

JUDGMENT AND DECISION MAKING

Baruch Fischhoff



earthscan
from Routledge

JUDGMENT AND DECISION MAKING

Baruch Fischhoff is a long-time leader in behavioral decision research. This book presents his most important work in the field, offering the reader both an understanding of how decisions are made and lessons on how to improve decision making. Relevant across a range of professional sectors and academic disciplines, this book is essential reading for those who need a breadth and depth of understanding of this field and how it can help us in the important decisions we face.

This book includes sections on hindsight bias, judgment under uncertainty, value elicitation, and decision making. Applied topics include adolescent risk taking, sexual assault, insurance, environment, medicine, technology, national security, and natural disasters. Introductory sections highlight important underlying themes: that understanding decision making requires research that integrates formal analysis (of how decisions should be made), empirical studies (of how decisions are actually made), and interventions (on how decision making can be improved). If science can clarify people's decision-making needs and abilities, while society provides the information and opportunities for them to steer their own lives, Fischhoff's view offers a feeling of guarded optimism regarding the opportunities for helping people to make sound choices. Thus, the readings provide scientific input to the ongoing debate on the proper roles of laypeople and technical experts in charting society's course.

Baruch Fischhoff is Howard Heinz University Professor, Carnegie Mellon University, USA, Department of Social and Decision Sciences, Department of Engineering and Public Policy.

EARTHSCAN RISK IN SOCIETY SERIES

Series editor: Ragnar E. Løfstedt

- 1. Calculating Political Risk**
Catherine Althaus
- 2. The Citizens at Risk**
From Urban Sanitation to Sustainable Cities
Gordon McGranahan, Pedro Jacobi, Jacob Songsor, Charles Surjadi and Marianne Kjellen
- 3. The Earthscan Reader on Risk**
Edited by Ragnar E. Løfstedt and Åsa Boholm
- 4. The Ethics of Technological Risk**
Edited by Lotte Asveld and Sabine Roeser
- 5. Facility Siting**
Risk, Power and Identity in Land-Use Planning
Edited by Åsa Boholm and Ragnar E. Løfstedt
- 6. The Feeling of Risk**
New Perspectives on Risk Perception
Paul Slovic
- 7. Foresight in Action**
The Practice of Dealing with Uncertainty in Public Policy
Marjolein B. A. van Asselt, Susan A. van 't Klooster, Philip W. F. van Notten and Livia A. Smits
- 8. Global Environmental Risk**
Jeanne X. Kasperson and Roger E. Kasperson
- 9. Hazards, Vulnerability and Environmental Justice**
Susan L. Cutter
- 10. The Perception of Risk**
Paul Slovic
- 11. Public Safety and Risk Assessment**
Improving Decision Making
David J. Ball and Laurence Ball-King
- 12. Risk Governance**
Coping with Uncertainty in a Complex World
Ortwin Renn
- 13. Risk Management in Post-Trust Societies**
Ragnar E. Løfstedt
- 14. Risk, Media and Stigma**
Understanding Public Challenges to Modern Science and Technology
Edited by James Flynn, Paul Slovic and Howard Kunreuther
- 15. Risk, Uncertainty and Rational Action**
Carlo C. Jaeger, Ortwin Renn, Eugene A. Rosa and Thomas Webler
- 16. The Social Contours of Risk (Volumes 1 & 2)**
Jeanne X. Kasperson and Roger E. Kasperson
- 17. Social Trust and the Management of Risk**
Edited by George Cvetkovich and Ragnar E. Løfstedt
- 18. The Spatial Dimension of Risk**
How Geography Shapes the Emergence of Risksapes
Detlef Müller-Mahn
- 19. The Tolerability of Risk**
A New Framework for Risk Management
Edited by Frédéric Boudier, David Slavin and Ragnar E. Løfstedt
- 20. Transboundary Risk Governance**
Rolf Lidskog, Linda Soneryd and Ylva Uggla
- 21. Transboundary Risk Management**
Edited by Joanne Linnerooth-Bayer, Ragnar E. Løfstedt and Gunnar Sjøstedt
- 22. Trust in Cooperative Risk Management**
Uncertainty and Scepticism in the Public Mind
Michael Siegrist, Timothy C. Earle and Heinz Gutscher
- 23. Trust in Risk Management**
Uncertainty and Scepticism in the Public Mind
Edited by Michael Siegrist, Timothy C. Earle and Heinz Gutscher
- 24. Uncertainty and Risk**
Multidisciplinary Perspectives
Edited by Gabriele Bammer and Michael Smithson
- 25. Judgment and Decision Making**
Baruch Fischhoff
- 26. Risk Analysis and Human Behavior**
Baruch Fischhoff

JUDGMENT AND DECISION MAKING

Baruch Fischhoff

earthscan

publishing for a sustainable future
LONDON AND NEW YORK

First published 2012
by Earthscan
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Earthscan
711 Third Avenue, New York, NY 10017

Earthscan is an imprint of the Taylor & Francis Group, an informa business

© 2012 Baruch Fischhoff, selection and editorial material; individual chapters, Baruch Fischhoff and the contributors

The right of Baruch Fischhoff to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Fischhoff, Baruch, 1946-

Judgment and decision making / Baruch Fischhoff.
p. cm. -- (Earthscan risk in society series)

Includes bibliographical references and index.

1. Judgment. 2. Decision making. 3. Risk-taking (Psychology) I. Title.

BF447.F57 2012

153.4'6--dc23

2011027664

ISBN: 978-1-84971-444-0 (hbk)

ISBN: 978-1-84971-445-7 (pbk)

ISBN: 978-0-203-14193-9 (ebk)

Typeset in Bembo
by GreenGate Publishing Services, Tonbridge, Kent

CONTENTS

<i>List of figures</i>	<i>viii</i>
<i>List of tables</i>	<i>ix</i>
<i>Acknowledgments</i>	<i>xi</i>
<i>On the cover</i>	<i>xii</i>
<i>Introduction</i>	<i>xiii</i>
PART I	
BACKGROUND	1
1 Judgment and decision making	
<i>Baruch Fischhoff</i>	3
2 Amos Tversky (1937–96)	
<i>Baruch Fischhoff</i>	23
PART II	
HINDSIGHT BIAS	31
3 Hindsight \neq foresight: the effect of outcome knowledge on judgment under uncertainty	
<i>Baruch Fischhoff</i>	33

4	For those condemned to study the past: reflections on historical judgment <i>Baruch Fischhoff</i>	50
5	An early history of hindsight research <i>Baruch Fischhoff</i>	65
PART III		
JUDGMENT		69
6	Hypothesis evaluation from a Bayesian perspective <i>Baruch Fischhoff and Ruth Beyth-Marom</i>	72
7	Knowing with certainty: the appropriateness of extreme confidence <i>Baruch Fischhoff, Paul Slovic, and Sarah Lichtenstein</i>	103
8	Fault trees: sensitivity of estimated failure probabilities to problem representation <i>Baruch Fischhoff, Paul Slovic, and Sarah Lichtenstein</i>	124
9	Assessing uncertainty in physical constants <i>Max Henrion and Baruch Fischhoff</i>	146
10	Fifty–fifty = 50% <i>Baruch Fischhoff and Wändi Bruine de Bruin</i>	162
PART IV		
EVALUATION		181
11	Predicting frames <i>Baruch Fischhoff</i>	184
12	Value elicitation: is there anything in there? <i>Baruch Fischhoff</i>	207
13	Informed consent for eliciting environmental values <i>Baruch Fischhoff</i>	232

PART V	
DECIDING	247
14 Giving advice: decision theory perspectives on sexual assault <i>Baruch Fischhoff</i>	249
15 The real world: what good is it? <i>Baruch Fischhoff</i>	272
16 Assessing adolescent decision-making competence <i>Baruch Fischhoff</i>	300
17 Questions of competence: the duty to inform and the limits to choice <i>Baruch Fischhoff and Sara Eggers</i>	320
Conclusion	343
<i>Index</i>	346

