



Measuring Constructed Preferences: Towards a Building Code

JOHN W. PAYNE

JAMES R. BETTMAN

Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708-0120

DAVID A. SCHKADE

Department of Management, University of Texas, Austin, TX 78712-1174

Abstract

A “building code” for preference measurement is needed in a world in which many expressions of preference are constructed when people are asked a valuation question. Construction of preferences means that preference measurement is best viewed as architecture (building a set of values) rather than as archaeology (uncovering existing values). We describe potential faults in the process of preference construction, offer guidelines for measuring constructed preferences (a “building code”) to mitigate these faults, and discuss how the code must be sensitive to the purpose of the valuation (design vs. prediction).

Key words: constructive preferences, value measurement, decision aiding

JEL Classification: D80

Introduction

Imagine that you are a member of a management team at the Environmental Protection Agency (EPA) charged with allocating scarce resources to alternative environmental management interventions. As part of your efforts, you need to establish the value of potential natural resource benefits achieved from different proposed interventions. To help in that valuation exercise you seek information about public values for different levels of these natural resources; obtaining citizen inputs is a natural extension of the principles of democratic governance (see Gregory et al. (1997) for a discussion of this position). One method that you consider for getting such information about public values is to conduct a contingent valuation (CV) study of peoples’ preferences for different levels of an environmental resource (for a general review of CV studies, see Mitchell and Carson, 1989). Those expressions of preferences will be used as inputs to *design* and select among alternative environmental management interventions.

Alternatively, imagine that you are the marketing manager for a new product and have been asked to generate a *prediction* regarding the likely sales of that product. The need to make such predictions is becoming both more important and more difficult because of rapid changes in technology and shortened product lives. You consider asking a sample of consumers to express their preference for your product in comparison to other competitor products that are, or might be, in the marketplace. Those expressions of preference will be used to predict future sales of your product.

How the EPA administrator or market manager approaches the measurement of preferences depends on the assumptions made about those preferences. One common viewpoint is that well-defined preferences exist for most objects, and therefore the task is to uncover (reveal) those preferences. Gregory, Lichtenstein, and Slovic (1993) have likened this task to an archaeological project, i.e., uncovering values that may be hidden but are assumed to exist. An alternative viewpoint is that preferences are generally constructed—not revealed—at the time a valuation question is asked. The measurement of constructive preferences is then more like a form of architecture, i.e., building a sturdy and “defensible” value or set of values, than archaeology (Gregory et al., 1993).¹ We argue below that a constructive view of preference expression requires a fundamentally different approach to the measurement of preferences than that implied by the view of preference expression as revealing well-defined values. That is, it matters greatly whether preference measurement is seen as a form of archaeology, uncovering values that are already there, or as a form of architecture, building some defensible set of values.

We believe that enough knowledge has been gained about decision processes to begin to develop a set of guidelines (a “building code”) for measuring constructed preferences. These building code guidelines should help avoid (or at least mitigate) some of the common faults that can occur in preference construction. In illustrating the development of such a code, we consider both the nature of the information that might be provided to the decision maker (the construction materials) and useful tools for thinking (the construction procedures). We also consider how these “building code” guidelines should vary depending upon whether the purpose of the exercise is to *guide* (*design*) future decisions (e.g., the EPA manager’s task) or is to *predict* individuals’ future responses (e.g., the marketing manager’s task).

The remainder of this paper is organized as follows. First, we briefly discuss the two perspectives on preferences. Then we consider how to define a well-constructed preference and principles for good preference construction. Next we discuss the stages of constructing a preference, outline the types of “faults” that are likely to occur in each stage, and provide guidelines for measuring constructed preferences that seek to mitigate these common faults at each stage. We initially focus on design tasks and then consider prediction situations. Finally, we discuss when our suggested remedies are most likely to be needed and provide some words of caution about developing a “building code” for measuring constructed preferences.

1. Expressions of preference

1.1. *Stable, coherent, and known preferences*

Although it is clear that economists vary in terms of the degree to which they believe well-defined preferences exist (McFadden, 1999), a common assumption in economics is that “each individual has stable and coherent preferences” (Rabin, 1998, p. 11). In addition, it is often assumed that “people know their preferences” (Freeman, 1993, p. 7), that they have the ability or skill in computation to identify (calculate) the option that maximizes received value, and that they will choose accordingly.

Under such assumptions, the task involved in a valuation exercise is to uncover (reveal) these well-defined, pre-existing preferences. However, there can be great difficulties in uncovering even such pre-existing values. For example, a key problem is to be sure that the good to be valued has been defined properly, so that the “right” preferences are being uncovered. Mitchell and Carson (1989) provide an impressive list of the potential biases that may result when the respondent in a valuation study views the good in a way that was not intended by the researcher (sometimes called the problem of “scenario misspecification”). The researcher must also attempt to design the valuation study so that the respondent is motivated to research his or her preferences and respond truthfully (i.e., not strategically misrepresent his or her preferences).

This approach to preference measurement, which assumes a decision maker with existing, well-defined preferences, has contributed greatly to the prediction and understanding of decisions. However, an increasing number of researchers believe that the assumption of well-articulated preferences is tenable only when people are familiar and experienced with the preference object, and that an alternative, constructive view of preferences is needed in most situations.²

1.2. *Constructive preferences*

Two major tenets of the constructive perspective on preferences are that 1) expressions of preference are generally constructed at the time the valuation question is asked and 2) the construction process will be shaped by the interaction between the properties of the human information processing system and the properties of the decision task (Payne, Bettman, and Johnson, 1992; Slovic, 1995), leading to highly contingent decision behavior. That is, we assume that people do not have existing well-defined values for many objects; in addition, when asked a valuation question, they will selectively use information that is part of the immediate task description, as well as information that is drawn selectively from memory, to construct a response on the spot. The constructive view also asserts that preferences are not necessarily generated by applying some invariant process such as expected utility maximization; instead, a wide variety of heuristics (strategies)

may be used in constructing a preferential response (Simon, 1955). Individuals may construct preferences because they lack the cognitive resources needed to compute and store well-defined preferences for many situations (March, 1978) or because they bring multiple, potentially conflicting processing goals to a decision problem (e.g., maximizing accuracy, minimizing effort, minimizing negative emotion, or maximizing ease of justification; see Bettman, Luce, and Payne, 1998).³

The constructive viewpoint does not necessarily mean that there is no "true" value to be measured. Expressed preferences (measured values for decision objects), in our view, generally reflect *both* a decision maker's basic values for highlighted attributes (e.g., more money is preferred to less) *and* the particular (contingent) heuristics or processing strategies used to combine information selectively in order to construct the required response to a particular situation. That is, in addition to random error, expressed preferences include two different sources of systematic variance—stable values associated with the attributes of the object being evaluated that are relatively constant across situations and a situation-specific component that is the joint effect of the task and context contingencies that are present. We believe that the situational component will often be large, perhaps much larger than either the random error or stable value components. The focus of this paper is on measuring preferences when the situational component is a major determinant of observed choice and judgment responses.

As suggested above, an important implication of the constructive nature of preferences (and evidence for such construction) is that decisions and decision processes are highly contingent upon a variety of factors characterizing decision problems. The major findings from the past two decades of behavioral decision research support that conclusion. It is beyond the scope of the present paper to review this evidence in detail; for reviews of the extensive evidence in support of the constructive view of decision behavior, see Payne, Bettman, and Johnson (1992, 1993), Bettman, Luce, and Payne (1998), Fischhoff (1991), Slovic (1995), and McFadden (1999). Instead, we will briefly summarize some of the major conclusions of that research.

First, choice among options is context (or menu) dependent; the relative value of an option depends not only on the characteristics of that option, but also upon characteristics of other options in the choice set. As noted by Tversky (1996), such context dependence indicates that "people do not maximize a precomputed preference order, but construct their choices in light of the available options" (p. 17). Second, preference among options also depends upon how the valuation question is asked; strategically equivalent methods for eliciting preferences can lead to systematically different preference orderings. Preference reversals due to response mode persist even in the face of substantial monetary incentives (see Grether and Plott, 1979). Third, choice among options depends upon how the choice set is represented (framed) or displayed, even when the representations would be regarded as equivalent by the decision maker upon reflection. As noted by Kahneman, Ritov, and Schkade (1999), "framing effects violate a condition of extensionality (Arrow, 1982), which is commonly taken for granted in economic analyses of

preference" (p. 206). Finally, the process used to make a choice depends on the complexity of the decision task; the use of simple (heuristic) decision processes increases with task complexity. Thus, people construct preferences using a variety of methods (strategies), with such usage contingent upon characteristics of the choice problem.

