

VALUING
ENVIRONMENTAL
GOODS

AN ASSESSMENT OF THE
CONTINGENT VALUATION
METHOD

R.G. Cummings
D.S. Brookshire
W.D. Schulze

Contributors

Richard Bishop, Don L. Coursey,
A. Myrick Freeman, Thomas A. Heberlein,
Alan Randall, V. Kerry Smith

Commentators

Kenneth Arrow, Daniel Kahneman,
Sherwin Rosen, Vernon Smith

ROWMAN & ALLANHELD
PUBLISHERS

Contents

| | |
|--|----------|
| <i>Tables and Figures</i> | ix |
| <i>Preface and Acknowledgments</i> | xi |
| Part I The Contingent Valuation Method | 1 |
| 1 <i>Introduction</i> | 3 |
| The Contingent Valuation Method 3; The CVM Has Achieved Acceptability 4; But on the Other Hand 5; The Need for a State of the Arts Assessments of the CVM 6; Notes 9 | |
| 2 <i>A Historical Perspective for the CVM Assessment</i> | 10 |
| Overview 10; Social Welfare: What Is It and How Is It Measured? 11; Benefit-Cost Analysis and the General Possibilities Theorem 11; Market Prices as Measures of Social Welfare 12; Valuing Public Goods 13; Developing the CVM 15; Related Research in Other Disciplines 17; The Structure for a CVM Assessment 19; Notes 20 | |
| 3 <i>Applications of the CVM: An Overview of Issues</i> | 21 |
| Overview 21; Strategic Bias and The CVM 21; Axiomatic Behavior and CVM Valuations 27; Bias Issues Related to the Design of CVM Questionnaires 28; Starting Point Bias 28; Vehicle Bias 31; Information Bias 33; Willingness to Pay vs. Willingness to Accept 35; Concluding Remarks 36 | |
| 4 <i>Experimental Economics: Implications for the CVM</i> | 37 |
| Introduction 37; Methodological Development in Experimental Economics 38; Recent Applications of Laboratory Methods Related to CVM Developments 41; Valuations Under Uncertainty Conditions: Relevant Results from Laboratory Experiments 44; An Example: Revelation of Compensating Income Variation 45; Solicited Compensating Variations 46; Tatonnement Version of the Second-Price Auction 46; Conclusions 48; Notes 48 | |
| 5 <i>Imputing Actual Behavior From Choices Made Under Hypothetical Circumstances</i> | 49 |
| The Issues 49; Hypothetical Payment: An Incentive for Accuracy? 50; Hypothetical Biases Related to Time 50; | |

ROWMAN & ALLANHELD

Published in the United States of America in 1986
by Rowman & Allanheld, Publishers
(a division of Littlefield, Adams & Company)
81 Adams Drive, Totowa, New Jersey 07512

Copyright © 1986 by Rowman & Allanheld

All rights reserved. No part of this publication
may be reproduced, stored in a retrieval system, or
transmitted in any form or by any means, electronic,
mechanical, photocopying, recording, or otherwise,
without the prior permission of the publisher.

Library of Congress Cataloging in Publication Data

Cummings, Ronald G.
Valuing Environmental Goods: An Assessment of the
Contingent Valuation Method

Bibliography: p. 247

Includes index.

I. Public goods—Cost effectiveness. I. Brookshire,
David S. II. Schulze, William D. III. Bishop, Richard C.
IV. Arrow, Kenneth Joseph, 1921— V. Title.

HB846.5.C86 1986 363 85-14298
ISBN 0-8476-7448-7

86 87 88 / 10 9 8 7 6 5 4 3 2 1
Printed in the United States of America

research community, more and more scholars have entered the debate as to the efficacy of the CVM, in real and potential terms, as a means for valuing public goods. At this point in time, a substantial literature has developed concerning the issue, in the most general terms, as to whether one can hope to derive meaningful measures of individual values from a method wherein all aspects "relevant" to value decisions are artificial or hypothetical. A brief overview of this literature will provide the reader with some flavor for this controversy and, therefore, with an appreciation for the major objectives of this book—a topic that will be discussed below. Thus, in the following two sections we consider arguments related to the proposition: "The CVM has achieved acceptability...., but on the other hand."