FIGURES

1.1 Six “laws of the new psychophysics,” depicting the influence of experimental design on the numerical response used to describe the psychological state (ψ) equivalent to a physical stimulus (ϕ)	9
7.1 Some representative calibration curves	104
9.1 Measurements of the velocity of light, 1875–1958	151
9.2 Recommended values for the velocity of light, 1929–1973	153
9.3 Recommended values for fundamental constants, 1952–1973	155
10.1 Distribution of probabilities given in response to the Study 1 question	168
10.2 Distribution of probabilities given in response to the question “What is the probability that you will develop cancer by 40?”	171
12.1 Trends in self-reported happiness, 1971–1973	212
12.2 Estimates of fertility expectations of American women	221
13.1 Materiality standard for prioritizing information	240
13.2 Manipulation checks	243
17.1 Decision tree for men considering saw palmetto (SP) as a treatment for lower urinary track problems	325
17.2 Decision tree for women considering the use of emergency contraception, contingent on suspecting failure of contraceptive method	329
17.3 Steps in the Q850 risk management decision-making process	339

TABLES

1.1 From artifact to main effect	11
1.2 Correlations between decision-making competence (DMC) and other variables	14
1.3 Debiasing methods according to underlying assumptions	16
3.1 Mean probabilities assigned to each outcome, Experiment 1	37
3.2 Mean data relevance judgments for Event A, Experiment 1	39
3.3 Mean probabilities assigned by subjects responding, Experiment 2	42
3.4 Mean probabilities assigned by subjects responding, Experiment 3	45
6.1 Potential sources of bias in Bayesian hypothesis evaluation	76
7.1 Analysis of certainty responses in Experiment 1	106
7.2 Lethal events whose relative frequencies were judged by subjects in Experiments 2 and 3	108
7.3 Percentage of correct answers for major odds categories	110
7.4 Frequency of extreme overconfidence	116
7.5 Percentage wrong with deceptive and nondeceptive items	117
7.6 Deceptive items in Experiment 3	120
7.7 Deceptive items in Experiment 4	120
8.1 Results from Experiment 1: pruning the tree	131
8.2 Effects of focusing subjects' attention on "all other problems"	134
8.3 Effect of presenting varying degrees of detail	136
8.4 Experiment 5: effects of splitting and fusing branches	140
8.5 Experiment 6: judgment of experts	142
9.1 Calibration statistics for measurements of physical constants	152
10.1 Some adolescents' uses of fifty-fifty	164
10.2 Probability responses in Study 1	167

x Tables

10.3 Percentages of 50 responses in Study 2	171
10.4 Probability responses in Study 2	173
11.1 Basic stimulus	187
11.2 Frame and option preferences in Experiment 1	190
11.3 Frame and option preferences in Experiment 2	192
11.4 Frame and option preferences in Experiment 4	195
11.5 Frame and option preferences in Experiment 5	198
11.6 Option preferences for variations of civil defense problem	202
12.1 Risk of misdiagnosis	210
12.2 Three paradigms for eliciting values	216
12.3 Conditions favorable to articulated values	227
13.1 Compatibility between respondent and researcher view of study intent	235
13.2 Framework for defining transactions	236
15.1 Option structure	277
15.2 Interpretations of AIDS risk questions	282
15.3 Coding framework	284
15.4 What people think women are likely to do if assaulted	288
15.5 Transaction analysis insurance/preference	291
16.1 Predictive validity and accuracy of adolescents' expectations in NLSY97	305
16.2 Possible reasons for evaluating goods differently at different times	308
16.3 Option structures in teens' open-ended decision descriptions	310
16.4 Correlations between decision-making competence (DMC) and other variables	314
17.1 Predicted optimality of consumer choices with alternative saw palmetto labels	327

ACKNOWLEDGMENTS

I am grateful to Mikey Jones, Alison Kuznets, Ragnar Löfstedt, Charlotte Russell and Rosa Stipanovic for their help in making this volume happen; to my teachers and mentors in the Detroit Public Schools, Wayne State University, and the Hebrew University of Jerusalem, for providing my education; to the governments and others that supported those institutions and my research; to my family for their love and support; and to all for helping me to see research as a form of public service.

ON THE COVER

In *Overgrown* the elk as ruminant encourages the viewer to ruminate across and on the image. As we traverse the branching nexus of antlers, entirely new and magical worlds open up. The experience is that of excited electrons surging through perpetually bifurcating pathways, forks in roadways that will never intersect, though trails extend toward each other. Endless rhizomatic possibility is promised. Yes, these are the branching decision trees, the individual trains of thought and the social networks that combine or proliferate to become the collective experience. The image conveys both the implications and consequences of individual decisions, if considering the elk head as undivided consciousness, and the tethering of myriad positions to a sovereign no longer central control, if looking through a political lens. The piece pivots on seeing a system so grand, a security system so exhaustive that it crashes under its own weight. It also conveys personal pain, anguish, turmoil, psychic trauma and physical paralysis. Considering the implications and risks of a potential decision or trying to retrace the (mis)steps back to a point of origin can be as damning as it can be enlightening.

Andrew Ellis Johnson

Overgrown, 1997

Oil on Belgian linen, 78" × 62"

<http://artscool.cfa.cmu.edu/~johnson/>

INTRODUCTION

From early in my career, I have been fortunate to have colleagues who valued research that was informed by both analytical and empirical sciences. Indeed, before I even had a career, Francine Wehmer and Ira Firestone, two psychology professors at Wayne State University, matched me with Samuel Komorita, as a research assistant for his pioneering work in what is now called behavioral game theory. At the time, I was too much of a novice to appreciate Sam's innovativeness in asking how people actually responded to prisoner's dilemma games—although I did realize that it was a suitable match for a math-psych major.

I had even less appreciation of the uniqueness of the guidance that I received from David Jonah, my advisor in the mathematics department, about the role of empirical knowledge in the evolution of that analytic discipline. As personal gifts, he gave me Otto Neugebauer's *The exact sciences in antiquity*, Carl Friedrich Gauss's *Disquisitiones Arithmeticae*, and G.H. Hardy's *A mathematician's apology*. I finished only the last, but from their gist (and from his confidence in me), I inferred that academic fields had no natural bounds, and could be informed by both thought and experience. My family prized education, but had little familiarity with academia. (Our father, an industrial chemist, was the first to go to high school, also at Wayne, along with my mother, sister, and brother.) The time being the Sixties also supported my feeling that anything was possible in academia—without being too specific about what that might be.

After graduating Wayne and three years of kibbutz living, I had the extraordinary good fortune to have Amos Tversky accept me as a student, bringing me into a world where anything was actually possible—for someone with his intellect and independence of mind. Soon afterward, he began his collaboration with Danny Kahneman, marked by their joint willingness to challenge any orthodoxy. At that time, the Hebrew University of Jerusalem was a remarkable place, in part because the university community included people who had grown up together, hence could reach across fields in ways that would have

been impossible without that long-standing trust. Philosophers, statisticians, historians, engineers, and others dropped in on Amos and Danny's weekly seminars on judgment and decision making.

For me, this liberating atmosphere was amplified by two unrelated seminars. One was Danny's course in "Applications of Psychology," where each week's assignment required addressing an applied problem (e.g., designing hard-to-counterfeit bills, resolving conflicts in a cooperative farming village), using whatever psychology had to offer. The seminar taught me that psychology's validity depends on its usefulness, that no single theory or method could solve any serious problem, and that the search for solutions required both analysis and evidence. In Alan Baddeley's insightful formulation, psychology progressed through a combination of applied basic research, seeing how well basic results hold up in applied settings, and basic applied research, extracting new fundamental problems from applications.