Clearly not all expressions of preference are constructed at the time the valuation question is asked; at times the expression of values reflects a reference to a well-defined value in memory. However, expressed preferences are likely to be constructed in a wide variety of situations in which we are interested in preference measurement. Therefore, a "building code" for measuring constructed preferences is likely to be widely applicable. We provide a more extensive discussion of the conditions leading to constructed preferences later in the paper.

In the following sections of the paper we address issues in measuring constructive preferences. Our initial focus is on preference measurement for design purposes, i.e., tasks similar to those facing the EPA administrator in our example above. Later in the paper we address assessing preferences for prediction problems such as those faced by the marketing manager in our example.

In design problems there is no obvious "correct" preference. Therefore, before we can address the measurement issues inherent in design problems, we first need to consider criteria for assessing the effectiveness of a preference construction and principles for good preference construction. Then we consider the process of preference construction, separating that process into several interrelated stages. For each stage we examine the possible "faults" in preference construction that might occur at that stage and suggest how those "faults" may be remedied or mitigated.

2. Assessing the effectiveness of preference construction

2.1. Defining a well-constructed preference

We believe that it is possible to judge the quality of a preference construction. For example, people express dismay when they learn that their preference orders vary as a function of how the question was asked (Tversky, Sattath, and Slovic, 1988). People also do not think it is a good idea to violate principles of coherent decision making such as dominance or transitivity of choices. Beyond coherence in decisions, we believe that the *processes* (methods) leading to an expression of preference can, and should, be judged in terms of quality. That is, we accept that "truth ultimately resides in the process" of the decision (Slovic, 1995, p. 369). This tradition of evaluating human reasoning according to standards applied to the reasoning itself rather than to its conclusions can be traced back at least as far as Socrates, Aristotle, and Plato. It also has a long tradition in psychology (see Baron (1988) for further discussion). Recently, Yankelovich (1991) has also argued that

mindfulness of the consequences of an opinion should be a key component in judging the quality of an expressed opinion.

Sunstein (1990) makes a related argument about the importance of judging the quality of decision processes in the case of public decisions. He argues that not all preferences should be treated equally in a democracy and that more weight should be given to preferences that are reached "with a full and vivid awareness of available opportunities, with reference to all relevant information, and without illegitimate or excessive constraints on the process of preference formation" (p. 11).⁴

Obviously, quality must be defined in terms of the goals of a preference measurement exercise, e.g., measuring values as inputs to a public environmental protection decision or measuring values as a prediction of future market shares. In the latter case, the distinction between a better or more poorly constructed expression of preference may not matter as long as the measured preference corresponds to actual preferential behaviors such as buying a product. When measuring preferences for the purpose of designing better public or private decisions, however, we think that the focus should be on the processes leading to a well-constructed expression of preference. In part, this reflects the fact that the processes of preference construction are to some extent, at least, controllable *ex ante*.

What characteristics of the construction process should be used to assess effectiveness? Along with many others, we argue that a well-constructed expression of preference is based on thorough processing of information (reason and reflection) that is transparent and in proportion to the importance of the question at hand (e.g., Baron, 1988; Frisch and Clemen, 1994; Janis and Mann, 1977; Slovic, 1995; for an alternative view of the effects of increased thought, see Wilson and Schooler, 1991). Such processing should include consideration of a range of alternative courses of action, consideration of the full range of objectives to be fulfilled, thorough consideration of the information most critical to the individual, the making of tradeoffs, and careful review of responses to detect inconsistencies. Thus, we argue that an expression of preference based on highly selective use of information, avoidance of tradeoffs, and which is not endorsed upon reflection is not well-constructed. Further, we believe that well-constructed expressions of preference should be sensitive to manipulations that should affect them, given the purposes of measurement, e.g., the quantity of the good provided, and should be insensitive to manipulations that should not affect them, e.g., framing effects (Baron, 1997; Fischhoff, 1997; also see Hammond (1996) for a general discussion of how one judges the quality of a decision).⁵

2.2. Principles for good preference construction

In the following sections of the paper we suggest some specific tools, procedures, and tests that might be used to mitigate some of the faults in the construction of

preferences. In other words, we outline some of the specifications in materials (information) and procedures (construction processes) that might make up a building code for the measurement of preferences. The procedures often involve greater work in the measurement of preferences, with a focus on doing more tasks with fewer respondents. This reflects the belief that it is systematic effects (e.g., the effects of the choice context or complexity) rather than random error that have the greatest impact on most constructed expressions of preference. The procedures also require a greater sensitivity to the psychology of preference construction in our measurement efforts, e.g., requiring of people only those types of judgments that they can do well (Norman, 1988; Kahneman, Schkade, and Sunstein, 1998).

3. Stages of preference construction, their faults, and proposed remedies

We are far from a complete theory of preferential decision making; however, there is general agreement that decision processes can usefully be viewed in terms of multiple, interacting stages (e.g., Hogarth, 1987). Generally, the stages include the initial cognitive representation of the decision problem, information acquisition and interpretation, information combination leading to an evaluation, and an expression or mapping of that valuation onto a response, e.g., a choice or willingness-to-pay amount.

Next, we briefly describe and discuss the faults and building code implications for each of these stages. Given space limitations, we only focus on what we believe are some of the most critical faults and possible remedies. For a summary of major faults and remedies at each stage, see Figure 1.

3.1. Problem representation

Our thought is controlled by mental models and frames, typically portrayed as associative networks of concepts interconnected by relationships of varying degrees of strength. Key components of a mental model for a decision (a decision frame) include the options to be considered, the states of nature to be evaluated, and the criteria used to measure the relative attractiveness of consequences. In addition, a decision frame affects both the boundaries one puts on a problem, such as whether prior costs (outcomes) are considered, and the reference points used to code a consequence as good or bad (Russo and Schoemaker, 1989).

When individuals construct mental models of a situation, they often make explicit as little information as possible in order to minimize the cognitive load. Such mental models affect what information we pay attention to and what we ignore. In particular, there is often a focusing phenomenon in which thoughts are restricted to what is explicitly represented in the mental model (Legrenzi, Girotto, and Johnson-Laird, 1993). At other times, the focal information will depend on past experiences (knowledge) that an individual brings to the problem. In either

Faults at Stages of the Preference Construction Process and Proposed Remedies

Stage	Fault	Remedies
Problem Representation	Myopic decision frames	<ul style="list-style-type: none"> • Explicitly encourage consideration of multiple options, events, and objectives <ul style="list-style-type: none"> • Expand set of options, use value ladders • Encourage consideration of multiple futures • Construct a value tree • Clarify the distinction between fundamental and proxy values
Information Acquisition and Interpretation	<p>Using an inappropriate problem representation</p> <p>Inappropriate selectivity and the focusing illusion</p> <p>Lack of comprehension of the information provided</p>	<ul style="list-style-type: none"> • Use extensive pretesting, focus groups • Use manipulation checks • Provide important information using formats which make it salient and easy to process • Use manipulation checks • Use common, anchored scale formats • Explicitly present range information
Information combination	Avoidance of tradeoffs	<ul style="list-style-type: none"> • Provide time and thinking tools such as multiattribute utility analysis or judgment policy analysis • Assess swing weights • Decompose complex judgments • Use tools to help improve attribute weighting
Expression or Mapping of Preferences	<p>Influences of scale compatibility</p> <p>Biases in scale usage</p>	<ul style="list-style-type: none"> • Triangulation (ask questions in multiple ways) • Use liability as an asset – ask for reconciliation of inconsistencies • Use explicit scale anchors, e.g., behaviorally anchored scales • Use less sensitive, more robust scales
All Stages		<ul style="list-style-type: none"> • Increased use of sensitivity analysis

Figure 1. Faults at stages of the preference construction process and proposed remedies.

case, the highlighting of some elements of a decision problem, while downplaying others, is very important in the context of a valuation exercise. For example, in assigning a value to “air quality,” many evaluative dimensions of that concept can come to mind, e.g., visibility and health effects. Which components of air quality figure most strongly in the evaluation will depend on the mental frame adopted for

the valuation exercise (Irwin et al., 1990). Finally, once a mental model has been constructed or retrieved that seems to work for the task at hand, people tend not to consider alternatives (Legrenzi et al., 1993).

As noted above, we believe that a well-constructed preference should be based upon consideration of a range of options and those objectives most critical to the individual. Two major faults related to the problem representation stage concern failures to meet these standards. First, decision makers often focus on too narrow a range of options or objectives (often only one); i.e., they adopt a myopic decision frame (problem representation). Second, decision makers may use an inappropriate representation, i.e., they may try to solve the wrong problem using options or objectives that would not be most critical to them upon reflection.