THE CVM HAS ACHIEVED ACCEPTABILITY...

Randall and others argue that research to date has established the acceptability of the CVM as a method for non-market benefit estimation and that the current task "is to identify and explain systematically the relationship between the structure and performance of contingent markets." (Randall et al. 1983, p. 642) Thus, Randall et al. assert:

At the outset, the research agenda in contingent valuation sought to establish, in the face of considerable skepticism, contingent valuation as an acceptable method of non-market benefit estimation (acceptable in the sense that it works about as well as available alternative techniques and is adaptable to at least some valuation tasks that alternative methods cannot handle). That objective has been attained. In addition, the experimental work of others has blunted traditional fears that strategic responses would inevitably dominate data sets of stated personal valuations.

Other authors, despite their critique of some CVM studies, suggest cautious optimism for the promise of the CVM. For example:

CVM studies are a promising approach for the estimation of non-market environmental values. There has been steady progress in minimizing biases, just as there has been progress with problems in other techniques; nevertheless, we are far from being out of the woods. (Rowe and Chestnut 1983, p. 408)

Since the relatively recent beginning of empirical experiments with the CVM,¹ progress of sorts has undeniably been made in the development of the CVM. As pointed out by Randall et al. (1983), bids obtained in CVM studies are generally shown to be significantly related to income, availability of substitute and complementary commodities, and demographic characteristics; i.e., CVM bids "not random numbers." (p. 639-40) Bids have been shown to be consistent with actual behavior. (Randall et al. pp. 1983, 639-40) As is discussed in some detail in a later chapter of this book, maximum willingness to pay measures derived from CVM studies have been shown to be consistent with market-demand-based values. Within this context, a basis exists for Randall et al.'s assertion (p. 639) that "several kinds of evidence generated by CVM studies support contingent valuation methods." Moreover, in a recent study by Schulze et al. (1981c), selected CVM studies were reviewed to the end of assessing the extent of various biases in CVM measures. The authors conclude that "Biases do

not appear to be an overriding problem" (p. 170), although they point out that "to establish a precise contingent market—the 'good' must be well-defined." (p. 170)²

BUT ON THE OTHER HAND

Notwithstanding the "progress" noted above, others within the economics profession, and many outside the profession, reject the above described notion that the CVM has attained anything near the level of "acceptability" ascribed to the method. In reviewing estimation methods, including the CVM, for valuing non-market goods, Feenberg and Mills (1980) offer the dreary conclusion that "In the absence of market data, demand or willingness to pay estimation would appear to be hopeless." (p. 58)³ Referring specifically to survey methods such as the CVM, Feenberg and Mills seemingly presume to speak for the economics profession in offering the following conclusion:

Economists are biased against such surveys because they believe crucial contrary-to-fact questions are unlikely to be answered accurately. People lack the incentive and ability to answer accurately questions such as, 'How much more often would you swim in lake L if ambient pollution concentrations were reduced 10%?' Most people presumably experiment and talk to others to ascertain the effect of pollution abatement on their utility-maximizing behavior. Thus, economists doubt the accuracy of survey responses regarding effects of pollution abatement. (p. 169)

Interestingly enough, the "incentives" criticism of measures drawn from the CVM, as couched above, is inextricably related to a second criticism of the CVM, that is, biases resulting from strategic behavior on the part of survey participants. Essentially, the strategic behavior hypothesis—discussed in detail below in chapter 2 – posits behavior by survey respondents whereby false responses are given when such responses may result in a gain to the individual; i.e., "it is in the selfish interest of each person to give false signals, to pretend to have less interest in a given collective consumption activity than he really has." (Samuelson 1954, p. 389) From empirical efforts to test the strategic behavior hypothesis, it is shown that the more hypothetical the question in a survey, the less the incentive for strategic behavior—the use of hypothetical questions could be a means of avoiding biases from strategic behavior. (Freeman 1979a, pp. 97-99) Herein lies the potential dilemma: the more hypothetical the question, the less the incentives for strategic behavior but, also, the less the incentives for accurate responses.