The second unrelated seminar was an informal gathering of graduate students from different departments, convened by Reuven Kaminer, associate dean for foreign students (and a fellow Detroit). Our main text was *The Sociological Imagination*, by C. Wright Mills (1959). Following Mills, we took our "personal worlds" (here, our chosen disciplines) as points of departure for examining our social worlds (here, the pressures shaping research questions). Although I remember few details (beyond the critique of Talcott Parsons as mind-numbingly dull), I absorbed something of Mills's turn of mind (or "imagination"), in which the scientific enterprise depends on continuously challenging how research problems are formulated.

The nascent state of decision-making research made it easier to adopt a challenging frame of mind. Economists, psychologists, statisticians, philosophers, and others had not yet sorted out their respective roles—sometimes even collaborating (as seen in Edwards, Lindman and Savage, 1963, and Edwards and Tversky, 1967). Sarah Lichtenstein's comprehensive "List" of researchers committed to collecting behavioral data framed in analytical terms had fewer than 100 members worldwide. Even that number included several in Eastern Europe whose opportunities were limited by their use of such suspect terms as "utility" and "subjective" probability. (As Stalin said, "We don't forecast, we plan.")

A further spur to my pursuing new research topics was the poor quality of scientific communication in those days. Far off in the Middle East, it seemed best to work on problems that no one else was likely to touch. My somewhat unusual preference for topics with both analytical and empirical elements made that somewhat easier. Compared to joining the fray on a hot problem, this research strategy has the benefit of getting to structure messy problems, along with the risk of working on marginal ones. Having stumbled across this strategy more or less as a matter of necessity, I have followed it more or less as a matter of taste. My next stop, working with Robyn Dawes, Lita Furby, Sarah Lichtenstein, and Paul Slovic, in Eugene, Oregon, afforded me superb help in cultivating these tastes. At Carnegie Mellon University, they are a norm.

The University views its strategic advantage as letting faculty do whatever they want, as long as they can eventually convince some serious audience of the value of their work.

The articles collected in this volume reflect my efforts to take advantage of the extraordinary mentoring that I've received and the freedom that academic life affords us. A companion volume collects basic applied research regarding a wide variety of risks ranging from waterborne diseases and nuclear power to sexual assault and adolescent sexuality. Given the diversity of these topics and the lack of boundaries on "real world" problems, my work has taken me into the fundamental topics around which this collection is organized: judgment, preference, and decision making. Along the way, I have had the further good fortune of wonderful collaborators. I hope that others have similar opportunities.

References

- Baddeley, A. 1979. Applied cognitive and cognitive applied research. In L. G. Nilsson (Ed.), *Perspectives on memory research*. Hillsdale, NJ: Erlbaum.
- Edwards, W., Lindman, H., and Savage, L.J. 1963. Bayesian statistical inference for psychologists. *Psychological Review*, 70, 193–242.
- Edwards, W. and Tversky, A. (Eds.). 1967. *Decision making*. Wadsworth: Penguin.
- Gauss, C.F. 1801, 1966. *Disquisitiones Arithmeticae* (translated by A.A. Clarke). New Haven: Yale University Press.
- Hardy, G.H. 1940. *A mathematician's apology*. Cambridge: Cambridge University Press.
- Mills, C.W. 1959. *The sociological imagination*. Oxford: Oxford University Press.
- Neugebauer, O. 1957. *The exact sciences in antiquity*. Mineola, NY: Dover.

PART I

Background

The field currently called behavioral decision research was first identified by Ward Edwards (1954), in “A theory of decision making.” Writing but a decade after von Neumann and Morgenstern’s classic (1944) exposition of utility theory (see also Savage, 1954), Ward showed the essential role of behavioral research in translating utility theory from a brilliant abstraction to a practical way to understand and improve actual behavior. The drift in the field’s name, through the intermediate behavioral decision making, reflects the growing realization that no single theory could cover the diversity of decision making. The field has, however, retained its unifying commitment to examining choices from both analytical and empirical perspectives, with utility theory as a point of departure for determining how decisions should be made. Decision analysis has always been the term for approaches trying to help people see their decisions in utility theory terms (Clemen and Reilly, 2003; Hammond, Keene and Raiffa, 1999; Raiffa, 1968).

“Judgment and decision making” (Chapter 1) offers a historical overview of the field. In the spirit of Ward’s vision, it emphasizes the interplay between normative analysis of decisions’ formal properties, descriptive research into how people actually behave, and prescriptive interventions for bridging the gap between the normative ideal and the descriptive reality. When interventions fail, any of the three stages could be at fault. The normative analysis might not have identified the issues that matter to the decision makers. The descriptive research might not have captured their intuitive perspectives faithfully. The intervention might not have addressed their problems well. The chapter traces research in confidence assessment, preference formation, clinical judgment, and other topics, often following the seminal formulations from the University of Michigan’s mathematical psychology group, which included

Edwards, Clyde Coombs, Robyn Dawes, David Krantz, and Anatol Rapoport, among others. The chapter concludes by analyzing alternative default assumptions about where problems lie when decisions seem wrong.

“Amos Tversky” (Chapter 2) describes some of the path-breaking research of one of my graduate advisors, Amos Tversky, often in collaboration with my other advisor, Daniel Kahneman. That research exemplifies formulating problems in clear analytical terms, broadly informed by empirical research. The brief summaries here provide context for the chapters that follow, along with a small measure of gratitude for the advising that I received.

References

- Clemen, R.T., and Reilly, T. (2003). *Making hard decisions*. Boston: Duxbury.
- Edwards, W. (1954). A theory of decision making. *Psychological Bulletin*, 54, 380–397.
- Edwards, W., Lindman, H.R., Savage, L.J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, 70, 193–242.
- Hammond, J.S., Keeney, R.L., and Raiffa, H. (199). *Smart choices*. New York: Broadway.
- Raiffa, J. (1968). *Decision analysis*. Belmont, MA: Addison-Wesley.
- Savage, L.J. (1954). *The foundations of statistics*. Wiley: New York.
- von Neumann, J., and Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

1

JUDGMENT AND DECISION MAKING

Baruch Fischhoff

© 2010 JOHN WILEY AND SONS, LTD. WIREs COGN Sci

Abstract

The study of judgment and decision making entails three interrelated forms of research: (1) *normative analysis*, identifying the best courses of action, given decision makers' values; (2) *descriptive studies*, examining actual behavior in terms comparable to the normative analyses; and (3) *prescriptive interventions*, helping individuals to make better choices, bridging the gap between the normative ideal and the descriptive reality. The research is grounded in analytical foundations shared by economics, psychology, philosophy, and management science. Those foundations provide a framework for accommodating affective and social factors that shape and complement the cognitive processes of decision making. The decision sciences have grown through applications requiring collaboration with subject matter experts, familiar with the substance of the choices and the opportunities for interventions. Over the past half century, the field has shifted its emphasis from predicting choices, which can be successful without theoretical insight, to understanding the processes shaping them. Those processes are often revealed through biases that suggest non-normative processes. The practical importance of these biases depends on the sensitivity of specific decisions and the support that individuals have in making them. As a result, the field offers no simple summary of individuals' competence as decision makers, but a suite of theories and methods suited to capturing these sensitivities.

Introduction

Decisions are easy when decision makers know what they want and what they will get, making choices from a set of well-defined options. Such decisions could

be equally easy, but reach different conclusions, for people who see the facts similarly, but have different goals, or for people who have the same values but see the facts differently, or for people who disagree about both facts and values.

Decision making can become more difficult when there is uncertainty about either what will happen or what one wants to happen. Some decisions are so sensitive to estimates of fact or value that it pays to invest in learning, before acting. Other decisions will work out just as well, for any plausible estimates.

Thus, any account of decision-making processes must consider both the decisions and the individuals making them. The field of *behavioral decision research* provides such accounts. It entails three forms of research: (1) *normative*, identifying the best possible choice, given the state of the world and decision makers' values; (2) *descriptive*, characterizing how individuals make decisions, in terms comparable to the normative standard; and (3) *prescriptive*, attempting to close the gap between the normative ideal and the descriptive reality.