3.1.1. Myopic decision frames. Although simplification of decision problems is generally required, given limited information processing capacity, the extent to which myopic problem representations are adopted is probably the major fault at the problem representation stage. For instance, people often focus on a single option, a single objective or attribute, or a single assumed state of the world when reasoning about a decision problem (Eisenhardt, 1989; Nutt, 1998; Shepard, 1964). This focusing is particularly acute when the decision is made under stressful conditions such as time pressure or heavy information load (e.g., Payne, Bettman, and Johnson, 1988). In the context of our example of the EPA administrator, the recent NOAA panel on CV studies (National Oceanic and Atmospheric Administration, 1993) highlighted this concern and emphasized being explicit about substitutes (i.e., multiple alternative courses of action) in well-designed CV studies. We believe that people generally will not consider alternative goods of the same type as the good evaluated unless explicitly reminded; see Baron (1997) for a similar argument.

3.1.2. Remedies for myopic decision frames. Remedies for this decision fault *explicitly* encourage people to think of multiple alternatives, multiple states of nature, and multiple attributes or objectives. First, it is important to expand the set of options considered. There should be at least two or more viable options considered when solving any decision problem if the goal is to make better decisions. Bazerman and Neale (1992, p. 69) capture this idea with their advice about house hunting: “fall in love with three, not with one.” In group decision making, there is a similar recommendation that each individual in a group be responsible for thinking of at least two valid solutions to the problem early in the process; this encourages critical thinking about the decision problem.

Keeney (1996) suggests making explicit how much money is spent on alternative public goods in order to encourage people to recognize alternatives in the context of a valuation of a specific public policy option, e.g., putting money into an environmental protection project. This technique might be thought of as a “value” ladder in the same way that “risk” ladders have been used to help communicate risks to people (Smith and Desvouges, 1987). A related idea in the CV literature is the use of “benchmark” amounts intended to “remind respondents that they are

already paying for many public goods and to provide a general idea of the magnitude of those payments" (Mitchell and Carson, 1989, p. 243). While there is an obvious danger that explicit benchmark amounts will provide an anchor for responses, we believe that the need to encourage consideration of alternative uses of resources when making decisions argues for making options as explicit as possible. This is particularly true when it is unlikely that people will retrieve options from memory in the normal course of solving a novel or complex valuation problem.

Does increasing the mental availability of options help in valuation? Although limited in quantity, research suggests that the answer is yes. For example, Baron and Greene (1996) found that explicitly reminding people of other goods helped reduce such problematic valuation effects as the adding up effect, where the value of good A assessed separately plus the value of good B assessed separately often exceeds the value of goods A and B when valued together. Another study by Posavac, Sanbonmatsu, and Fazio (1997) shows that the consistency between attitudes and decisions is higher when alternatives are made salient (explicit) in the decision context. This is most important when people are unlikely to generate (retrieve from memory) the complete set of options on their own. Finally, Nutt (1998) reports that consideration of more than one alternative course of action is associated with greater success in strategic decision making.

A second recommendation is to encourage consideration of multiple futures, because it is clear that people often fail to consider multiple uncertain futures when assessing preferences. One technique which can be used to address this problem is scenario generation (Schoemaker, 1991). The focus of most scenario exercises is not to list all possible futures but to generate at least two, three, or four different future states of the world so that people will be more cognizant of the possible range of uncertainties when making decisions.

Third, individuals should develop a multi-attribute problem representation. One of the basic tools used in decision analysis to help people consider multiple objectives or attributes when making a decision is the construction of a value tree (Keeney, 1992), particularly as an early step in a decision process. A critical issue related to such a representation of objectives is the distinction between fundamental and means objectives or between fundamental values and proxy values (Keeney, 1992). Proxy values reflect means to an end and reflect beliefs about the extent to which satisfying the proxy value will satisfy more fundamental values. Baron (1997, p. 83) suggests that "the failure to make respondents consult their fundamental values seriously may be the largest source of error in valuation of natural resources, as practiced." One solution to this problem mentioned by Baron is to provide respondents with expert opinions regarding which dimensions should be considered in assessing a value for an object (see DeKay and McClelland, 1996).

3.1.3. Inappropriate decision frames. A second class of faults at the problem representation stage is using an inappropriate representation. For example, a person may use an inappropriate analogy in generating a representation. One type of

inappropriate problem representation is from the perspective of the decision maker himself or herself; that is, a person may construct a preference (answer) for a different problem than that which he or she would endorse upon reflection. Inappropriate can also mean that a person is using a different frame than that intended by the researcher, i.e., the problem of scenario misspecification. More generally, use of a frame different from that intended by the researcher is a critical problem in any form of attitude research.

As an example, individuals sometimes use the analogy of charitable giving in constructing an answer to a contingent valuation task (Schkade and Payne, 1994). As a result, the WTP amounts that are expressed are often similar to the amounts given by the respondent to charities, e.g., \$25 or \$50 for a “worthwhile” cause. Clearly, the retrieval and use of a response to a situation perceived to be similar can be an efficient solution to a difficult problem. However, there can be costs associated with such solutions. For example, in solving a decision problem through the use of analogies, respondents may not be properly sensitive to important characteristics of the good being valued, e.g., the quantity of the good (Baron and Greene, 1996). Also, as noted earlier, once a mental model has been retrieved that seems to work for the task at hand, people may not consider alternatives (Legrenzi, Girotto, and Johnson-Laird, 1993) and therefore not appreciate what is omitted from the particular problem representation that they are using.

3.1.4. Remedies for inappropriate decision frames. Mitchell and Carson (1989) suggest that the use of inappropriate representations can be minimized in a contingent valuation study by an intensive program of questionnaire development, including the use of focus groups. In their words, “if the study is well designed and carefully pre-tested, the respondents’ answers to the valuation questions should represent valid WTP responses” (p. 3).

We believe that in addition to careful pre-testing, manipulation checks should be used to confirm the representation respondents are using. In fact, we feel that one obvious property of a well-designed preference measurement process is the use of manipulation checks to ensure that the information presented was understood by the respondent in sufficient detail to have confidence in his or her responses. Fischhoff et al. (1993) provide an example of the manipulation checks used in answering questions about the value of cleaning up rivers near Pittsburgh. Unfortunately, as noted by Fischhoff et al. (1993), the use of manipulation checks is not as common as it should be among valuation researchers. We argue that a minimum standard for a building code dealing with the measurement of preferences should be the use of manipulation checks (for a discussion of probing procedures developed by survey researchers, see DeMaio and Rothgeb, 1996).

Other methods for probing respondents’ thought processes can also be useful as manipulation checks. For example, the researcher can take a verbal protocol from the respondent during the construction of a preference (e.g., Schkade and Payne, 1994). The reasoning processes used by the respondent can then be checked to see if the respondent was answering the question as the researcher understood it.

3.2. Information acquisition and interpretation

In most situations of importance to an individual, a great deal of information may be available from many sources. Some of these sources are external to the individual (e.g., catalogues and brochures, special magazine issues, guidebooks, or opinions from friends for a college choice), and some information is internal, in the individual's memory. Once acquired, information is then interpreted in light of the representation that has been adopted for the decision task. One key component of this interpretation is the coding of potential outcome information in terms of a gain or loss relative to some reference value.

Even though a great deal of information may be available, human processing limitations require selectivity. That is, only a subset of the available information will be examined. Aspects of the individual's information environment can influence selectivity by affecting the salience of various pieces of information. For example, the organization of an information display can make some piece of information either more salient or easier to process (e.g., Bettman, Payne, and Staelin, 1986; Kleimuntz and Schkade, 1993; Russo, 1977). There is a general "concreteness principle" in information acquisition (Slovic, 1972): people will tend to use only explicitly displayed information and will use it untransformed from the form in which it is displayed. A related point is that respondents may use conversational norms to infer that the information presented (and not other, unmentioned information) is the appropriate focus (Schwarz, 1996). Thus, individuals will of necessity be selective; the important issue is how the individual chooses to be selective.

We have stated above that a well-constructed preference should be based upon thorough consideration of the range of information most critical to the individual. Two of the major faults at the information acquisition and interpretation stage that impede attaining this objective are inappropriate selectivity and lack of information comprehension.

3.2.1. Inappropriate selectivity. Selective acquisition of information is related to the focusing phenomenon in constructing mental representations of problems, and therefore many of the faults of these two stages are similar. Thus, one of the major faults of the information acquisition stage is that individuals are too selective and do not acquire and consider enough relevant information. This problem can be compounded by the fact that the information that is most salient (the focus of attention) in a task environment can be uncorrelated or even negatively correlated with the diagnosticity or importance of the information for the individual's goals. Of course, it is also possible that individuals may seek to examine too much information, while not paying enough attention to the subset of information that is most important. A similar problem can occur when a decision maker is overloaded with too much information.