In addition to the above, two related sets of considerations that pose questions as to the efficacy of the CVM emanated from outside of the profession per se, that is, from the branch of psychology referred to as "cognitive psychology". The first of these (noted above) questions the extent to which responses derived in CVM studies are expressions of attitudes as opposed to intended behavior (as is presupposed in CVM studies) and a related controversy in the discipline of psychology concerning the extent to which attitudes are reliable predictors of behavior (Bishop and Heberlein 1979). A second set of considerations received from psychology which is of potential relevance for the CVM strikes at one of the most basic concepts in economic analysis: the concept of rational behavior. A number of recent studies point to stark discrepancies between actual

decision-making behavior and the postulates of rationality, particularly in circumstances involving uncertainty.⁴ Arrow (1982) notes that "these failures of the rationality hypothesis are in fact compatible with some of the specific observations of cognitive psychologists." (p.5) The "observations" referred to by Arrow will receive considerable attention in later sections of this book. For present purposes, two of these observations from psychological research are germane. In direct contrast to expected utility theory wherein subjective probabilities based on prior information play a major role, cognitive psychologists argue that individuals, in evaluating uncertain events, tend to ignore *both* prior information and the quality of present evidence. (Tversky and Kahneman, 1974, 1981) Second, also in direct contrast with the rationality precepts underlying expected utility theory, cognitive psychologists essentially argue that an individual's valuation of a commodity, along with many other commodities, is not simply dependent on the commodity set (prices, income and commodities), but on *how the set is described*—different descriptions of the *same* commodity space may yield different values for specific commodities. (Tversky and Kahneman, 1981)

Implications of these observations for potential biases in results from CVM studies are obviously a matter of some concern. For example the first issue—excessive reaction to current information—may imply that obtained CVM values are susceptible to the influence of (often) temporary "media events"; in terms of efforts to value environmental quality, the Three-Mile Island incident and the furor over Love Canal—a popular media topic in 1980—come to mind. Moreover, the applicability of CVM values obtained in one "current information" climate to values relevant for a different climate is questionable. The second issue—the dependence of commodity values on how commodities are described—implies potential biases arising from the framing of willingness-to-pay questions in the CVM questionnaire; thus, for any given public/environmental commodity to be valued via the CVM, different descriptions of the same basic commodity could yield different estimates of values of the commodity.⁵

THE NEED FOR A STATE OF THE ARTS ASSESSMENT OF THE CVM

It is important that the reader understand the context for which the controversy described above is relevant. President Reagan's Executive Order 12,291 (46 Fed. Reg. 13, 193, Feb. 17, 1981) requires that federal agencies such as the EPA consider the benefits and costs of federal regulations/actions prior to their implementation. For EPA regulations, such as air and/or water quality standards and regulations on hazardous waste disposal practices, costs may be amenable to estimation, but benefits attributable to a large part of these regulations are non-market, "public goods" in nature: cleaner air and water, a safer environment. Agencies such as the EPA then have strong incentives and interests in identifying and developing means by which benefits attributable to public goods—such as environmental improvements—may be assessed.

Methods other than the CVM exist for valuing public goods, primarily the Travel Cost Method (TCM)⁶ and the Hedonic Price Method (HPM)⁷ The environmental (and other public good) "commodities" for which the TCM or HPM might be used for valuation purposes are very limited, however.⁸ For the broad range of air quality and environmental safety issues of potential

regulatory concern to the EPA, the CVM is, metaphorically, the only game in town for estimating relevant benefits. Obviously, the fact that the CVM is no worse than other methods or is the only game in town is not a sufficient reason for the use of CVM values as "acceptable" economic measures of social benefits in policy assessments. However, one sees rationales like these suggested as justifications for the continued development of the method. For example, Burness et al. (1983) conclude their discussion of caveats relevant for reported CVM results with the observation:

Continued interest and research in this (the CVM) area are clearly warranted given, first, the importance of the public goods issue and, second, the lack of apparent alternatives to some form of the survey method in deriving valuations for large classes of public (environmental) goods. (p. 682)

On the other hand, the fact that the CVM is "the only game in town" for providing information of relevance to critical policy issues of the day is a powerful incentive for scholars to meet the intellectual challenge to devise means by which the CVM (or other methods) can be effectual in responding to society's needs.

