that investors not only think of future monetary benefits, but also value the choices' implications regarding their self-esteem in decision making. Self-esteem is one's subjective evaluation of the self. Most people want to maintain a positive self-image. When they decide to invest in a project, people expect to receive financial rewards, and they also hope to enhance their self-esteem through the success of the project. Thus, when their initial investment produces negative economic return, they not only suffer financial loss, but also encounter challenges to their self-esteem. They can withdraw from the project to minimize their monetary loss, or they may keep throwing additional money into the project to demonstrate that their initial decision was correct. It is painful to admit a mistake because it poses negatively to the investors' self-concept. As a result, investors may be entrapped within a losing project and suffer accumulated financial loss. The present theory suggests that when investors encounter conflicts between money and self-esteem in decision making, they may choose to give up money in order to defend their self-esteem.

Keywords: *Money, Self-esteem, Investment, Motivation, Decision Making.*

To some extent, we are all investors. We invest money and time in education. We devote much effort into specialized training before we start a career. The return that we expect to get is not merely to receive some regular paycheck and make a living, but also to validate our

* Correspondence regarding this article should be addressed to Liqing Zhang, Department of Psychology, Peking University, 5 Yi-He-Yuan Avenue, Beijing, P. R. China, 100871. Email: liqingzhang@pku.edu.cn.

self-worth. Similarly, in financial investment, people not only expect to receive economic benefits, but also hope that their investments succeed and thus glorify their self-concepts.

SELF-ESTEEM

Self-esteem refers to the subjective evaluation of the self. It is a fundamental motivation of human beings. Research in psychology has provided abundant empirical evidence that people want to defend, maintain, and enhance a positive self-image (Baumeister, 1998; Crocker and Park, 2004; Greenwald, 1980; James 1890; Steele, 1988; Taylor and Brown, 1988). People attribute success to their internal traits, but blame external factors for their failure (Miller and Ross, 1975). They derogate outgroup members in order to enhance their own group identity (Crocker and Luhtanen, 1990). They may compare themselves with others inferior to them so that they can enhance their self-esteem (Wills, 1981). In many situations, self-esteem plays an adaptive role in human functioning. Self-esteem fosters confidence, optimism and controllability, and thus it helps people to survive hardships (Taylor and Brown, 1988). For instance, cancer patients with high self-esteem tend to live longer after they are diagnosed with the disease than those with low self-esteem (Taylor and Brown, 1988). However, the pursuit of self-esteem may have costs. Attributing failure to external factors defends people's self-esteem, but it also precludes them from learning from their experiences. The belief that one's contribution is valuable may validate their self-worth, but it produces conflicts in resource allocation. For instance, in distributing the reward of a collaborative work, people tend to come up with different fair allocations, and the solution each side proposes usually favors their own position (Messick and Sentis, 1979; Van Avermaet, 1974). The present paper argues that economic loss may be a cost in pursuing self-esteem.

CONFLICT BETWEEN MONEY AND SELF-ESTEEM IN INVESTMENT

The conflict between money and self-esteem becomes salient after people have chosen to invest in a project and the project produces negative economic return. When they decide to invest money into the project, they tie their pride to it, at least to a certain degree. When the project does not go on well, they can either withdraw or continue. Withdrawal from the project implies that the decision makers have made a bad decision. Therefore, instead of terminating a losing project, people may put additional money into the project to justify their original decision (Brockner and Rubin, 1985; Brockner, 1992; Staw, 1976, 1997). They may assume that the difficulty is temporary and there is still the prospect of making a profit. The motivation to defend their positive self-image thus produces costly over-investment in failing projects. In sum, people encounter conflict between money and self-esteem when they receive negative financial reports of their chosen project. They may defend their self-esteem at the expense of money.

Research has shown that ego-threat may increase the tendency to incur monetary loss to defend people's self-esteem. Ego-threat refers to the situation in which people perceive threats to their self-esteem. Ego-threat is usually produced by negative information about the

self, such as failure on an important task. People become more defensive after their self-esteem is challenged (Baumeister, 1998; Steele, 1988). They often make efforts to restore their positive self-image. If they encounter a decision dilemma of terminating or continuing a losing project, ego-threated people may be more likely to invest repeatedly so as to defend their self-esteem.