3.2.2. Remedies for inappropriate selectivity. One approach for generating well-constructed preferences, therefore, is to give people high quality information to be used in constructing a preference. For example, many CV researchers argue for the use of in-person interviews, at least in part because it is easier to convey a good deal of high quality information, such as pictures and other visual aids, in such interviews. Urban et al. (1997) have also emphasized the use of various multimedia tools to accelerate the learning of preferences about novel products through extensive, and hopefully more realistic, information provision.

However, as noted by Fischhoff and Furby (1988), the provision of information is not as straightforward a task as might be imagined. Although one does not want to leave out important information from the description of the valuation problem given to a respondent, providing all the relevant information for a valuation task can be a formidable task in itself. For example, Fischhoff and Furby give a checklist of the many components of a contingent valuation task that may need to be explained to the respondent in order for the respondent to generate a well-defined preference. As noted above, it is also clearly possible to overload the respondent with too much information; consequently, comprehension may be affected and various simplifying heuristics may be used as a way to deal with the stress caused by the information load (Payne, 1976). As noted below, we generally need to provide cognitive reasoning tools to help individuals cope with such information load problems.

Given the tension between providing too much information and too little, it is important to consider how one might make information easier to process so that individuals are more likely to examine the information most critical to them. There is a good deal of research suggesting that careful design of information display formats can be extremely useful. A classic example of such an information format effect is Russo's (1977) demonstration that the use of unit price information by consumers was affected by simple changes in how information was made available to the consumer (by presenting unit prices ranked in a list instead of on individual shelf tags). Guidelines also have been developed for such tasks as providing warnings on product labels (e.g. Bettman, Payne, and Staelin, 1986).

3.2.3. Lack of information comprehension. A second major fault at the information acquisition and interpretation stage is that individuals simply may not comprehend the information available or comprehend it in ways not intended by the researcher. Comprehension failures can be due to confusing presentation, lack of knowledge, or other factors (Fischhoff, Welch, and Frederick, 1999). For example, Hsee et al. (1999) review a growing body of research demonstrating that it is more difficult for individuals to evaluate the desirability of values on some attributes than on others. Hsee et al. also make clear that a more difficult to evaluate attribute will be given less weight when options are evaluated separately than when these options are evaluated jointly, which can lead to preference reversals. More generally, if the individual does not understand the object of the preference assessment, then we cannot expect meaningful responses (Fischhoff et al., 1993).

3.2.4. Remedies for lack of information comprehension. We can attempt to ensure that individuals understand the information presented by using manipulation checks and by attempting to make the meaning of the information clear. As stressed earlier, one feature of a well-designed preference measurement process for design purposes should be the use of manipulation checks to ensure that the information presented was understood by the respondent in sufficient detail to have confidence in his or her responses.

Selective attention and comprehension faults due to information format also can be avoided by making the formats of the information as comparable as possible (e.g., using a common format to the extent possible) and by making the meaning of the information as clear as possible. In addition, decision analysts often attempt to make information more easily interpretable by explicitly presenting range information, such as mileage ranges for cars. They often do this by using a common scale such as 0 to 100 for the values of various attributes and anchoring the endpoints of the scales for each attribute by the worst (0) and best (100) attribute values provided by the alternatives in the choice set. Thus, the car in a choice set offering the best gas mileage (e.g., 30 MPG) would be given a value of 100, while the car offering the poorest gas mileage (e.g., 14 MPG) would be given a value of 0. The use of both absolute scale information, i.e., MPG numbers, and relative scale information (0 to 100) in combination may help in interpreting attribute information. The use of meaningful endpoints for scales also helps in the expression stage of preference construction, discussed below.

3.3. Information combination

How people combine information and make tradeoffs has been a focal point of research on preferences. For many, this stage is the heart of the preferential decision making process. It is in making tradeoffs between more of one thing and less of another that one's values are most often revealed to oneself and to outside observers; according to Freeman (1993), the tradeoff ratios between pairs of goods that matter to people are at the core of the economist's concept of value. Many have argued that making tradeoffs is a crucial aspect of high-quality, rational decision making (e.g., Frisch and Clemen, 1994).

It is clear that people do make tradeoffs. Much of market research, for example, has been aimed at measuring tradeoffs using a variety of techniques such as conjoint analysis (e.g., Green and Srinivasan, 1990) and direct methods for assessing exchange rates between attributes (e.g., Aaker, 1991). Similarly, asking people to make tradeoffs explicitly is a key part of decision analysis techniques (e.g., Clemen, 1996). However, the major fault of the information combination stage is that decision makers often avoid making explicit tradeoffs, relying instead on an array of non-compensatory decision heuristics.

3.3.1. Tradeoff avoidance. The avoidance of tradeoffs by individuals is reflected in behaviors such as the selection of the status quo option, the expression of a “protest” zero in CV studies, delaying choice, or selection based on choice set context, violating menu-independence (Luce, 1998). As noted above, the major fault of the information combination stage clearly is avoidance of tradeoffs.

Decision makers may avoid tradeoffs for several reasons. One explanation for tradeoff avoidance is simply that making tradeoffs is a cognitively demanding task that people will try to minimize (Payne, Bettman, and Johnson, 1993). Thus, one observes frequent use of decision heuristics that emphasize comparing alternatives on just one (or a few) attributes as tasks become more complex (i.e., lexicographic-type heuristics; Payne, 1976).

Another explanation is that tradeoffs can be difficult for emotional as well as cognitive reasons (Hogarth, 1987; Luce, Payne and Bettman, 1999). It can be emotionally trying to think about giving up something one values. Values that people are relatively unwilling to trade off have been called “protected” values (Baron and Spranca, 1997), and decisions involving such difficult tradeoffs have termed “cruel choices” (Russell, 1990). Keeney (1996) suggests that public policy decision-makers “rarely” make explicit tradeoffs between costs and the non-economic impacts of programs.

3.3.2. Remedies for tradeoff avoidance. We believe that multiple approaches should be used to help individuals make explicit tradeoffs. Because thinking about tradeoffs is difficult for people, providing more information may be necessary, but not sufficient, for helping people think more deeply about their tradeoffs. As a general building code principle, we suggest that providing cognitive tools can help a respondent think about tradeoffs.

Some simple cognitive techniques for improving the making of tradeoffs have already been mentioned. For instance, expanding the set of options and thinking in terms of two or more alternatives helps people to appreciate that every option is likely to have some advantages and disadvantages. However, there are even more sophisticated thinking tools to facilitate making explicit tradeoffs, such as multiattribute utility analysis (MAUA) and judgment policy analysis.

Thinking tools like MAUA incorporate ways to avoid, or at least minimize, some of the more common faults in thinking about tradeoffs. For example, one of the most important errors identified by decision analysts is the failure to consider an adequate range of levels for particular attributes when deciding tradeoffs. A simple technique for helping people to do a better job of considering attribute ranges is the assessment of swing weights. Typically, the respondent compares alternatives that “swing” between the worst and best levels represented in the given (or plausible) set of alternatives for each attribute and assesses the extent to which the swings in each attribute contribute to overall value differences.

Another feature or characteristic of thinking aids for the making of tradeoffs is the decomposition of complex value judgments into a series of less complex judgments. A simple model like the additive model of preferences is then used to

combine the simpler judgments into an overall valuation. Consistent with the idea that cognitive capacity constraints are a source of error in value construction, decomposing a complex judgment into a series of smaller judgments and then combining those judgments mechanically can improve judgment performance (see Jako and Murphy, 1990; Kleinmuntz, 1990; Einhorn, 1972).

Other thinking tools are available that may help in the construction of preferences through improving attribute weighting. Social judgment theory (Hammond et al., 1975), for instance, provides feedback to an individual about his or her judgment policy and the relative weights apparently placed on various attributes of value. The individual can then use that knowledge to decide how much, if at all, to change the weighting of attributes. Harris, Tinsley, and Donnelly (1988) discuss how judgment policy analysis might be applied to the valuation of natural resources.

However, as noted above, tradeoffs may be avoided for emotional as well as cognitive reasons. Even providing cognitive tools may not be sufficient to overcome individuals' emotional and moral reasons for avoidance. Therefore, an important area for future research is understanding the emotional costs of making tradeoffs and developing techniques which help to both alleviate such costs and encourage reasoned tradeoffs. For example, one advantage of the judgment, feedback, learn, and respond procedure of judgment policy analysis is that it might help overcome the emotional reaction to making even more explicit tradeoffs, although this needs to be investigated.