Although they can be described as an orderly progression, these three forms of research are deeply interrelated. Descriptive research is needed to reveal the facts and values that normative analysis must consider. Prescriptive interventions are needed to assess whether descriptive accounts provide the insight needed to improve decision making. Normative analyses are needed to understand the facts that decision makers must grasp and the practical implications of holding different values. Thus, understanding choices requires an iterative process, cycling through the three stages. This chapter follows the evolution of theory and method for seeking that understanding.

Behavioral decision research

Behavioral decision research emerged from normative models of decision making developed by philosophers, mathematicians, and economists.^{1,2} These models describe how to determine the best possible course of action, given what individuals believe about the world and what they want from it. Individuals who follow these rules are said to be *rational*. Their choices are *optimal*, if they are well informed about the world and about their own values. Although normative models take strong positions on *how* decisions should be made, they are mute regarding *what* options, facts, and values should be considered. As a result, they require the empirical content provided by descriptive and prescriptive research to be anything but formalisms.

Comparability with normative analysis imposes important constraints on descriptive and prescriptive research. They cannot begin without first examining the world from decision makers' perspective. They cannot criticize choices without asking whether they might be rational, given what people want and believe. They cannot assess the importance of imperfections in decision-making processes, without embedding them in normative analyses, showing their practical implications. Imperfections can be theoretically informative without

matter much. Indeed, nonrational processes may survive because they have too little practical significance to provide the sharp negative feedback sometimes needed to change behavior.

Psychology progresses, in part, by applying what Berkeley and Humphreys³ call the “bias heuristic,” identifying departures from normative standards.⁴ However, unless those standards are well defined, vaguely similar biases may proliferate. Different biases might share a common name (e.g., confirmation bias); the same bias might have different names (e.g., saliency, availability), impeding scientific progress.⁵ Indeed, as discussed next, a major advance in early behavioral decision research was discovering that seemingly different theories were often indistinguishable.

Clinical judgment

World War II was a turning point for psychology, which showed its ability to assess efficiently the skills and problems of masses of individuals. After the war, attention turned to the effectiveness of those efficient assessments. These studies of *clinical judgment* quickly spread to topics as diverse as how psychologists decide whether clients are “neurotic” or “psychotic,” radiologists sort ulcer X-rays into “benign” or “malignant,” bank officers classify loans as “nonperforming,” and brokers weigh stocks’ prospects.^{6–8}

Conducting studies of clinical judgment is straightforward: collect many judgments of cases described on a common set of possibly relevant *cues*. Use statistical methods (e.g., multiple regression) to predict those judgments from the cues. For example, Dawes⁹ studied University of Oregon Psychology Department graduate admission committee evaluations of 384 applicants. Although applicants’ files had many cues (e.g., letters of recommendation, full transcripts), the committee’s ratings could be predicted well from just three: Graduate Record Examination (GRE) score, undergraduate grade point average (GPA), and quality of undergraduate institution (QI):

$$0.0032 \text{ GRE} + 1.02 \text{ GPA} + 0.0791 \text{ QI} \quad (1)$$

This study illustrates four frequently replicated patterns:^{9,10} (1) A simple model predicts a seemingly complex process. (2) Judges describe using very different strategies than that “captured” in the model. For example, committee members claimed that they considered more cues and used these three in more nuanced ways than just weighting and adding. (3) Even simpler models, replacing regression weights with unit weights on normalized variables, predict equally well. (4) Simple models predict the actual criterion (graduate school success) well.

There are at least three reasons why simple models predict surprisingly well. One is that people have difficulty introspecting into their own decision making.^{11,12} A second is that people have difficulty executing complex strategies reliably, so that only simple patterns appear consistently. The third is that simple linear models can

predict well without capturing the underlying processes,^{9,13} as long as they use reliably measured correlates of the variables that actually affect decision making.

This good news for predictive research is bad news for explanatory research. Models using different variables, implying different processes, often predict equally well. As a result, regression weights need not capture how decisions are made. In many applications, good prediction suffices. For example, the health belief model^{14,15} provides a structured way to identify variables correlated with health-related choices. Its application would, however, be misguided, if the weights on those variables were taken as reflecting how individuals think.¹⁶

For analogous reasons, behavioral decision researchers typically avoid the *revealed preference* analyses that are a staple of economics research.^{17,18} For goods traded in efficient markets, prices show rational decision makers' values. If goods are characterized on common attributes, regression weights show those attributes' usefulness as predictors. For example, house prices might be predicted from their size, condition, age, school district, commuting distance, construction, and so on. Unfortunately, when predictors are correlated, regression weights can be unstable, complicating their interpretation as measures of importance.

One strategy for undoing these confounds is generating stimuli with uncorrelated cues. For example, one might create hypothetical graduate school candidates, using all possible combinations of GRE, GPA, and QI. A drawback to this *ANOVA design* is violating behavioral decision research's commitment to *probabilistic functionalism*,^{4,19} the view that behavior is shaped by naturally occurring correlations among uncertain cues. Stimuli that violate these relationships lack ecological validity and require unnatural behavior, such as evaluations of implausible cue combinations (e.g., low GRE, high QI). An ANOVA design also gives equal weight to all cue combinations, however common or possible. Moreover, as with any design that presents many stimuli with a transparent cue structure, respondents may either lose focus (producing unreliable judgments) or improvise a mechanical response strategy (producing reliable, but unnatural judgments).

How people respond to novel tasks (e.g., grad candidates with low GRE and high QI) can be revealing. However, because *importance* is inherently context dependent, artificial contexts produce artificial importance weights. For example, although money is generally relevant to consumer decision making, other factors may dominate choices among similarly priced options. One possible reason why Eq. (1) did not include the variable "strength of letters of recommendation" is that candidates had similarly strong letters, written by faculty advisors who sell their students similarly. (QI should capture the reputations of those letter writers.)

As it discovered these limits to the explanatory value of predictive models, behavioral decision research shifted its focus from *what* choices people make to *how* they make them. As a result, studies describe decision-making processes that *can* come into play, as revealed by tasks with which it is relatively clear how a process would express itself. Applied researchers must then determine which of the possible processes are evoked by a specific decision.

Subjective expected utility

The normative analysis underlying behavioral decision research is founded on *expected utility theory*, classically codified by von Neumann and Morgenstern.²⁰ Its basic logic is straightforward: List the possible action options. For each option, enumerate its possible outcomes. For each such outcome, assess the value, or *utility*, of it happening. Assess the probability of its occurrence should each option be selected. Compute the *expected utility* of each option by multiplying the utility and probability of each outcome, should it be undertaken, then summing across outcomes. Choose the action with the greatest expected utility. When the probabilities reflect decision makers' beliefs, rather than scientific knowledge, the calculation produces *subjective expected utility*.²¹ (As discussed below, some scholars view all probabilities as subjective.)

Descriptive research can look at how people undertake each element of this process: assessing the probabilities of possible outcomes, evaluating their utility (should they occur), and combining probabilities and utilities to identify the best option. The decisions can range from completely described and static to incompletely described and dynamic. Normative analyses exist for many kinds of decision.²²

Individuals' performance on these tasks can be evaluated by *correspondence* or *coherence* tests. Correspondence tests ask how accurate their answers are. For example, how well can they predict whether they will graduate college or enjoy their major? Coherence tests ask how consistent responses are. For example, are probability judgments for an event at least as large as those for a subset ($p[A] \geq p[A \cap B]$)? Are outcomes equally valued, when described in formally equivalent ways (e.g., succeeding vs. not failing).

Predicting outcomes

Studies of how well people predict uncertain events have produced seemingly contradictory results. Sometimes, people do quite well; sometimes, quite poorly. To a first approximation, the difference depends on whether the task requires counting or inferences. With counting studies, the evidence is all of one type; with inference studies, the evidence is of different types. A counting study might display stimuli drawn randomly from a hidden population, then elicit estimates of a summary statistic (e.g., mean, range).²³ An inference study might require integrating *base-rate* evidence about what usually happens, with *individual information* about a specific case.