Recent empirical research supports the idea that people may sacrifice material payoffs in order to protect their self-esteem in entrapping situations. Ego-threat increases people's desire to restore self-esteem, and thus promotes their over-investment in a failing project. For instance, in Experiment 1 conducted by Zhang and Baumeister (2006), ego-threat was manipulated by suggesting that participants may have a negative personality—choking under pressure. In the ego-threat condition, participants were told, "If you're the kind of person who usually chokes under pressure or if you don't think that you have what it takes to win the money, then you might want to play it safe. But it's up to you." In the control condition, participants were not informed of any ego-threatening information. The entrapping situation was a gamble procedure that was adapted from Brockner and Rubin (1985). It represents chance-based situations, or other situations where people are not in control, such as the stock market, or waiting for a bus that does not come on time. In this game, participants were given $5 dollars as deposit. They were told that they had the opportunity to win a jackpot of $10. They were not informed of the exact probability of winning. This experiment was conducted using a computer program. A counter on the computer screen ran from 0 to 500. For every 25 numbers that the counter ran, participants invested in a quarter. When the counter reached a multiple of 25, such as 25, 50,---, the counter stopped and participants had to click on one of the buttons: continue or exit. Thus, each time participants decided to continue, they invested an additional quarter. They were told that they would win $10 if a bell rang when the counter stopped. This game finished when participants decided to withdraw or when they used up all of their deposit (the counter reached 500). In fact, the program was set up so that no beep would be sound. Thus, each time they invested in some money, participants received negative economic return. The results demonstrated that participants in the ego-threat condition (M = 3.67, SD = 1.59) invested a larger amount of money than those in the control condition (M = 2.43, SD = 1.94). Thus, people who were eagerly motivated to restore their self-esteem invested and lost more money in the failing project than other people.

The motivation to maintain their self-esteem also impacts people's investment when their success depends on their ability and effort. In Experiment 2 conducted by Zhang and Baumeister (2006), ego-threat was manipulated by informing participants that they may have the negative personality of choking under pressure. The entrapping situation was a skill-related task, specifically solving a jigsaw-puzzle (Brockner and Rubin, 1985). Thus, participants might have thought that their skills and efforts could improve their probability of success. The puzzle was solvable, though the time limit--15 minutes made the task difficult. In fact, no participant solved it. Participants were given $4 as deposit. The prize for solving the puzzle was $10. Participants were told that solving a complicated puzzle under time pressure reflected one's capacity to perform well in challenging situations. They were first given 19 free puzzle pieces. They then had to purchase additional puzzle pieces at the price of 5 cents a piece. They were not informed of the exact number of additional pieces required to solve the puzzle. If they solve the puzzle within 15 minutes, they would gain a $10 prize, and the money they used to purchase the puzzle pieces would be given back. There were no differences in the time that participants spent in solving the puzzle. Participants in the ego-

threat condition (M = 2.12, SD = .91) invested more money in buying the puzzle pieces than those in the control condition (M = 1.47, SD = .77).

In the jigsaw-puzzle situation, people face the conflict between money and self-esteem after they have purchased some puzzle pieces and the puzzle is still not solved. The money that they have spent to buy the puzzle pieces is forgone. They can stop their investment in this task so that they minimize their financial loss. The findings suggest that they may rather continue investing in order to demonstrate that they are able to deal with the challenging situation. Ego-threat spurs the desire for self-esteem and increases the entrapment in losing endeavors.

It is conceivable that ego-threat influences choices in the potential investing situations because withdrawal poses negatively to one's self-concept (Staw and Ross, 1987). In Experiments 1 and 2 in Zhang and Baumeister (2006), the manipulation of ego-threat was related to the subsequent investment. It linked the decision of playing it safe to a personality weakness—choking under pressure. Therefore, it raised the question of whether ego-threat still increased the over-commitment to a losing project when the ego-threat was produced by a source irrelevant to the potential entrapping situation.

Empirical research shows that self-esteem influences people's investment choices, even when the ego-threat manipulation does not have any relevant connection with the subsequent decision making. Experiment 4 in Zhang and Baumeister (2006) was conducted to address this question. In it, ego-threat was manipulated by failure performance feedback on a creativity test (Baumeister, Tice and Heatherton, 1993). All participants completed a creativity test. In the ego-threat condition, participants received failure performance feedback. In the control condition, participants were given success performance feedback. The entrapping situation was the counter game as discussed earlier in this paper. The results showed that participants in the ego-threat condition (M = 4.04, SD = 1.34) invested more money than those in the control condition (M = 2.40, SD = 1.94). Therefore, ego-threat increased entrapment in losing projects even when the source of the ego-threat was irrelevant to the entrapping situations.