3.4. Expression or mapping of preferences

The essence of the expression or mapping stage is that an individual must take an internal preferential response and express that response using some specified response mode (e.g., a choice or a scale). However, even when a person has in mind a well-constructed preference, it is not always the case that he or she will be able to translate that preference or value easily and without error (bias) into the response called for in a valuation task (Kahneman, Ritov, and Schkade, 1999). The outcome of the expression or mapping stage should be a response that best reflects the individuals' values and is free of inconsistencies. Two major faults that hinder reaching this goal are influences of scale compatibility and biases in scale usage.

3.4.1. Influences of scale compatibility. One major fault at the preference expression stage is the influence of scale compatibility. Expressions of preference are likely to be overly influenced by the compatibility between an attribute of the stimulus and the response scale or mode. For example, the weight of an attribute whose values are expressed in monetary terms is generally enhanced if the required response is also in monetary terms (Tversky, Sattath, and Slovic, 1988). Similarly, Gregory et al. (1993) have suggested that although a person may hold a strong value for an environmental resource, that value is unlikely to be represented in memory in

dollar terms. Consequently, in attempting to translate such a value into a monetary equivalent, the person may exhibit systematic errors.

3.4.2. Remedies for influences of scale compatibility. We can attempt to overcome scale compatibility effects via triangulation of responses. When the goal of preference measurement is design, one important building code principle is that a preference question should not be asked in just one way; instead, questions about values that are likely to be the result of a constructive process should be asked in multiple ways (see Huber et al., 1993 for evidence of the effectiveness of such an approach). Differences in responses to different but strategically equivalent forms of a value question provide insight into the extent to which the expressed preference is constructed. In addition, one can then ask the respondent to consider the inconsistencies implied by his or her responses. Thus, lability of responses can provide an opportunity for the decision maker to think carefully about the inconsistencies and thereby gain greater insight into the decision (von Winterfeldt and Edwards, 1986). A related point is that asking for multiple types of responses allows different task orders; as suggested by a reviewer, different task orders should in fact be used and any resultant order effects can be used as a source of insight into the construction process.

3.4.3. Biases in scale usage. A second set of expression or mapping faults is due to a number of classic psychophysical phenomena. For example, there are biases such as the tendency to use the center of the scale. It is also likely that any response scale will suffer from anchoring effects. Finally, scale use is affected by the meaningfulness of the scale's labels. If scale labels are vague and do not provide a context for the evaluation, expressed preferences may have little stability. For example, in a study that has implications for an area of debate in public policy, Kahneman, Schkade, and Sunstein (1998) have shown that unpredictability in punitive damage awards may result from the difficulty that people have in mapping their preference for the degree of severity of punishment onto an unbounded dollar scale. This difficulty can be traced to the well-known difficulties people have with magnitude estimation scales in psychophysics. Kahneman, Ritov, and Schkade (1999) provide further discussion of psychophysical phenomena related to valuation tasks.

3.4.4. Remedies for biases in scale usage. A general principle of our proposed building code for the measurement of preferences is that it is better to be explicit about anchors rather than leave the anchors implicit; this helps avoid large individual differences in the implicit anchors used by respondents. Scale usage and the ease of interpretation of a scale can also be improved by using meaningful anchors. For instance, it is recommended in performance appraisal that scales be behaviorally anchored. That is, rather than simply using the term "excellent" on a performance rating scale, one would provide an explicit example of excellent task performance, e.g., responding to a customer's service complaint.

There also may be times when it is better to use a less sensitive, but more robust, response scale. Kahneman, Schkade, and Sunstein (1998) have suggested that one solution for the problem of unpredictability in punitive damage awards is for jurors to use a less sensitive category-type scale (e.g., a seven category scale ranging from no punishment to extremely severe punishment) on which jurors can make more reliable and robust judgments. Those responses might then be converted into punitive damage amounts using a formula. This suggestion of Kahneman et al. is consistent with the general building code principle of only asking people to do those things that they can do well (Norman, 1988).

3.5. A more prominent role for sensitivity analysis

A final suggestion for better measurement cuts across the various stages of the preference construction process. Sensitivity analysis provides insight into the robustness of a decision to changes in the parameters and values used to define it. In our experience, one of the major factors in the increasing use of decision aiding techniques to solve real-world problems has been the ability of computer-based decision support tools to easily do, and display, the results of sensitivity analyses. The presence of an irreducible arbitrariness in measured preferences makes sensitivity analysis an essential tool in the construction of a defensible expression of value, particularly when the value is going to be used as an input for the design and selection of a public policy option.

Typically, sensitivity analysis is done after the problem structuring and valuation stages of a decision. However, given that expressed values may be constructed, we propose that in some cases it may be useful to start the analysis of a decision with a sensitivity analysis. For example, in looking for public values to be used as inputs into a policy decision about environmental protection levels, the decision maker might start by asking under what values a decision would shift. If the value of a natural resource must be above \$X in order to warrant the cost of some action to improve or protect the environment, then the decision maker might use that knowledge to directly ask whether the constructed value is likely to be above or below \$X. Of course, there is a danger that the values assessment could be directed by the decision maker to either support or not support an expression of value greater or less than \$X; however, we suggest that the danger can be ameliorated by making the assessment process as explicit as possible and by including other anchors that are above or below \$X.

Until now, when reviewing construction faults and possible remedies for those faults, we have emphasized situations in which the goal of the measurement task is to design or select a future course of action, particularly in the public domain. When the task is to predict rather than design preferences, another set of construction guidelines is called for, presented next.

4. Measuring preferences for prediction: context matching of information environments

When the goal of measurement is prediction, the “quality” of the measurement is determined by how well the measured preference corresponds to preferential behaviors of interest (e.g., buying a product). The basic building code principle for achieving such correspondence is context matching. That is, one attempts to match the task environment (the context) presented to the respondent for purposes of preference measurement as closely as possible to the task environment that the decision maker will actually face. In implementing context matching, the analyst attempts to determine the relevant factors that might influence preferences in the individual’s environment and then matches the values of those factors in the measurement environment (Simmons, Bickart, and Lynch, 1993; Wright and Kriewall, 1980). For example, we now know from behavioral decision research that such factors as the response mode, the number of options and attributes, time pressure, information display, the context provided by competing options (choice set), and others can affect choices (Payne, Bettman, and Johnson, 1992). The environment in which preferences are elicited should attempt to approximate the individual’s environment on all these factors, particularly if the individual has little familiarity with the decision. The predictive validity of the preference data supplied by a respondent depends crucially on his or her constructing the same preference during the staged measurement episode as he or she will when coming to grips with the real choice (Wright and Weitz, 1977; see also Marder (1997) for a similar idea, the congruence principle). In sum, the measurement situation should attempt to mimic to the extent possible the major factors that affect how people deal with the actual decision problems that are the predictive focus.

Context matching thus demands a thorough knowledge of the properties of choice environments, which requires a fairly complete task analysis of the situations an individual will face when making the actual marketplace or public policy choice. This may not be an easy task to accomplish (see Huber (1997), however, for an attempt to relate properties of market environments to properties of evaluation tasks). First, individuals may not have detailed knowledge of the situations for which they are predicting (Fischhoff, 1991). In addition, in some cases factors may differ systematically across choice environments (e.g., the set of available options may vary). In that case, measurements may need to be taken for each of the major variants of the choice environment and then aggregated based upon the relative frequency of those variants. This will be more difficult to the extent that these variants change over time and hence must also be predicted. Finally, it may be virtually impossible to match on all aspects of the environment. How can we determine which dimensions are most important to match on? As noted above, behavioral decision research has provided evidence for a set of plausible candidate properties (e.g., response mode); perhaps effect sizes from this prior research could be used as broad guidelines for selecting the most critical aspect of the environment on which to match.

5. Discussion

Fischhoff (1997) notes that psychology “is still grappling with how to handle situations in which people might not know what they want” (p. 209). We would extend that view and say that researchers interested in the measurement of preferences are grappling with what it means to measure values in a world of constructed preferences. The purpose of this paper is to begin a dialogue that will hopefully lead to agreement on at least some elements of a building code for the measurement of preferences. We believe that enough research now exists to identify some common “faults” in the construction of preferences that can, and should be, addressed. As noted above, many of our suggestions for addressing preference construction faults involve the provision of tools as well as information (i.e., the materials) for thinking about values. Our suggestions also involve greater effort devoted to the measurement of preferences, with a focus on doing more tasks with fewer respondents. In the next section we consider the conditions under which such efforts should be undertaken.