Counting tasks take advantage of individuals' ability to estimate the relative frequency of events that they observe—even without preparing to do so. For example, after producing rhymes for a set of words, people can estimate the number beginning with different letters.²⁴ Indeed, encoding frequencies has been called an automatic cognitive function, with research focusing on whether it relies on *tokens*, records of individual observations, or on *types*, category representatives reinforced with each observation.²⁵

Assuming that individuals trust their frequency-encoding ability, Tversky and Kahneman²⁶ proposed the *availability heuristic*, whereby individuals estimate an event's probability by their ability to retrieve instances (tokens) or imagine them (types). Reliance on availability produces biased judgments when the observed events are an unrepresentative sample—and individuals cannot correct for the sampling bias. Researchers have identified many other possible biases, arising from reliance on judgmental heuristics. The strength of any claim of bias depends on the strength of the normative analysis.^{27,28} The usefulness of any heuristic depends on how well its application can be predicted (e.g., how memory is searched for examples).

Inference studies tap individuals' lack of intuition and training for combining different kinds of evidence. Here, the normative standard has been the Bayesian approach to hypothesis evaluation.^{5,29} Bayes theorem is an uncontroversial part of probability theory. Bayesian inference is more controversial, because it treats probabilities as subjective, thereby allowing inferences that combine diverse kinds of evidence.³⁰ *Frequentistic* probabilities require evidence of a single kind (e.g., coin flips, weather records). Subjective judgments are only probabilities if they pass coherence tests. Thus, probabilities are not just any assertion of belief.

A widely studied inferential bias is the “base-rate fallacy”. Attributed to reliance on the representativeness heuristic,²⁶ it involves allowing even weak information about specific cases to outweigh knowledge of what generally happens (the base rate). Inadequately regressing judgments is the same bias with continuous variables. Absent strong information about specific cases, one should predict the mean of a distribution.

To avoid artifactual sources of bias,³¹ behavioral decision research draws on the century-plus of psychophysics research into factors affecting quantitative judgments.^{32–34} For example, because people avoid decimals, they are more likely to overestimate small risks in studies eliciting percentages (e.g., 0.1%) than in studies eliciting odds (e.g., 1 in 1000).³³ Knowing that, researchers can choose the method best suited to their question and reduce measurement artifacts.

Figure 1.1 depicts six such design features, critical to eliciting numbers. For example, Figure 1.1A shows that stimuli [S1, S2] elicit less of the response range when embedded in a larger range [L1, L2]. Figure 1.1C shows how values assigned to larger stimuli are cramped if the initial (standard) stimulus is large, relative to others in the set. Such effects occur because respondents must translate their perceptions into the investigators' terms. Where those terms are unnatural, respondents rely on *response preferences*.³⁵ For example, they try to use the entire response scale; they look for patterns in randomly ordered stimuli; they deduce an expected level of precision, such as what trade-off to make between speed and accuracy.

Ignoring response preferences leads to misinterpreting judgments. For example, subjects produced much higher estimates of annual US death toll, from⁴¹ causes,

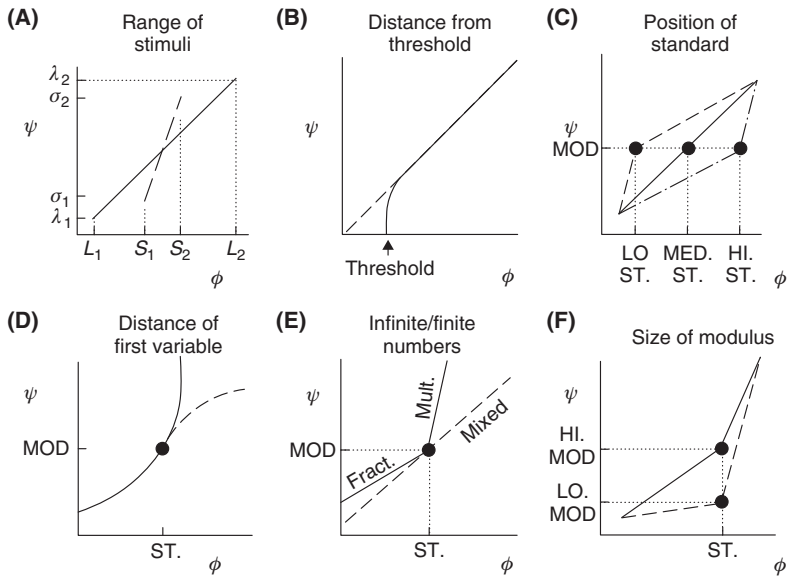


FIGURE 1.1 Six “laws of the new psychophysics,” depicting the influence of experimental design on the numerical response used to describe the psychological state (Ψ) equivalent to a physical stimulus (Φ). (A) A narrower stimulus range (S_1, S_2) will use a proportionately larger portion of the response range than would the same stimuli, when embedded in a larger response range (L_1, L_2). (B) The effects of assumptions regarding the treatment of stimuli below the threshold of perception or evaluation. (C) The effects of where a standard stimulus falls in the response range, after it has been assigned a numerical valuation (or modulus). (D) The effects of where the first judged stimulus is relative to the standard. (E) The effects of using fractional or integer response values, for stimuli smaller than the standard. (F) The reverse effects where a modulus value, for a given standard stimulus, falls within the response range (Reprinted with permission from Ref. 17 Copyright 2005 Elsevier).

when they received a high *anchor* (50,000 motor vehicles deaths), rather than a low anchor (1000 accidental electrocutions). Low frequencies were greatly overestimated with the high anchor, much less so with the low one. Estimates of relative frequency were similar, however the question was asked, suggesting robust risk perceptions, whose translation into numerical judgments was method dependent. The estimates were also biased in ways consistent with relying on the availability heuristic (e.g., homicides were overestimated relative to, less reported, suicides).³⁶

Eliciting values

There are two streams of research into how people form preferences.¹⁷ One follows psychophysics, treating the intensity of preferences like the intensity of physical experiences. The second follows the precepts of *decision analysis*, a consulting process designed to help individuals follow decision theory's normative model.^{21,22}

Research in the psychophysical stream has individuals report their feelings directly, perhaps with a rating scale or a judgment of willingness-to-pay for a good. Attitude research is the archetype of this paradigm.

The correspondence test for psychophysical research asks how well elicited values predict behavior. Some attitude researchers hold that a fair test must elicit attitudes that are directly comparable to the target behavior.³⁷ For example, many behaviors could follow endorsement of "my faith is very important to me." Stronger predictions follow from "daily prayer with like-minded worshipers is very important to me." Even stronger predictions follow from specifying the form of worship. At the extreme, these judgments become statements of intention, rather than attitudes, representing general values. As such, their validity depends on how well people can predict their own experiences.³⁸

The coherence standard for psychophysical judgments is *construct validity*. Expressed values should be sensitive to relevant changes in questions and insensitive to irrelevant ones. Applying this standard requires independently assessing relevance. For example, assuming that more is better, *scope tests* ask whether people put higher values on larger quantities of a good. Scope insensitive judgments represent incoherent preferences—except for individuals who feel that there can be too much of a good thing (e.g., rich food, conspicuous consumption). An "inside view" on individuals' basic values is needed to evaluate the coherence of their preferences.

Research in the decision analysis stream assumes that people cannot know what they want, in all possible situations. Rather, they must *construct* specific preferences from more basic values. In making these inferences, people may seek cues in a world that might be helpful, indifferent, or manipulative. The better people understand the factors shaping their inferences, the better chance they have of figuring out what they want.^{39,40} Decision analysis structures that process. Its measurement is *reactive*, in the sense of changing people in the process of trying to help them discover their preferences. If successful, it deepens individuals' understanding of themselves.

Correspondence tests for constructed preferences compare elicited values with those that emerge from similar real-world processes. Thus, an intensive electoral campaign might be the standard for a study eliciting candidate preferences.^{41,42} Intensive medical consultation might be the standard for preferences elicited with a medical decision aid.^{43,44} Coherence tests for constructed preferences ask whether the elicitation session has included all perspectives that

individuals might want to consider, while avoiding ones that would apply irrelevant influences.