The motivation to defend self-esteem also contributes to the costly entrapment in interpersonal competition. For instance, in Experiment 3 conducted by Zhang and Baumeister (2006), ego-threat was produced by failure performance on a creativity test. The procedure that represented interpersonal competition was called dollar auction (Brockner and Rubin, 1985; Shubik, 1971; Teger, 1980). In this game, participants competed with another person in buying a dollar. A special feature of the game was that the person who paid the second highest price also had to pay their bid, although they did not get the dollar. In this experiment, participants were run in groups of four people. They were told that they were randomly paired with another person to bid for one dollar. They were given $5 as deposit. They were told that they could choose whether to bid or not. If they chose to bid, they could decide how much to bid. The rule of the game was that the person who bided the highest paid their bid and attained the dollar. The person who had the second highest bid was also charged for their bid, even though this person did not gain anything in return. In the opening round of the auction, each person wrote down their own bid. Starting from the second round, participants were told how much their opponent had bided and they decided whether to bid more or not. The game ended when one person decided to withdraw from the auction. The highest bid that one could offer was their total deposit of five dollars. In fact, each time the participant offered a bid, the experimenter told them that their opponent had bided more. The results showed that

participants in the ego-threat condition (M = 3.71, SD = 1.58) bided more than those in the control condition (M = 2.46, SD = 1.89).

In the dollar auction, participants may first want to put in a little bit of money to gain one dollar. As the auction went on, their motivation changed from gaining a small amount of money to winning the game (Teger, 1980). When they bided more than one dollar, it was obvious that the motivation to win, instead of earning money, was the driving force of their decisions. By additionally analyzing the data of Experiment 3 in Zhang and Baumeister (2006), participants' highest bids were divided into two categories: below (or equal to) one dollar, and over one dollar. The results of the Chi Square analysis showed that participants in the ego-threat condition (30 out of 35) were more likely to bid more than one dollar than those in the control condition (21 out of 34), Chi^2 (1, N = 69) = 5.13, $p < .02$.

Thus, when people encounter conflicts between money and self-esteem in interpersonal competition, their reluctance to withdraw may produce financial loss for both parties. This theory provides a perspective in understanding a variety of destructive behaviors in interpersonal and intergroup conflicts, such as price wars, strikes, and costly law suits in divorce (Teger, 1980).

In the counter game and the jigsaw-puzzle game, it is conceivable that people want to put additional money into the losing project so that they can change the losing situation into a profitable one by winning the jackpot or solving the puzzle. However, in the dollar auction, it is apparent that no one is able to turn the losing situation into a winning one if people bid more than a dollar in exchange for one dollar. It is obvious that people increase their bid in the dollar auction because they do not want to be defeated by others.

Maximizing monetary payoffs is usually considered the bench mark for rational choices. The present theory indicates that people may consider factors other than money during decision making, and specifically, people may strive for self-esteem at the expense of economic payoffs. It should be noted that the pursuit of self-esteem at the expense of money may help people defend their pride in the short-run, but it may eventually produce much larger costs of both money and self-esteem (Baumeister, Heatherton, and Tice, 1993; Fox and Staw, 1979; Ross and Staw, 1986, 1993). For example, British Columbia's decision to host a world fair (Expo 86) in Vancouver ended up in a costly entrapping situation (Ross and Staw, 1986). The initial budget in 1978 was $78 million, but the cost eventually accumulated to $1.5 billion with a deficit of $311 million. The Provincial Premier William Bennett originally decided to hold the fair and tied his pride to it. He declined the fair director's suggestion to cancel it despite the negative financial reports about the project. Therefore, giving up monetary payoffs may not be a best strategy to defend one's self-esteem.

GENERAL DISCUSSION

Research has demonstrated that people throw good money away into losing projects (Bazerman, Giuliano, and Appelman, 1984; Brockner and Rubin, 1985; Brockner, 1992; Garland, 1990; Staw, 1976, 1997; Teger, 1980). Entrepreneurs over-invest in projects that have produced negative financial return (Antonides,1995; McCarthy, Schoorman and Cooper, 1993). People fail to change their unpromising career choices (Drummond and Chell, 2001).

The present paper argues that people may forgo financial benefits in chasing their positive self-image.