5.1. *When is a building code needed?*

Given that our guidelines often involve substantial effort on the part of respondents, it is important to consider when such guidelines are necessary. Just as housing construction codes often vary depending upon the type or purpose of a building (e.g., whether for a high-rise office, a residence, a school, or a hospital) and the location of the building (e.g., whether the area is prone to earthquakes), we believe that the application of our building code should be contingent on characteristics of the measurement situation. In particular, we believe that the building code should be applied when preferences are more likely to be constructive and to the degree that the decision for which the preferences are relevant is more critical or important. We consider each of these two aspects of a measurement situation next.

5.1.1. *When are preferences likely to be constructed?* Clearly not all expressions of preference are constructed at the time the valuation question is asked. There are occasions when the expression of values reflects a reference to a well-defined value in memory. For example, for the first two authors of this paper, asking for the name of their favorite college basketball team would yield a quick preferential response from memory, i.e., Duke. In Bayesian terms, there are times when individuals have strong priors.

Fischhoff, Slovic, and Lichtenstein (1980), argue that “people are most likely to have clear preferences regarding issues that are familiar, simple, and directly experienced” (p. 118). Experience allows a person to obtain feedback on the

outcomes of prior judgments and choices and thus to learn preferences. In the field of attitude research, clear preferences have been called “crystallized” values (Schuman and Presser, 1981). Crystallized values are more likely when a person has had the opportunity to think about and/or obtain experience with a good prior to being asked a valuation question (Irwin et al., 1990). Thus, when the objects to be valued are ones with which the respondent is familiar and has directly experienced, the common assumptions of economics are more likely to be justified, i.e., preferences exist to be uncovered.⁶

However, even if a preference is stored in memory, it may not always be retrieved; thus, preferences may be partially constructed even when there has been prior experience. The likelihood that a particular prior preference or value will be retrieved from memory is a positive function of (a) the accessibility in memory (i.e., ease of retrieval) of the prior valuation and (b) the perceived diagnosticity of the prior valuation (i.e., the perceived degree to which that prior value aids the person in achieving his or her goals). On the other hand, the probability of a particular value's being retrieved will be a negative function of the (c) accessibility and (d) diagnosticity of alternative values available in memory (Feldman and Lynch, 1988). Accessibility can be influenced by a variety of factors, including various priming effects, information format, prior usage of the information, and so on (e.g., Alba and Chattopadhyay, 1985; Simmons, Bickart, and Lynch, 1993).

When the object to be valued has any novelty or complexity, e.g., valuing a change in an environmental resource, a retrieved value or preference may simply serve as a starting point or anchor for a more elaborate inferential process in which existing values are related to the problem at hand (Fischhoff and Furby, 1988). The more the valuation exercise requires inferences to be drawn from past situations to a current problem and the more unfamiliar the terms and issues (e.g., the sustainability of an ecological system), the more the expression of preference is likely to reflect a constructive process. Also, the greater the conflict among existing values that might exist, e.g., environmental protection versus economic development, the greater the uncertainties about the future, and larger the number of options to be considered, the more the expression of preferences is likely to reflect a constructive process. Thus, expressed preferences are likely to be constructed in a wide variety of situations.

5.1.2. When does a decision problem warrant use of the guidelines? We have only begun to examine how the building code guidelines should vary according to properties of the measurement situation, e.g., the purpose of the valuation exercise. We have characterized the differences between the measurement of preferences for the two different purposes of decision design and prediction. However, as suggested in several places in this paper, one might also want different “building code” guidelines for different situations involving decision design. Obviously, one would want to make more extensive use of the guidelines the more important the

decision. One might also want to use the guidelines more extensively the greater the degree to which the expression of preference by one individual might impact the consequences experienced by another. The problem of the EPA administrator clearly fits into the category of situations in which we think a more complete building code approach is needed. The characterization of when the guidelines should be used to a greater or lesser extent is a major topic for future research.

One final issue to be addressed is which guidelines are the most critical to implement when a decision to use the building code has been made. We believe that the following guidelines are the most critical: ensuring consideration of multiple decision frames and options, using manipulation checks to ensure understanding, encouraging explicit tradeoffs, use of more than one response mode, and the use of sensitivity analysis.

In the next section we consider a learning perspective on the construction of preferences that presents another view of the guidelines. Then we provide some words of caution for using the building code.

5.2. A learning perspective on the construction of preferences

One way to view the construction of preferences is from a learning perspective. That is, one could view the construction of a preference as a process by which a decision maker comes to “learn” his or her value for an object (see Plott (1996) for a somewhat similar view). From such a perspective, a building code guideline can be seen as a method to encourage more effective learning. For example, facilitating explicit tradeoffs through the use of matching tasks (Tversky et al., 1988) can be viewed as a method for encouraging people to think (learn) about values through the making of tradeoffs. Similarly, the use of multimedia tools to provide complex information (e.g., Urban et al., 1997) is a way to accelerate the learning of preferences. Finally, we argue that helping people learn and consider the distinction between means and fundamental objectives is a critical learning step in the construction of a well-formed preference. However, the contingencies of decision making make clear that the learning of preferences can be highly path-dependent, i.e., what is learned can depend on initial anchor values. Further, convergence on a preference may, at best, be a slow process. Thus, the building code guidelines suggested above represent suggestions for helping people to follow better and perhaps quicker paths in learning their preferences.

5.3. Words of caution

A number of cautions need to be expressed. First, we are sensitive to the danger of overreaching when one talks about differentially treating better and more poorly constructed preferences. However, we argue that the more that measured prefer-

ences are to play a role in an important decision, e.g., a public policy decision, the greater the weight that should be given to the better constructed preferences. As noted earlier in the paper, we make this argument cognizant that there are situations in which an individual's preferences must be given weight, no matter how poorly constructed they may be.

A related point is that some of the suggestions for improving preference measurement appear paternalistic. We do not intend to imply that the analyst should impose values on the individual. Nevertheless, we believe that providing procedures and tools that help individuals discover their own preferences is in the best interests of those individuals, even though this may also influence those preferences. A reviewer made the cogent and related point that the influence of the measurement process on values may be subtle enough that it does not arouse respondents' defenses and hence may lead to "persuading" the respondents.

Second, not all examples of contingencies in decision making should necessarily be viewed as faults in the construction of preferences. For instance, coding an outcome as a gain or a loss relative to a particular reference value may reflect how that outcome in fact will be experienced by the decision maker and hence need not be a fault. Again, however, we argue that clarity in the use of a reference value should be a principle of a building code for preference measurement.

Third, to the extent that a measured preference today is really a prediction about a future preference, there is some research suggesting that people often get it wrong when predicting future preferences (Kahneman, 1994; Huber et al., 1997; Loewenstein and Schkade, 1998). People fail to fully appreciate, for example, the impact of adaptation on the experience of outcomes over time. One implication of a more dynamic view of preferences is that the closer the measurement of preferences is to the likely consumption experience the better. Another implication is that special instructions may be needed to help make the salience of time intervals greater (e.g., Read and Loewenstein, 1995). More generally, the guidelines for what constitutes a well-constructed preference may need to be modified when preferences develop substantially over time and people have great deficiencies in their ability to predict how future outcomes will be experienced.

Fourth, one reviewer expressed a caution about "protected" values or preferences (Baron and Spranca, 1997). Is it appropriate to even try to cause people to rethink such values during a design process? We argue that the importance of considering explicit tradeoffs must be very carefully weighed against the possible emotional costs that may be incurred by the respondent. We suspect that in at least some cases the benefits from making the tradeoffs will in fact outweigh these emotional costs.

Finally, we have suggested a number of ways in which preferences might be better constructed. Each of the suggested methods for improvement has a foundation in the research on human judgment and choice; however, some of the suggested methods reflect much more research than do others. Thus, there is much need for research that would verify that the methods we suggest for improving the construction of preferences are in fact likely to lead to better expressions of values.

6. Conclusion

We argue that there is a need for a “building code” for preference measurement in a world in which many expressions of preference are constructed by people at the time they are asked a valuation question. As in the case of the historical development of building codes, much of the impetus for a building code for constructed preferences comes from an awareness of faults in the processes typically used in preference construction. We have sought to begin the development of a building code for preference measurement by identifying some principles and techniques for preference construction and measurement that should mitigate some of the most common and important construction faults. Many of our suggestions for addressing preference construction faults build on the work of others and involve the provision of tools for thinking about values as well as providing information (i.e., the materials) for thinking about a given expression of preference. This reflects our belief that many construction faults and associated difficulties in the measurement of preferences are frequently due to cognitive limitations interacting with novel and complex task demands. We have also tried to begin a discussion of how a building code’s guidelines should vary as a function of the purposes of the valuation exercise. Clearly, an architectural, constructive view of expressed preferences requires a fundamentally different approach to the measurement of preferences than that which is implied by an archaeological, well-defined existing preferences view.