Identifying the factors influencing behavior is, of course, psychology's central challenge. To study theoretically relevant factors, researchers must control irrelevant ones. Understanding these processes is an ongoing

TABLE 1.1 From artifact to main effect

<i>Liability in judgment due to</i>	<i>Led to</i>
Organism Inattention, laziness, fatigue, habituation, learning, maturation, physiological limitations, natural rhythms, experience with related tasks	Repeated measures Professional subjects Stochastic response models Psychophysiology Proactive and retroactive inhibition research
Stimulus presentation Homogeneity of alternatives, similarity of successive alternatives (especially first and second), speed of presentation, amount of information, range of alternatives, place in range of first alternative, distance from threshold, order of presentation, areal extent, ascending or descending series	Classic psychophysical methods The new psychophysics Attention research Range-frequency theory Order-effects research Regression effects Anticipation
Response mode Stimulus-response compatibility, naturalness of response, set, number of categories, halo effects, anchoring, very small numbers, response category labeling, use of end points	Ergonomics research Set research Attitude measurement Assessment techniques Contrasts of between-and within-subject design Response-bias research Use of blank trials
"Irrelevant" context effects Perceptual defenses, experimenter cues, social pressures, presuppositions, implicit payoffs, social desirability, confusing instructions response norms, response priming, stereotypic responses, second-guessing	New look in perception Verbal conditioning Experimenter demand Signal-detection theory Social pressure, comparison, and facilitation research

Reprinted with permission from Ref. 46 Copyright 1980 L. Erlbaum Associates.

enterprise, which McGuire⁴⁵ depicted as turning “artifacts into main effects,” worthy of independent investigation. Table 1.1 assembles parts of this history, in terms of the four essential elements of any behavior: the organism, the stimulus being evaluated, the response mode for expressing preferences, and potentially distracting contexts.⁴⁶ In terms of correspondence tests, these are all factors that could undermine the match between the conditions in which values are measured by researchers and expressed in life. In terms of coherence tests, these are all factors whose effects on expressed values could be compared with independent assessments of their relevance. That is, do changes in these factors affect valuations when, and only when, they should make a difference? Given the sheer number of potentially relevant factors, value elicitation requires broad understanding of behavioral science.

Making decisions

Non-normative theories

Knowing the limits to the theoretical insights possible with predictive models, applied in complex settings, behavioral decision researchers have focused on processes observed most clearly under experimental conditions. The robustness of observations in the lab is tested by varying those conditions (e.g., increasing economic incentives for good performance, changing information displays) and by identifying real-world analogs, in which a theoretically interesting process might play a practical role.

Foremost among these models is Kahneman and Tversky’s⁴⁷ *prospect theory*. Its initial formulation identified several utility theory assumptions that were implausible psychologically. One is that people evaluate expected outcomes in terms of changes in their *net asset position*, namely, everything they have in the world. However, people are actually highly sensitive to changes and tend to forget the big picture—as witnessed in reminders to “count your blessings.”⁴⁸ A second psychologically implausible assumption is that numerically equivalent changes in probabilities are equally important. However, the psychophysics of probability weighting places a premium on changes that lead to certain outcomes (e.g., from 90% to 100%) compared to mid-range changes (e.g., from 30% to 40%). A third such assumption is that people get increasingly averse as losses mount up, whereas psychology finds them increasingly apathetic.

One widely studied corollary of these principles is the *status quo bias*. It reflects how easily *reference points* can be shifted, varying how changes are viewed. For example, organ donation rates are much higher when drivers must opt out, when getting their drivers licenses, compared to when they must opt in.^{49,50} Opting in makes surrendering organs seem like a loss, hence aversive. That formulation also suggests a social norm of organ donation and perhaps even a weaker right to refusal.

Any behaviorally realistic approach to decision making must accommodate the limits to cognitive computational capacity. Prospect theory accepts utility theory's cognitively implausible calculation of expected values. However, it uses more intuitively plausible elements and, as a linear model, is relatively robust to misestimating its parameters. Applying the theory requires identifying its elements with real-world equivalents, such as the reference points that decision makers use when assessing changes.⁵¹ Fuzzy-trace theory⁵² studies the processes by which individuals master the gist of recurrent decisions. Approaches building on the classic work of Herbert Simon⁵³ have examined individuals' ability to match simple decision-making heuristics to choices that would, otherwise, be unduly complex.⁵⁴ Query theory,⁵⁵ support theory,⁵⁶ and others⁵⁷ formalize the notion of weighting retrieved beliefs, embodied in the availability heuristic.

Emotions

Normative analyses can accommodate emotions as valued outcomes, such as the utility of being happy or the disutility of being fearful. For example, there are formal methods for incorporating such "psychological" outcomes, in analyses of risk decisions.^{17,31} Descriptive research can accommodate emotions in terms of their effects on each element of decision making (defining options, predicting events, assessing personal values, integrating beliefs and values). For example, *cognitive appraisal theory*⁵⁸ predicts that anger increases the perceived probability of overcoming problems. In a field test with a nationally representative US sample, Lerner et al.⁵⁹ found that respondents were about 5% more optimistic, regarding their vulnerability to terror-related events, after an anger induction than after a fear induction. Prescriptive research can accommodate emotions by helping people to getting the right mix for particular choices.^{60,61} For example, formal analyses might be used cautiously when they "anaesthetize" moral feeling⁶²; decision aids for adolescents have focused on controlling emotions.⁶³

The importance of emotion effects depends on their size. A 5% shift might tip a close decision, but not a clear-cut one. von Winterfeldt and Edwards²¹ showed, mathematically, that decisions with continuous options (e.g., invest \$X) are often insensitive to changes in input variables (i.e., probabilities, values). Thorngate⁶⁴ used simulations to examine the sensitivity of stylized decisions to errors due to imperfect heuristics, an approach that others have pursued extensively.^{65, 66}

Decision-making competence (DMC)

The fundamental premise of experimental decision research is that people who master the skills that it studies make better real-world decisions.⁶⁷ Table 1.2 presents results from a study evaluating the external validity of seven experimental tasks, chosen to span the space of cognitive decision-making competencies.⁶⁸

TABLE 1.2 Correlations between decision-making competence (DMC) and other variables

DMC correlated with	Pearson <i>r</i>	Semi-partial correlation, controlling for		
		Vocabulary	ECF	Vocabulary and ECF
Cognitive ability				
Vocabulary	.50	—	.28	—
ECF	.48	.26	—	—
Overall	<i>p</i> < .0001	<i>p</i> = .0009	<i>p</i> = .0008	—
Cognitive style				
Polarized thinking	-.34	-.20	-.24	-.19
Self-consciousness	.20	.14 ^b	.05	.11
Self-monitoring	.24	.29 ^b	.30 ^b	.32
Behavioral coping	.32	.27 ^a	.28 ^a	.26
Overall	<i>p</i> < .0001	<i>p</i> < .0001	<i>p</i> < .0001	<i>p</i> < .0001
Risk behavior				
Antisocial disorders	-.19	-.18 ^b	-.05	-.09
Externalizing behavior	-.32	-.28 ^b	-.18	-.20
Delinquency	-.29	-.28 ^b	-.18	-.21
ln(lifetime # of drinks)	-.18	-.22 ^b	-.15	-.18
ln(lifetime marijuana use)	-.25	-.30 ^b	-.20	-.25
ln(# times had sex)	-.24	-.30 ^b	-.21	-.27
ln(# sexual partners)	-.30	-.33 ^b	-.29 ^a	-.31
Overall	<i>p</i> = .0004	<i>p</i> = .0002	<i>p</i> = .009	<i>p</i> = .002
Social and family influences				
Risk status (HAR = 1; LAR = 0)	-.35	-.27	-.23	-.21
SES	.35	.20	.21	.15
Social support	-.30	-.21	-.23	-.19
Positive peer environment	.33	.35 ^b	.32 ^a	.35
Overall	<i>p</i> = .0002	<i>p</i> = .002	<i>p</i> = .006	<i>p</i> = .007

ECF = executive cognitive function; HAR = high risk family; LAR = low risk family; SES = socio-economic status.