Implications to Other Theories on Entrapment

The present idea is consistent with the self-justification theory on entrapment (Brockner, 1992; Brockner, Houser, Birnbaum, etc, 1986; Brockner and Rubin, 1985; Staw, 1976, 1979, 1981, 1997). Self-justification theory deems that people put more money into a previously chosen course of action because they want to justify that their initial decision of investment is wise. The notion that self-esteem is the motivation underlying self-justification was once suggested by Staw (1981). Empirical research also suggests that people are more likely to be entrapped into a losing project when the project is diagnostic of an important ability than when the task is irrelevant to any central attribute (Brockner and Rubin, 1985). Decision makers may perceive that their self-esteem is more at stake when the success of the project signifies their ability than when the failure of the task does not pose any threat to their competence. Research also demonstrates that people put in a larger second investment when they themselves decided the initial investment than when someone else chose the project (Schoorman and Holahan, 1996; Staw, 1976). Apparently, people's self-esteem was more at stake when they themselves made the original choice. To some extent, the present theory advances the self-justification theory by demonstrating self-esteem as the underlying motivation of self-justification.

Furthermore, the present theory suggests that motivational theories are able to provide new predictions that are not easily explained by competing cognitive theories. Even though self-justification theory is an influential theory on entrapment, alternative cognitive theories have been proposed (for review, see Brockner, 1992; Wilson and Zhang, 1997). Specifically, prospect theory (Kahneman and Tversky, 1979; Whyte, 1986) and decision dilemma (Bowen, 1987) theory provide two competing alternatives. According to prospect theory, decision makers prefer to take risks when they encounter a choice between accepting a sure loss and gambling to turn the situation around. After they have invested a small amount of money and receive negative financial return, decision makers have to accept the sure loss of the previous investment if they withdraw. Due to loss aversion, they would rather put additional money into the project so that they can have the opportunity (though small) to turn the situation around. It has been difficult to differentiate the driving forces between prospect theory and self-justification theory because the predictions of both theories are usually similar. Some researchers even proposes that prospect theory is able to account for all the findings that support self-justification theory and can also explain a broader range of phenomena (e.g., Whyte, 1986). Even though prospect theory offers an important alternative explanation, in particular for situations that do not involve the motivation to maintain an integrative self-concept (e.g., Arkes, and Blumer, 1985), self-esteem contributes to costly entrapment. Prospect theory cannot explain the findings reviewed in this paper (Zhang and Baumeister, 2006). Risk seeking in loss domains cannot interpret why people become more locked into losing endeavors after their self-esteem is threatened.

Decision dilemma theory (Bowen, 1987) suggests that ambiguity may account for the over-investment in losing situations. Since the possibility of success is uncertain or unknown, people may think that they may eventually succeed by investing repeatedly.

Even though ambiguity may influence entrapment in some situations, the decision dilemma theory can not explain why challenges to self-esteem increase investors' risk taking. Additionally, in the dollar auction, there is clearly no possibility to turn bad into good after people pay more than one dollar for a dollar, but people, especially investors who have received ego-threat and thus have greater desire for self-esteem, bid much more than one dollar. Therefore, the present theory advances the motivational viewpoint of entrapment.

Indications for De-Escalation

The present theory suggests that the interpretation of the situation may influence investors' choices when their previously chosen project has produced economic loss. If they view withdrawal positively, investors may more likely stop investing in a failing project than if they regard withdrawal as a personal weakness (Brockner and Zubin, 1985; Simonson and Staw, 1992; Staw and Ross, 1987). If people do not view withdrawal as an indication of failure, they may be more likely to stop an unprofitable project. Re-interpretation of the situation reduces withdrawal's negative implication to the investors' self-concept and thus contriutes to de-escalation.

The present theory suggests that self-affirmation may de-escalate commitment to a losing project. Self-affirmation refers to the verification of one's positive self-image. Research has documented that self-affirmation reduces defensiveness (Steele, 1988). Future research should test empirically whether people become less entrapped in a failing project after their self-esteem has been assured.

CONCLUSION

The present theory proposes that people may encounter conflicts between money and self-esteem in investment. The conflict between money and self-esteem becomes salient when people have invested in a project and the project produces monetary loss. People thus face the tough choice: withdrawal or continuing. Withdrawal may minimize financial loss, but it may also pose negatively to investors' self-esteem. Continual gives investors a small possibility to turn the losing situation into a winning one, but it may also more likely produce a large amount of accumulated loss. Furthermore, continuing may indicate that investors' initial decisions were wise and thus protect their self-esteem. The present paper reviews evidence showing that people are entrapped more deeply in losing endeavors when their self-esteem is threatened and thus they are more motivated to restore their positive self-image. Therefore, people may sacrifice monetary payoffs in pursuing self-esteem.