Acknowledgment

Preparation of this paper was supported by grants from the National Science Foundation and from the Environmental Protection Agency.

Notes

1. The notion of constructed preferences is consistent with the “philosophy of basic values,” which holds that people lack well-differentiated values for all but the most familiar evaluation tasks. The notion of well-defined preferences, on the other hand, is consistent with the “philosophy of articulated values,” which assumes that people have values for all (most) valuation questions and the trick is just to ask the question in the right way (Fischhoff, 1991).
2. Lucas (1986, p. S402) has argued that economists tend “to focus on situations in which the agent can be expected to ‘know’ or to have learned the consequences of different actions so that his observed choices reveal stable features of his underlying preferences.” In a similar vein, Plott (1996) has argued that individuals have a consistent set of preferences but that such preferences only become known to the individual (are “discovered”) through thought and experience.
3. See Sen (1997) for a different but related discussion of how preferences are sensitive to choice processes and the goals evoked by those processes.
4. Clearly there are situations in which an individual’s preferences must be given weight no matter how poorly constructed they may be. There are sometimes “rights” to preferences; see Sunstein (1990) for a further discussion of this point.

5. Other properties of a well-constructed preference might include such things as high levels of test-retest reliability (stability); however, as noted by one reviewer, stability may result from the repeated use of similar information and processes in the construction process, and should not be taken as a sufficient indicator of the retrieval of a well-defined preference (see Sudman, Bradburn, and Schwartz, 1996) for a related discussion.
6. However, even such crystallized values can still be subject to task and context effects (Krosnick and Schuman, 1988). For a further discussion of such effects, see Kahneman, Schkade, and Sunstein (1998).

References

- Aaker, David A. (1991). *Managing Brand Equity*. New York: The Free Press.
- Alba, Joseph W. and Amitava Chattopadhyay. (1985). "Effects of Context and Part-Category Cues on Recall of Competing Brands," *Journal of Marketing Research* 22, 340–349.
- Arrow, Kenneth J. (1982). "Risk Perception in Psychology and Economics," *Economic Inquiry* 20, 1–9.
- Baron, Jonathan. (1988). *Thinking and Deciding*. Cambridge, England: Cambridge University Press.
- Baron, Jonathan. (1997). "Biases in the Quantitative Measurement of Values for Public Decisions," *Psychological Bulletin* 122, 72–88.
- Baron, Jonathan and Joshua Greene. (1996). "Determinants of Insensitivity to Quantity in Valuation of Public Goods: Contribution, Warm Glow, Budget Constraints, Availability, and Prominence," *Journal of Experimental Psychology: Applied* 2, 107–125.
- Baron, Jonathan and Mark D. Spranca. (1997). "Protected Values," *Organizational Behavior and Human Decision Processes* 70, 1–16.
- Bazerman, Max H. and Margaret A. Neale. (1992). *Negotiating Rationally*. New York: The Free Press.
- Bettman, James R., Mary Frances Luce, and John W. Payne. (1998). "Constructive Consumer Choice Processes," *Journal of Consumer Research* 25, 187–217.
- Bettman, James R., John W. Payne, and Richard Staelin. (1986). "Cognitive Considerations in Designing Effective Labels for Presenting Risk Information," *Journal of Marketing and Public Policy* 5, 1–28.
- Clemen, Robert T. (1996). *Making Hard Decisions*, 2nd edition. Belmont, CA: Duxbury Press.
- DeKay, Michael L. and Gary H. McClelland. (1996). "Probability and Utility Components of Endangered Species Preservation Programs," *Journal of Experimental Psychology: Applied* 2, 60–83.
- DeMaio, Theresa J. and Jennifer M. Rothgeb. (1996). "Cognitive Interviewing Techniques: In the Lab and in the Field." In Norbert Schwarz and Seymour Sudman (eds.), *Answering Questions: Methodology for Determining Cognitive and Communication Processes in Survey Research*. San Francisco: Jossey-Bass.
- Einhorn, Hillel J. (1972). "Expert Measurement and Mechanical Combination," *Organizational Behavior and Human Performance* 7, 86–106.
- Eisenhardt, Kathleen M. (1989). "Making Fast Strategic Decisions in High Velocity Environments," *Academy of Management Journal* 32, 543–575.
- Feldman, Jack M. and John G. Lynch, Jr. (1989). "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intentions, and Behavior," *Journal of Applied Psychology* 73, 421–435.
- Fischhoff, Baruch. (1991). "Value Elicitation: Is Anything in There?" *American Psychologist* 46, 835–847.
- Fischhoff, Baruch. (1997). "What Do Psychologists Want? Contingent Valuation as a Special Case of Asking Questions." In Raymond J. Kopp, Werner W. Pommerehne, and Norbert Schwartz (eds.), *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Boston: Kluwer Academic Publishing.
- Fischhoff, Baruch and Lita Furby. (1988). "Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility," *Journal of Risk and Uncertainty* 1, 147–184.

- Fischhoff, Baruch, Marilyn Jacobs Quadrel, Mark Kamlet, George Loewenstein, Robyn Dawes, Paul Fishbeck, Steven Klepper, Jonathan Leland, and Patrick Stroh. (1993). "Embedding Effects: Stimulus Representation and Response Mode," *Journal of Risk and Uncertainty* 6, 211–234.
- Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein. (1980). "Knowing What You Want: Measuring Labile Values." In Thomas S. Wallsten (ed.), *Cognitive Processes in Choice and Decision Behavior*. Hillsdale, NJ: Lawrence Erlbaum.
- Fischhoff, Baruch, Ned Welch, and Shane Frederick. (1999). "Construal Processes in Preference Assessment," *Journal of Risk and Uncertainty* 19, 139–176.
- Freeman, A. Myrick, III. (1993). *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future.
- Frisch, Deborah and Robert T. Clemen. (1994). "Beyond Expected Utility: Rethinking Behavioral Decision Research," *Psychological Bulletin* 116, 46–54.
- Green, Paul E. and V. Srinivasan. (1990). "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing* 45, 33–41.
- Gregory, Robin, James Flynn, Stephen M. Johnson, Theresa A. Satterfield, Paul Slovic, and Robert Wagner. (1997). "Decision-Pathway Surveys: A Tool for Resource Managers," *Land Economics* 73, 240–254.
- Gregory, Robin, Sarah Lichtenstein, and Paul Slovic. (1993). "Valuing Environmental Resources: A Constructive Approach," *Journal of Risk and Uncertainty* 7, 177–197.
- Grether, David M. and Charles R. Plott. (1979). "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review* 69, 623–638.
- Hammond, Kenneth R. (1996). *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice*. New York: Oxford University Press.
- Hammond, Kenneth R., Thomas R. Stewart, Berndt Brehmer, and D. O. Steinmann. (1975). "Social Judgment Theory." In M. F. Kaplan and S. Schwartz (eds.), *Human Judgment and Decision Processes*. New York: Academic Press.
- Harris, Charles C., Howard E. A. Tinsley, and Dennis M. Donnelly. (1988). "Research Methods for Public Amenity Resource Valuation: Issues and Recommendations." In George L. Peterson, B. L. Driver, and Robin Gregory (eds.), *Amenity Resource Valuation: Integrating Economics With Other Disciplines*. State College, PA: Venture Publishing.
- Hogarth, Robin M. (1987). *Judgment and Choice*, 2nd ed. New York: John Wiley.
- Hsee, Christopher H., George F. Loewenstein, Sally Blount, and Max H. Bazerman. "Preference Reversals between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis," *Psychological Bulletin* 125, 576–590.
- Huber, Joel. (1997). "What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments." In *1997 Sawtooth Conference Proceedings*. Sequim, WA: Sawtooth Software.
- Huber, Joel, John G. Lynch, Jr., Kim P. Corfman, Jack Feldman, Morris C. Holbrook, Donald R. Lehmann, Bertrand Munier, David A. Schkade, and Itamar Simonson. (1997). "Thinking About Values in Prospect and Retrospect: Maximizing Experienced Utility," *Marketing Letters* 8, 323–334.
- Huber, Joel, Dick R. Wittink, John A. Fiedler, and Richard Miller. (1993). "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice," *Journal of Marketing Research* 30, 105–114.
- Irwin, Julie R., David Schenk, Gary H. McClelland, William D. Schulze, Thomas Stewart, and Mark Thayer. (1990). "Urban Visibility: Some Experiments on the Contingent Valuation Method." In C. V. Mathai (ed.), *Visibility and Fine Particles*. Pittsburgh, PA: Air and Waste Management Association.
- Jako, Robert A. and Kevin R. Murphy. (1990). "Distributional Ratings, Judgment Decomposition, and Their Impact on Interrater Agreement and Rating Accuracy," *Journal of Applied Psychology* 75, 500–505.
- Janis, Irving L. and Leon Mann. (1977). *Decision Making*. New York: The Free Press.
- Kahneman, Daniel. (1994). "New Challenges to the Rationality Assumption," *Journal of Institutional and Theoretical Economics* 150, 18–36.