^a Test A rejects the one-mediator null hypothesis.

^b Test B rejects the one-mediator null hypothesis.

Reprinted with permission from Ref. 69 Copyright 2005 John Wiley and Sons, where the tasks are described more fully.

Respondents were 110 18- to 19-year-old males in a longitudinal study involving extensive assessments beginning at age 10. DMC scores, extracted from a factor analysis of performance on the seven tasks, showed good test-retest reliability, as did scores on an adult version.⁶⁹

The first section shows positive correlations between DMC and standard measures of verbal and fluid intelligence (Vocabulary and ECF, respectively). The second section shows positive correlations between DMC and four “constructive” cognitive styles. The third section shows negative correlations between DMC and several important risk behaviors. The fourth section shows that DMC is higher for teens coming from low-risk (LAR) families, higher socioeconomic status (SES) families, and more positive peer environments. (The negative correlation with social support may reflect low DMC teens’ greater gang membership.) Most correlations remained statistically significant after partialing out the two intelligence measures.

These results support the construct validity of DMC as a measure of decision-making skills that both cause and reflect important aspects of teens’ lives. For example, teens with higher DMC come from families that might both model and reward good decision making. Bruine de Bruin et al.⁶⁹ found similar correlations between adult DMC and scores on a psychometrically validated Decision Outcome Inventory, eliciting self-reports of outcomes suggesting poor decisions, varying in severity (threw out food, bought clothes that were never worn, missed a train or bus, had a mortgage foreclosed, had a driver’s license revoked, had an unplanned pregnancy) and inversely weighted by their frequency.

Prejudices about biases—and the rhetoric of competence

Over the past 40 years, the study of judgment and decision making has spread widely, first to social psychology,⁷⁰ then to application areas like accounting, health, and finance, finally penetrating mainstream economics under the banner of behavioral economics. That success owes something to the power of the approach, which liberated researchers previously bound by rational-actor models for describing behavior. It also owes something to the fascination of results that address a central aspect of the human condition, individuals’ competence to manage their own affairs.^{50,66,67} Very different social institutions may suit rational actors (e.g., free markets, civic engagement) and irrational ones (e.g., strong regulatory protection, deference to paternalistic experts).

Those seeking to extract general messages from this complex research literature have adopted several archetypal rhetorical stances. Familiarity with these stances can help in seeing the research through the stances. [Table 1.3](#) summarizes several common themes, formulated in terms of their advocates’ interpretation of the demonstrations of bias that tend to dominate the field.

It is not true

Examining research for possible flaws is central to any science. However, as seen in [Table 1.1](#), the set of features that might conceivably change a research

TABLE 1.3 Debiasing methods according to underlying assumptions

<i>Assumption</i>	<i>Strategies</i>
Faculty tasks Unfair tasks	Raise stakes Clarify instructions Dispel doubts Use better response modes Discourage second guessing Ask fewer questions
Misunderstood tasks	Demonstrate alternative goal Demonstrate semantic disagreement Demonstrate impossibility of task Demonstrate overlooked distinction
Faulty judges Perfectible individuals	Warn of problems Describe problem Provide personalized feedback Train extensively
Incorrigible individuals	Replace them Recalibrate their responses Plan on error
Mismatch between judges and task Restructuring	Make knowledge explicit Search for discrepant information Decompose problem Consider alternative situations Offer alternative formulations
Re-education	Rely on experts Educate from childhood

Reprinted with permission from Ref. 71 Copyright 1982 Cambridge University Press.

result is very large, allowing endless criticisms by those who dislike a result (“the bias would disappear had you just changed ...”). Such radical skepticism may be met by radical counter-skepticism (“you can’t test for every conceivable confound”). A compromise asks whether confounds have general effects. The “unfair tasks” section of [Table 1.3](#) lists common methodological criticisms (e.g., biases would vanish with higher stakes or clearer instructions). An early review of all studies studying these factors found no effect on hindsight bias or on overconfidence in beliefs. A more recent review found that financial incentives had mixed effects, sometimes improving performance, sometimes degrading it, but most often making no difference.¹³

It is true, but you should not say so

Demonstrations of bias allow researchers, who claim to know the answers, to fault others, who do not. Charging others with incompetence undermines their right to make decisions. As a result, researchers should avoid sweeping statements about human competence and stick to the details of domain-specific studies. They should convey both the “figure” of biases and the “background” of the heuristics producing them. They should recall that optical illusions reveal important properties of vision without hindering most activities.²¹ They should resist those who promote their research because it serves their political ends.

People are doing something quite different—and doing it quite well

Describing decisions as suboptimal presumes a normative analysis, informed by knowledge of what people know and want. Without that analysis, evaluations can be unduly harsh (e.g., charging overconfidence when people have strategically overstated their beliefs) or lenient (e.g., excusing mistakes as attempts to learn by trial and error). Table 1.3’s “misunderstood tasks” section lists some ways that actors and observers can interpret decisions differently. In experiments, *manipulation checks* can assess whether subjects understand tasks as intended. In the world, observers are typically left guessing. For example, there is an unresolved controversy over whether some Americans increased their travel risk, by driving rather than flying, right after the 9/11 attacks. However, the interpretation of their decisions requires knowing how they saw the costs, risks, and hassles of flying and driving. Without evidence on these beliefs, any evaluation of their choices is speculative.

But look at how well people do other things

Claims of bias seem strikingly at odds with the complex tasks that people routinely accomplish (including driving and flying). Perhaps, the biases are just laboratory curiosities, theoretically informative, but of limited practical importance. Or, perhaps the research denies people support that life typically affords them. The “restructuring tasks” section in Table 1.3 lists manipulations that have improved performance under lab conditions. For example, when prompted, people can generate reasons why they might be wrong (reducing overconfidence), ways that events might have turned out otherwise (reducing hindsight bias), or estimates of what normally happens (reducing base-rate neglect). If life provides similar cues, then these “debiasing” studies are most relevant for extrapolation to actual behavior.

Facing the problems

Arguably, by mid-adolescence, most people have the cognitive ability to acquire most of the skills needed to make better decisions.^{52,67,69,72} Whether they do depends on the help that they get. Unfortunately, people often receive little training, feedback, and help in making decisions. Indeed, they often face marketers, politicians, and others trying to manipulate their choices.^{44,73} Table 1.3's "perfectible individuals" section lists strategies that seem able to enhance individuals' decision-making abilities—recognizing that their success, in any specific setting, is an empirical question.^{74,75} The "incurable individuals" section lists ways to live with fallibility. A historical example of recalibration was doubling engineers' chronic underestimates of the repair time for power plants.⁷⁶ A currently popular compromise is "nudging" people toward better decisions, by choosing better default choices (e.g., being an organ donor, contributing to pension plans).⁵⁰

Conclusions

Judgment and decision making research both requires and allows an unusual degree of collaboration among scientists with diverse expertise. The core discipline of behavioral decision research entails familiarity with normative analyses, descriptive studies, and prescriptive interventions. Its execution involves input from experts in the subject matter of specific decisions, the other (social and affective) pressures on them, and the opportunities for change.⁵² For example, Downs et al.⁶¹ helped young women make better sex-related decisions, with an interactive DVD whose content reflected medical research (about sexually transmitted infections), behavioral decision research (about risk perceptions), and social psychology (about self-efficacy).

Behavioral decision research also provides a research platform where theoretical and practical research is mutually reinforcing. In the study of clinical judgment, such interactions showed the predictive power of simple models, a result that was invisible to researchers immersed in domain-specific research. In the study of judgment under uncertainty, these interactions revealed suboptimal strategies that survive because they are good enough to avoid major problems. In the study of value elicitation, they revealed the constructive nature of preference formation, as individuals infer what they want in the novel situations created by life and researchers. In the study of choice, they revealed the positive and negative interplay of cognition and affect. The field's future may exemplify Allan Baddeley's⁷⁷ call for the integrated pursuit of *applied basic* research, testing theory by its application, and *basic applied* research, creating theory from new phenomena observed through those tests.