Within this milieu (chapter 2 traces the character of historical efforts to develop the CVM), it seems fair to say that all scholars—whatever their predilection towards the CVM—who are directly or indirectly involved with the method appreciate the immediate need for a reflective pause in CVM experiment/application activities. Such a pause is required for thinking through the many (again, intuitive) propositions that have been posed as indicative of sources for bias in CVM measures, as well as related (again, often intuitive) counter-arguments. Most importantly, a reflective pause is required for a re-examination of means by which we can effectively apply the scientific method in our efforts to assess the CVM. In this regard, Joan Robinson's (1962) polemic concerning the difficulty in social sciences of applying the scientific method, is relevant for our discussions:

[Referring to why economics is a branch of theology] "... the process of science ... consists in trying to *disprove* theories ... The great difficulty in social sciences ... of applying scientific method, is that we have not yet established an agreed standard for the disproof of an hypothesis (pp. 22-3) (theories become religions in the social sciences because) first, the subject matter has much greater political and ideological content, so that other loyalties are ... involved ... (and secondly) it has been sometimes remarked that economists are more queazy and ill-natured than other scientists. The reason is that, when a writer's personal judgement is involved in an argument, disagreement is insulting." (pp. 23-24)

As will be seen in later discussions, it is not rare to find one writer questioning the judgement of other writers in the CVM literature and there exists considerable disagreement, if not confusion, as to standards for proving or disproving hypotheses relevant for important aspects of the method. Thus, developments with the CVM have reached an important watershed at which a state of the arts assessment of the method is timely. The purpose of this book is to provide such an assessment.

The critical assessment of the literature relevant for the CVM is the substance of the remaining five chapters in part I of this book. Given that the intent of this literature review is to go beyond a simple description of literature to an *assessment* of the strengths and weaknesses of the CVM, we begin in chapter 2 with the development of an historical setting for the CVM within which an assessment framework for evaluating the state of the arts of the method is promulgated. Arguments developed in chapter 2 will set the stage for the central thrust of remaining chapters in part I.

The arguments developed by the authors in these five chapters are intended to serve as a point of departure for a critical examination of the state of the arts for the CVM. Obviously, the author's assessment of the CVM is in no way "the profession's" assessment and, as noted above, what is needed at this point in time is a profession-wide evaluation of the CVM. An effort to obtain something akin to a broader, profession-wide assessment is accomplished via an Assessment Conference, which has the following form.

"A Conference on Valuing Environmental Improvements: A STATE OF THE ARTS ASSESSMENT OF THE CVM" was held in Palo Alto, California, on July 2, 1984. The purpose of the Conference was to elicit a review panel's judgments as to the promise of the CVM as a means for valuing public/environmental goods. The panel consisted of leading scholars in the economics profession and included:

Kenneth Arrow, Stanford University
Daniel Kahneman, University of British Columbia
Sherwin Rosen, University of Chicago
Vernon Smith, University of Arizona

The review panel's consideration of the CVM was based, in addition to their general knowledge and expertise in the science of public goods valuation, upon two sets of information. The first information set was the author's critical assessment of the CVM as set out in part I of this book; part I was made available to Panel members well in advance of the Conference. The second information set was papers and presentations provided by four leading scholars involved in research related to the CVM. Paper/presentations by these scholars focused first on the critical assessment of part I of this book and secondly on their individual assessments of the promise, strengths, and weaknesses of the CVM. The four scholars offering presentations at the conference were:

Richard Bishop, University of Wisconsin
A. Myrick Freeman, Bowdoin College
Alan Randall, University of Kentucky
V. Kerry Smith, Vanderbilt University

Results from the conference are reported in part II of this book. The authors' assessment of the CVM—the substance of part I—and a more general, profession-wide assessment of the CVM—part II of the book—allow us to conclude with what the authors hope will be regarded as an objective, benchmark evaluation of the CVM. Drawing from the diverse sources described above, in chapter 13 the authors will offer final conclusions as to the current state of the arts for the CVM.