- Kahneman, Daniel, Ilana Ritov, and David A. Schkade. (1999). "Economic Preferences or Attitude Expression?: An Analysis of Dollar Responses to Public Issues," *Journal of Risk and Uncertainty* 19, 203–236.
- Kahneman, Daniel, David A. Schkade, and Cass R. Sunstein. (1998). "Shared Outrage and Erratic Awards: The Psychology of Punitive Damages," *Journal of Risk and Uncertainty* 16, 49–86.
- Keeney, Ralph L. (1992). *Value-Focused Thinking: A Path to Creative Decisionmaking*. Cambridge, MA: Harvard University Press.
- Keeney, Ralph L. (1996). "Valuing Billions of Dollars," In Richard J. Zeckhauser, Ralph L. Keeney, and James K. Sebenius (eds.), *Wise Choices: Decisions, Games, and Negotiations*. Boston: Harvard Business School Press.
- Kleinmuntz, Don N. (1990). "Decomposition and the Control of Error in Decision-Analytic Models." In Robin M. Hogarth (ed.), *Insights in Decision Making: A Tribute to Hillel J. Einhorn*. Chicago: University of Chicago Press.
- Kleinmuntz, Don N. and David A. Schkade. (1993). "Information Displays in Decision Making," *Psychological Science* 4, 221–227.
- Krosnick, Jon A. and Howard Schuman. (1988). "Attitude Intensity, Importance, and Certainty and Susceptibility to Response Effects," *Journal of Personality and Social Psychology* 54, 940–952.
- Legrenzi, P., V. Girotto, and P. N. Johnson-Laird. (1993). "Focusing in Reasoning and Decision Making," *Cognition* 49, 37–66.
- Loewenstein, George and David A. Schkade. (1998). "Wouldn't It Be Nice? Predicting Future Feelings," In Ed Diener, Norbert Schwartz, and Daniel Kahneman (eds.), *Hedonic Psychology: Scientific Approaches to Enjoyment, Suffering, and Well-Being*. New York: Russell Sage Foundation.
- Lucas, Robert E., Jr. (1986). "Adaptive Behavior and Economic Theory," *Journal of Business* 59, S401–S426.
- Luce, Mary Frances. (1998). "Choosing to Avoid: Coping with Negatively Emotion-Laden Consumer Decisions," *Journal of Consumer Research* 24, 409–433.
- Luce, Mary Frances, John W. Payne, and James R. Bettman. (1999). "Emotional Trade-Off Difficulty and Choice," *Journal of Marketing Research* 36, 143–159.
- March, James G. (1978). "Bounded Rationality, Ambiguity, and the Engineering of Choice," *Bell Journal of Economics* 9, 587–608.
- Marder, Eric. (1997). *The Laws of Choice*. New York: Free Press.
- McFadden, Daniel. (1999). "Rationality for Economists," *Journal of Risk and Uncertainty*, 19, 73–106.
- Mitchell, Robert C. and Richard T. Carson. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.
- National Oceanographic and Atmospheric Administration. (1993). "Report of the NOAA Panel on Contingent Valuation," *Federal Register* 58, 4602–4614.
- Norman, Donald A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.
- Nutt, Paul C. (1998). "How Decision Makers Evaluate Alternatives and the Influence of Complexity," *Management Science* 44, 1148–1166.
- Payne, John W. (1976). "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance* 16, 366–387.
- Payne, John W., James R. Bettman, and Eric J. Johnson. (1988). "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14, 534–552.
- Payne, John W., James R. Bettman, and Eric J. Johnson. (1992). "Behavioral Decision Research: A Constructive Processing Perspective," *Annual Review of Psychology* 43, 87–131.
- Payne, John W., James R. Bettman, and Eric J. Johnson. (1993). *The Adaptive Decision Maker*. Cambridge, England: Cambridge University Press.
- Plott, Charles R. (1996). "Rational Individual Behaviour in Markets and Social Choice Processes," In Kenneth J. Arrow, Enrico Colombatto, Mark Perlman, and Christian Schmidt (eds.), *The Rational Foundations of Economic Behavior*. New York: St. Martins Press.
- Posovac, Steven S., David M. Sanbonmatsu, and Russell H. Fazio. (1997). "Considering the Best Choice: Effects of the Salience and Accessibility of Alternatives on Attitude-Decision Consistency," *Journal of Personality and Social Psychology* 72, 253–261.
- Rabin, Matthew. (1998). "Psychology and Economics," *Journal of Economic Literature* 36, 11–46.

- Read, Daniel and George Loewenstein. (1995). "Diversification Bias: Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices," *Journal of Experimental Psychology: Applied* 1, 34–49.
- Russell, Milton. (1990). "The Making of Cruel Choices." In P. Brett Hammond and Rob Coppock (eds.), *Valuing Health Risks, Costs and Benefits for Environmental Decision Making: Report of a Conference*. Washington, DC: National Academy Press.
- Russo, J. Edward. (1977). "The Value of Unit Price Information," *Journal of Marketing Research* 14, 193–201.
- Russo, J. Edward and Paul J. H. Schoemaker. (1989). *Decision Traps*. New York: Fireside.
- Schkade, David A. and John W. Payne. (1994). "How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis of Willingness-to-Pay for an Environmental Regulation," *Journal of Environmental Economics and Management* 26, 88–109.
- Schoemaker, Paul J. H. (1991). "When and How to Use Scenario Planning: A Heuristic Approach with Illustration," *Journal of Forecasting* 10, 549–564.
- Schuman, Howard and Stanley Presser. (1981). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.
- Schwarz, Norbert. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Lawrence Erlbaum.
- Sen, Amartya. (1997). "Maximization and the Act of Choice," *Econometrica* 65, 745–779.
- Shepard, Roger N. (1964). "On Subjectively Optimum Selection among Multiattribute Alternatives." In Maynard W. Shelley and Glenn L. Bryan (eds.), *Human Judgments and Optimality*. New York: Wiley.
- Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch, Jr. (1993). "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research* 20, 316–329.
- Simon, Herbert A. (1955). "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics* 69, 99–118.
- Slovic, Paul. (1972). "From Shakespeare to Simon: Speculations—and Some Evidence—About Man's Ability to Process Information," *Oregon Research Institute Bulletin* 12.
- Slovic, Paul. (1995). "The Construction of Preference," *American Psychologist* 50, 364–371.
- Smith, V. Kerry and William Desvouges. (1987). "An Empirical Analysis of the Economic Value of Risk Changes," *Journal of Political Economy* 95, 89–114.
- Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- Sunstein, Cass R. (1990). "Preferences and Politics," *Philosophy and Public Affairs* 20, 3–34.
- Tversky, Amos. (1996). "Contrasting Rational and Psychological Principles in Choice." In Richard J. Zeckhauser, Ralph L. Keeney, and James K. Sebenius (eds.), *Wise Choices: Decisions, Games, and Negotiations*. Boston: Harvard Business School Press.
- Tversky, Amos, Shmuel Sattath, and Paul Slovic. (1988). "Contingent Weighting in Judgment and Choice," *Psychological Review* 95, 371–384.
- Urban, Glen L., John R. Hauser, William J. Qualls, Bruce D. Weinberg, Jonathan D. Bohlmann, and Roberta A. Chicos. (1997). "Information Acceleration: Validation and Lessons from the Field," *Journal of Marketing Research* 34, 143–153.
- Von Winterfeldt, Detlof and Ward Edwards. (1986). *Decision Analysis and Behavioral Research*. Cambridge, England: Cambridge University Press.
- Wilson, Timothy D. and Jonathan W. Schooler. (1991). "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions," *Journal of Personality and Social Psychology* 60, 181–192.
- Wright, Peter L. and Mary Ann Kriewall. (1980). "State-of-Mind Effects on the Accuracy with Which Utility Functions Predict Marketplace Choice," *Journal of Marketing Research* 17, 277–293.
- Wright, Peter L. and Barton Weitz. (1977). "Time Horizon Effects on Product Evaluation Strategies," *Journal of Marketing Research* 14, 429–443.
- Yankelovich, Daniel. (1991). *Coming to Public Judgment: Making Democracy Work in a Complex World*. Syracuse, NY: Syracuse University Press.