Notes

- 1 Coombs CH, Dawes R, Tversky A. 1970. *Mathematical Psychology*. Englewood Cliffs, NJ: Prentice Hall.
- 2 Edwards W. 1954. A theory of decision making. *Psychol Bull* 54:380–397.
- 3 Berkeley D, Humphreys PC. 1982. Structuring decision problems and the “bias heuristic.” *Acta Psychol* 5:201–252.
- 4 Hammond KR. 1966. *The Psychology of Egon Brunswik*. New York: Holt, Rinehart and Winston.
- 5 Fischhoff B, Beyth-Marom R. 1983. Hypothesis evaluation from a Bayesian perspective. *Psychol Rev* 90:239–260.
- 6 Goldberg LR. 1968. Simple models or simple processes? Some research on clinical judgments. *Am Psychol* 23:483–496.
- 7 Kelly EL, Fiske DW. 1951. *The Prediction of Performance in Clinical Psychology*. Ann Arbor: University of Michigan Press.
- 8 Meehl PE. 1954. *Clinical Versus Statistical Prediction: A Theoretical Analysis and A Review of the Evidence*. Minneapolis: University of Minnesota Press.
- 9 Dawes RM. 1979. The robust beauty of improper linear models in decision making. *Am Psychol* 34:571–582.
- 10 Dawes RM, Faust D, Meehl P. 1989. Clinical versus actuarial judgment. *Science* 243:1668–1674.
- 11 Ericsson A, Simon HA. 1994. *Verbal Reports as Data*. Cambridge, MA: MIT Press.
- 12 Nisbett RE, Wilson TD. 1977. Telling more than we know: Verbal reports on mental processes. *Psychol Rev* 84:231–259.
- 13 Camerer CF, Hogarth RM. 1999. The effects of financial incentives in experiments: A review and capital-labor–production framework. *J Risk Uncertainty* 19:7–42.
- 14 Becker MH. 1974. The health belief model and personal health behavior. *Health Educ Monogr* 2(4):324–473.
- 15 Hochbaum G. 1956. Why people seek diagnostic x-rays. *Publ Health Rep* 71:377–380.
- 16 Ogden J. 2003. Some problems with social cognition models: A pragmatic and conceptual analysis. *Health Psychol* 22:424–428.
- 17 Fischhoff B. 2005. Cognitive processes in stated preference methods. In: Mäler K-G and Vincent J, eds. *Handbook of Environmental Economics*. Amsterdam: Elsevier 937–968.
- 18 Viscusi WK. 1983. *Risk by Choice* Cambridge, MA: Harvard University Press.
- 19 Slovic P, Lichtenstein S. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organ Behav Hum Perform* 6:649–744.
- 20 von Neumann J, Morgenstern O. 1944. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
- 21 von Winterfeldt D, Edwards W. 1986. *Decision Analysis and Behavioral Research*. New York: Cambridge University Press.
- 22 Clemen RT. 2003. *Making Hard Decisions: An Introduction to Decision Analysis*. Belmont, CA: Duxbury.
- 23 Peterson CR, Beach LR. 1967. Man as an intuitive statistician. *Psychol Bull* 68:29–46.
- 24 Jonides J, Naveh-Benjamin M. 1987. Estimating frequency of occurrence. *J Exp Psychol Hum Learn Mem* 13(2):230–240.
- 25 Hasher L, Zacks RT. 1981. Automatic processing of fundamental information. *Am Psychol* 39:1372–1386.
- 26 Tversky A, Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–1131.
- 27 Gilovich T, Griffin D, Kahneman D, eds. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press.

- 28 Kahneman D, Slovic P, Tversky A, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- 29 Edwards W, Lindman HR, Savage LJ. 1963. Bayesian statistical inference for psychological research. *Psychol Rev* 70:193–242.
- 30 Phillips LD. 1973. *Bayesian Statistics for Social Science* London: Nelson.
- 31 Fischhoff B. 2009. Risk perception and communication. In: Detels R, Beaglehole R, Lansang MA, Gulliford M, eds. *Oxford Textbook of Public Health*. 5th ed. Oxford: Oxford University Press; 940–952.
- 32 Poulton EC. 1968. The new psychophysics: Six models for magnitude estimation. *Psychol Bull* 69:1–19.
- 33 Poulton EC. 1994. *Behavioral Decision Making* Hillsdale, NJ: Lawrence Erlbaum.
- 34 Stevens SS. 1975. *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*. New York, NY: John Wiley & Sons.
- 35 Tune GS. 1964. Response preferences: A review of some relevant literature. *Psychol Bull* 61:286–302.
- 36 Lichtenstein S, Slovic P, Fischhoff B, Layman M, Combs B. 1978. Judged frequency of lethal events. *J Exp Psychol Hum Learn Mem* 4:551–578.
- 37 Ajzen I, Fishbein M. 1980. *Understanding Attitudes and Predicting Social Behavior* Englewood Cliffs, NJ: Prentice-Hall.
- 38 Gilbert DT. 2006. *Stumbling on Happiness*. New York: Knopf.
- 39 Fischhoff B. 1991. Value elicitation: Is there anything in there? *Am Psychol* 46:835–847.
- 40 Lichtenstein S, Slovic P, eds. 2006. *Construction of Preferences*. New York: Cambridge University Press.
- 41 Lupia A. 1994. Shortcuts versus encyclopedias—information and voting behavior in California insurance reform elections. *Am Polit Sci Rev* 88:63–76.
- 42 Schläpfer F. 2008. Contingent valuation: a new perspective. *Ecol Econ* 64:729–740.
- 43 Politi MC, Han PKJ, Col N. 2007. Communicating the uncertainty of harms and benefits of medical procedures. *Med Decis Making* 27:681–695.
- 44 Schwartz LM, Woloshin S, Welch HCG. 2009. Using a drug facts box to communicate drug benefits and harms. *Annals of Internal Medicine* 150:516–527.
- 45 McGuire W. 1969. Suspiciousness of experimenter's intent. In: Rosenthal R, Rosnow RL, eds. *Artifact in Behavioral Research*. New York: Academic Press.
- 46 Fischhoff B, Slovic P, Lichtenstein S. 1980. Knowing what you want: Measuring labile values. In: Wallsten T, ed. *Cognitive Processes in Choice and Decision Behavior*. Hillsdale, NJ: Erlbaum; 117–141.
- 47 Kahneman D, Tversky A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–281.
- 48 Helson H. 1964. *Adaptation Level Theory: An Experimental and Systematic Approach to Behavior* New York: Harper and Row.
- 49 Johnson EJ, Goldstein D. 2003. Do defaults save lives?. *Science* 302:1338–1339.
- 50 Thaler RH, Sunstein CR. 2008. *Nudge* New Haven: Yale University Press.
- 51 Fischhoff B. 1983. Predicting frames. *J Exp Psychol Learn Mem Cogn* 9:113–116.
- 52 Reyna VF, Farley F. 2006. Risk and rationality in adolescent decision making: Implications for theory, practice, and public policy. *Psychol Sci Publ Interest* 7:1–44.
- 53 Simon HA. 1957. *Models of Man*. New York: John Wiley & Sons.
- 54 Gigerenzer G, Selten R. 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.
- 55 Johnson EJ, Haubl G, Keinan A. 2007. Aspects of endowment: A Query Theory of value construction. *J Exp Psychol* 33:461–474.
- 56 Tversky A, Koehler DJ. 1994. Support theory: A nonextensional representation of subjective probability. *Psychol Rev* 101:547–567.
- 57 Weber EU, Johnson EJ. 2009. Mindful judgment and decision making. *Annu Rev Psychol* 60:53–85.