NOTES

1. As examples, see Davis (1963a) and Bohn (1971).
2. These conclusions are challenged, however, in Rowe (1983).
3. As part of the authors' context for the cited conclusion, the authors also assert that "almost no empirical work has been based on careful theoretical analysis" (p.58). Excepting the use of surveys, this conclusion is softened somewhat in their Chapter 10 however.
4. As examples, see S. Lichtenstein and P. Slovic (1971); D. Grether and C. Plott (1979); Kunreuther et al. (1978); H. Simon (1979).
5. For related discussions, see M.C. Weinstein and R.J. Quinn, (1983). Furthermore, it may be tempting to set this source of bias aside as one which can be readily eliminated through questionnaires with alternative question frames. A careful consideration of the example given in Arrow (1982, p. 7) belies the ease by which this problem may be mitigated by questionnaire design or administration.
6. See R. Mendelsohn and G.M Brown, Jr. (1983).
7. See S. Rosen (1974).
8. See Freeman (1979a), Chapters 4-6; particularly pp. 85-87.

development of the method, and of the major issues which require resolution for further developments. The need for such a pause was made manifest by our review of the myriad "criticisms" of the CVM, all of which pointed to the disarray and confusion amongst CVM researchers attributable to two central facts. First, there has been a lack of consensus among researchers as to the priority issues and hypotheses that warrant empirical focus. Research efforts appeared scattered and diffuse as we repeatedly addressed *asserted* "biases" in the CVM (e.g., starting point, information, vehicle biases, etc.) in the "heuristic" manner described in chapter 3, with seemingly but one basis for accepting or rejecting a "bias": some ill-defined "preponderance of evidence". In large part, this lack of a well-defined, prioritized research agenda for the CVM reflects the *ad hoc*, "chemistry set" approach to CVM research noted by V. Kerry Smith, Bishop-Heberlein, and other Conference participants. Empirical applications of the CVM have outstripped intellectual inquiry—via formal models or otherwise—as to how individuals may behave within contingent market settings and implications for questionnaire design and implementation practices. Secondly, following perhaps from the preceding observations, CVM researchers have been apologetic, or defensive, vis-à-vis the "rest of the profession" due to the pervasive feeling that interrogated responses by individuals to hypothetical propositions must be, at best, inferior to "hard" market data or, at worst, off-the-cuff attitudinal indications which might also be expected to reflect efforts by individuals to manipulate the survey to their selfish ends.

The difficulties involved in efforts to provide some state of the arts context for the controversies surrounding the viability of the CVM for estimating values for public goods are made manifest by the assessment *process* seen in parts I and II of this book. Thus, many of the positions and conclusions presented by us in part I were altered or retracted in this chapter as a result of the focused dialogue concerning priority issues in CVM research between the authors, four other prominent CVM researchers, and leading economics and psychology scholars whose interest in public goods valuation is a step removed from the CVM *per se*. Of course, the reader will judge the success of this process in providing a start of the arts context for the CVM. In this regard, *our* general view of this context is described as follows:

The CVM Without Apology

It is surely time for replacing apologies for the CVM with a positive research agenda to be described below. As a first step in this direction, we must eschew the joys of self-flagellation over our lack of knowledge of the "truth": we don't and won't know it, nor will our colleagues in the "rest of the profession" vis-à-vis *their* value estimates, nor will scientists in other disciplines. Following Arrow's exhortations, we must directly address the question, what *is* accuracy, and then look to calibration methods which provide us with a means to achieve accuracy levels that are reasonable and consistent with those levels obtained in other areas of economics and in other disciplines.

What is accuracy in a CVM estimate? It is a subject's valuation of a commodity which "reasonably" reveals his/her preference for the commodity. What does "reasonable" mean? "Reasonableness" is established by criteria—Reference Operating Conditions—which allow us to measure the magnitude of

probable errors in any given application of the CVM. Thus, whether resulting ranges for Reference Accuracy associated with applications of the CVM are never better than $\pm 50\%$ or $\pm 500\%$, our focus is on defining the reference accuracy range. As with any other estimates, the "usefulness" of estimates with any range of error is determined by the purposes to which the estimates are to be put.