REFERENCES

Antonides, G. (1995). Entrapment in risky investments. *The Journal of Socio-Economics*, 24, 447-461.

Arkes, H. R., and Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*, 124-140.

Baumeister, R. F., Dale, K., and Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality, 66(6),* 1081-1124.

Baumeister, R. F., Heatherton, T., and Tice, D. M. (1993). When ego threats lead to self-regulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology, 64*, 141-156.

Bazerman, M. H., Beekun, R. I., and Schoorman, F. D. (1982). Performance evaluation in a dynamic context: A laboratory study of the impact of a prior commitment to the ratee. *Journal of Applied Psychology, 67*, 873-876.

Bowen, M. G. (1987). The escalation phenomenon reconsidered: Decision dilemmas or decision errors? *Academy of Management Review, 12*, 52-66.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review, 17*, 39-61.

Brockner, J., Houser, R., Birnbaum, G., LIoyd, K., Deitcher, J., Nathanson, S., and Zubin, J. Z. (1986). Escalation of commitment to an ineffective course of action: The effect of feedback having negative implications for self-identity. *Administrative Science Quarterly, 31*, 109-126.

Brockner, J., and Rubin, J. Z. (1985). *Entrapment in Escalating Conflicts: A Social Psychological Analysis*. New York: Springer-Verlag.

Crocker, J., and Park, L.E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, 130(3), 392-414.

Crocker, J., and Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology*, 58, 60-67.

Drummond, H., and Chell, E. (2001). Life's chances and choices: A study of entrapment in career decisions with reference to Becker's side bets theory. *Personnel Review, 30*, 186-202.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*, 603-618.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47*, 263-291.

James, W. (1890). *The principle of psychology (Vol.1)*. Cambridge, MA: Harvard University Press.

McCarthy, A. M., Schoorman, F. D., and Cooper, A. C. (1993). Reinvestment decisions by entrepreneurs: Rational decision-making or escalation of commitment? *Journal of Business Venturing, 8*, 9-24.

Messick, D. M., Sentis, K.P. (1979). Fairness and preference. *Journal of Experimental Social Psychology, 15*, 418-434.

Miller, D. T., and Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82*, 213-225.

Ross, J., and Staw, B. M. (1986). Expo 86: An escalation prototype. *Administrative Science Quarterly, 31*, 274-297.

Schoorman, F. D., and Holahan, P. J. (1996). Psychological antecedents of escalation behavior: Effects of choice, responsibility, and decision consequences. *Journal of Applied Psychology, 81*, 786-794.

Shubik, M. (1971). The dollar auction game: A paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution, 15*, 109-111.

Simonson, I., and Staw, B. M. (1992). Deescalation strategies: A comparison of techniques for reducing commitment to losing course of action. *Journal of Applied Psychology, 77*, 419-426.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to chosen course of action. *Organizational Behavior and Human Performance, 16*, 27-44.

Staw, B. M. (1979). Rationality and justification in organizational life. In B. M. Staw and L. L. Cummings (Eds.) *Research in organizational behavior. (Vol. 2).* Greenwald Connecticut: JAI Press.

Staw, B. M. (1981). The escalation of commitment to a course of action. *Academy of Management Review, 6*, 577-587.

Staw, B. M. (1997). The escalation of commitment: An update and appraisal. In Z. Shapira (Ed.) *Organizational decision making* (pp191-215). New York, NY: Cambridge University Press.

Staw, B. M., and Ross, J. (1987). Behavior in escalation situations: antecedents, prototypes, and solutions. *Research in Organizational Behavior, 9*, 39-78.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261—302). New York: Academic Press.

Taylor, S. E., and Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Teger, A. I. (1980). *Too Much Invested to Quit*. Elmsford, New York: Pergamon Press.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization, 1*, 39-60.

Van Avermaet, E. (1974). Equity: A theoretical and experimental analysis. *Doctoral Dissertation*, University of California, Santa Barbara.

Whyte, G. (1986). Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review*, 11, 311-321.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.

Wilson, R. M. S., and Zhang, Q. (1997). Entrapment and escalating commitment in investment decision making: A review. *British Accounting Review, 29*, 277-305.

Zhang, L., and Baumeister, R. F. (2006). Your money or your self-esteem: Threatened egotism promotes costly entrapment in losing endeavors. *Personality and Social Psychology Bulletin, 32*, 881-893.

INDEX

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y