Conclusions Concerning Accuracy

While perhaps useful in pointing to needed research, the above is little more than a definition of accuracy. Given, as was argued above, that efforts to develop ROCs for the CVM have just begun, and that we are almost totally ignorant of the error implications associated with the few ROC's that seem palatable at this time, must we then agree with V. Kerry Smith's judgement (chapter 11) that no conclusions about the accuracy of CVM measures can be drawn based on research accomplished to date? We think not. At this point of reflective pause in the development of the CVM, one fails to see implications for the accuracy of CVM measures from received research *only* if one's view of "acceptable" implications is limited to evidence that demonstrates some degree of precision—narrow ranges of error. This is to say that while we cannot build the case for ranges of Reference Accuracy for the CVM of magnitudes that would make CVM value estimates of practical use in many cases, at this point in the method's development a "useful conclusion" in the sense of V. Kerry Smith's assertion might well be that the method produces *order of magnitude* estimates—but we think one can argue that error ranges are much smaller.

Before continuing this argument, it is relevant for our purposes to recall V. Kerry Smith's demonstration (chapter 11) of the wide range of potential for hypothetical and reporting errors in "hard" data commonly used, without apology, in economic analyses. Such data are seemingly accepted in total ignorance of ROC's relevant for their collection and the resulting ranges of Reference Accuracy. This observation, when combined with Coursey and Nyquist's findings of potential errors in ordinary demand analysis and Mitchell-Carson's general comments regarding sampling errors, should serve—to paraphrase Freeman (chapter 10)—as a chilling reminder of the limitations of empirical analysis/models in most areas of economic analysis. It seems fair to say that, in the general economics literature, questions of accuracy are not prominent. This is not to suggest a nihilistic approach to CVM research: the whole world is wrong (inaccurate), so why should we be concerned with accuracy. We mean to suggest the perspective: economists' typical preoccupation with such things as standard errors, etc., may have misled us into viewing value estimates as "precise" in terms of narrow error ranges, $\pm 5\%$, 10% or even 20% . Couched in the broader terms of Reference Accuracy, such "precision" in general economic value estimates may quickly dissipate. Again, that such broader views of accuracy are generally ignored in economics is made manifest by V. Kerry Smith's provocative discussion in chapter 11.

Returning to our discussion of what one *can* conclude regarding the accuracy of CVM measures, we begin by recalling an earlier discussion of the "truth". We do not and will not know it. But something analogous to "truth" may be attributed to values derived from, as examples, *actual* cash trades in Bishop-Heberlein's Sandhill study and in Vernon Smith's laboratory experiments.

Eschewing arguments as to how Bishop-Heberlein's auction formats might have been improved in one way or another, their cash offers/payments are certainly the "truth" vis-à-vis preference revelation in the sense that folks clearly paid (were paid) for a well-defined commodity and then used the commodity. For the limited, most likely nonequilibrium, "simulated" market used by them, we can surely attribute preference revelations to these values. The differences between mean cash and CVM-WTA values was roughly 42%; between cash and CVM-WTP values, differences ranged from about 38% to 124% across their four auction formats (tables 9.2 and 9.30). Do these differences imply nothing vis-à-vis conclusions as to the accuracy of CVM measures? If accuracy is viewed as involving "small" ranges for Reference Accuracy, one would lament the "large" differences, as do Bishop and Heberlein, and concur with V. Kerry Smith that nothing (positive) can be concluded. If orders of magnitude are relevant, one might find Bishop-Heberlein's results startling: CVM and cash offers are virtually the same (see figure 6.1). Our colleagues in environmental engineering may well envy such accuracy. In these regards, we note Bishop-Heberlein's later "surprise" at how well the CVM does work—cash-CVM differences were not "outrageous".

Questions of the transferability of laboratory results to real-world conditions aside, hypothetical responses in Vernon Smith's experiments were consistently within 10% of actual market outcomes. In the Coursey et al. (1983b) laboratory experiment (figure 4.1), differences between values derived from final Vickrey auctions and hypothetical questions were less than 20% for WTP and approximately 100% for WTA. The central point in all of this is apparent, however. In terms of the standard for comparisons of CVM values, we can continually argue as to how well preference revelations are made manifest by Bishop-Heberlein's cash offers, Vernon Smith's securities values, Coursey et al.'s measures related to tasting sucrose octa-acetate, or, moving to public goods, TCM and HPM values derived by the eight sets of authors given in table 6.12. But however well *any* of these measures reflect meaningful revelations of preferences by individuals, *every* piece of evidence that we have demonstrates that the CVM yields value estimates that are indistinguishable from those standards in order of magnitude terms. Indeed, and herein lies the relevance of our $\pm 50\%$ arguments, in *most* instances CVM values are within $\pm 50\%$ of values derived from alternative methods for estimating preference revealed values.

Final Remarks

Thus, our *final* (c.f. our stronger, pre-conference, reservations in chapter 1–6, *ad passim*) assessment of the state of the arts of the CVM is generally positive. We find impressive the accuracy of CVM measures inferred by the available evidence at this stage of the method's development. We find encouragement in the Conference results, particularly those reported by Arrow, Kahneman, and Bishop-Heberlein, which suggest that breaking the "hypothetical barrier" in the CVM may not be as hopeless as we and others earlier believed.

"Promise" is not "performance", however, and our assessments given above refer only to the potential promise of the CVM as a viable method for estimating values for public goods. The realization of that promise implies *real* challenges for theoretical and empirical research for those involved with the method's

further development. In concluding this book, we now focus attention on critical issues for any research agenda which are relevant for guiding future CVM research.

CRITICAL ISSUES FOR FUTURE CVM RESEARCH

In most general terms, it must be hoped that greater focus can be achieved in future research with the CVM. Both Bishop-Heberlein and V. Kerry Smith emphasize the *ad hoc* character of the bulk of CVM research to date—a characterization aptly described by Bishop-Heberlein as reflecting a "chemistry set" approach. To a large extent, the *ad hoc* quality of CVM research has resulted from the emphasis or priority given empirical results—necessitated in many cases by data needs of the entities providing research funding—as opposed to theoretical and design issues. Results from this empirical emphasis are made manifest by the profession's preoccupation, without resolution, with such operational "biases" as starting point, information and vehicle issues as noted in chapter 3, issues some of which, upon reflection by Conference participants, may now be viewed as *not* implying biases *per se* but rather implying limits on questionnaire design and the manner in which CVM values are interpreted. Thus, the first critical issue for future CVM research is the metaphorical realignment of the empirical cart and the theoretical horse. There is a critical need for modeling efforts focused on individual behavior in contingent market settings which may serve as a basis for formulating hypotheses for empirical testing. This need for modeling efforts underlies virtually all of the additional issues for further CVM research discussed below.

A second critical issue for future research involves the specification and measurement of Reference Accuracy for CVM measures. In this regard, imaginative and innovative thought is required for defining relevant ROCs (e.g. table 13.1 above) and for calibrating errors with deviations from ROCs. Thus, we must ask questions exemplified by: What is "familiarity" or "experience" vis-à-vis a CVM commodity; what is "uncertainty" and what constitutes "ideological content"; what variables may perform best as measures of cognition and/or affectation and how are attitudinal variables calibrated with measures of attitude-behavior correspondence; how can we better structure value questions so as to enhance *a priori* our expectations that preference are obtained which are at least consonant with incentive-compatible revelations in market context? In addressing these issues we will need to profit from and exploit the lessons learned in laboratory and field experiments, as well as in research in other disciplines.

A final critical issue for future CVM research involves our need to resolve the "incentives" question. In this regard, our concern extends beyond the hypothetical payment question. We concur with Arrow's suggestion that question settings that are sufficiently pseudo-real may be expected to result in satisfactorily pseudo-real responses and we are not convinced as to the extent to which one can distinguish between payment effects and those attributed to familiarity and experience questions. Of interest in these regards is the threads of an argument, seem implicitly in Randall's paper, as well as in Kahneman's Comments, and explicitly in Mitchell-Carson's paper (appendix), that valuations of contingent changes in provision levels of public goods might be better obtained via processes which attempt to simulate results from institutions