

Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility

BARUCH FISCHHOFF

Eugene Research Institute and Carnegie Mellon University

LITA FURBY

Eugene Research Institute

Key words: contingent valuation, value measurement, transactions, atmospheric visibility

Abstract

People express their value for a good when they pay something for it. Interpreting *good* and *payment* very broadly, we offer a general analytical framework for characterizing such transactions. This framework is suitable for interpreting actual transactions as well as for creating hypothetical transactions for research purposes. It is described here both in general terms and with special application to one particular kind of transaction, contingent valuation studies in which individuals estimate the value of possible changes in atmospheric visibility. In these transactions, as in many others, risk (of undesired changes in visibility) is one principal feature; at least some uncertainty often surrounds other transaction features as well (For example: How much will visibility really change if I promise to pay for it? Will I really have to pay?). The framework presented here conceptualizes any transaction as involving (a) a good, (b) a payment, and (c) a social context within which the transaction is conducted. Each of these aspects in turn has a variety of features that might and in some cases should affect evaluations. For each such feature, the framework considers first the meaning of alternative specifications and then the difficulties of ensuring that they are understood and evaluated properly. As a whole, the framework provides an integrated approach to designing evaluation studies and interpreting their results.

Finding out what something is worth to people is easiest when they engage in observable transactions involving it. Thus, the value of a consumer item is related to how much money people pay to acquire it. The value of a recreational experience is revealed, in part, by the amount of time people spend on it. The value of a political cause is, to some extent, reflected in the energies that people expend pursuing it—getting, in return, either the goal itself or the satisfaction of knowing they at least tried to obtain it. In all such overt transactions, latent predispositions are revealed in concrete actions.

People may reveal the value to themselves of a particular thing in a variety of

transactions. For example, they may spend money to visit a scenic area (through admission fees and travel costs), spend time getting there and back, and spend energy lobbying their legislators to preserve the area. The total value of a thing should at least be equal to the sum of all these "payments"—assuming that all the transactions are satisfactory ones. A *satisfactory* transaction is one involving individuals who are fully informed, uncoerced, and able to identify their own best interests. A thing might be worth more than the sum of transaction payments made for it in cases where the opportunity to acquire it has not arisen, or where people get a bargain (e.g., if something is so abundant that people can obtain it for less than they would be willing to pay if pressed).

Social scientists concerned with the worth of things have been quick to seize on these overt expressions of value. Economists collect and interpret prices for a great variety of marketed goods. Sociologists examine the time spent on different activities. Political scientists look at the money and effort devoted to various causes. Anthropologists consider the diverse sacrifices made to secure social and spiritual status.

To interpret these transactions credibly, investigators need procedures to undo complicating factors, such as paying for a thing in multiple ways, acquiring substitutes for it, or getting it at bargain prices. Even more involved procedures are needed for cases in which the thing of interest is not traded by itself, but only as part of more complex things. For example, neighborhoods are worth something to most people. However, neighborhoods per se are not exchanged in transactions. Rather, they are one factor affecting the monetary price paid for homes and the time spent looking for them. Determining the value of neighborhoods requires isolating their impact on those transactions (Maler, 1977). Similarly, people are seldom asked directly to accept a risk of death in return for a monetary payment, but are often asked to do so indirectly when the risk is one feature of a job (Jones-Lee, 1976; Thaler & Rosen, 1976; Viscusi, 1983).

Whether something is traded by itself or as part of a package, the interpretation of a transaction depends on how satisfactory the transaction was. Were any of the parties involved in it denied relevant information? Were they too lazy to collect information, perhaps exaggerating how much they knew already? Were they coerced by sales or social pressure? Did they fail to identify or think through all relevant options? Could they anticipate the enjoyment that would come with securing the thing being offered, or that would be foregone by declining it?

Where a transaction has imperfections, its interpretation requires determining what transaction would have been made under more satisfactory conditions. That means undoing the imperfections by imagining counterfactual states, such as how much neighbors would have protested a hazardous waste facility had they understood its risks better, or how much a couple would have paid for a house had they not been so much better negotiators than the sellers, or how much time that couple would have spent at the Miró exhibition had they not had their children in tow.

Even with a simple, satisfactory transaction, interpreting people's values means

being able to see the transaction from their perspective. That is, what did they think that they were getting—and giving? For example, it would be a mistake to assume that the price paid for a concert ticket reflects only the value of hearing the music, when attendance at the concert is a social necessity. It might also be a mistake to assume that people do not value their leisure time just because they put less effort into planning it than they put into rationing their working hours; not thinking too hard may be part of their leisure experience. Or, it might be a mistake to assume that students who skip a lecture on AIDS do not care about its risks; they may think that they know it all already, or may just not like the lecturer, or may value more the opportunity to sneak off with a friend. Thus, even though a transaction is open to view, its meaning may depend on less observable perceptions.

Real-life transactions are often so rich that people can construe (or misconstrue) them in many different ways, each of which might have to be considered by those hoping to interpret transactions. Investigators might also face an equally long list of possible threats to a transaction's satisfactoriness. As a matter of practicality, most disciplines develop some consensus on which of these possible complications to consider when interpreting transactions. Failure to address these canonical factors leads to rejected manuscripts and diminished professional reputations. Scientific disciplines differ on which factors are essential—and, as a result, have difficulty cooperating and accepting one another's conclusions. The complexity of divining how participants interpret real-life transactions makes some investigators so uncomfortable that they prefer artificial settings, where they can restrict the set of factors potentially relevant to interpreting a transaction. Unfortunately, the histories of experimental psychology and survey research show the continual discovery of seemingly irrelevant factors that can, in fact, affect behavior (McGuire, 1969; National Research Council, 1982). Despite their apparent (and intended) simplicity, experiments and interviews provide rich cues as to how participants should respond. Indeed, their very novelty and artificiality may prompt participants to search for clues as to what these studies are all about.

Controlled settings are also used to create transactions that cannot otherwise be observed. Some transactions are too scattered in time or space for efficient observation; some occur only privately; some just do not occur at all. For example, market researchers ask people to evaluate products that might some day be introduced. Psychologists offer unusual gambles in order to test specific hypotheses about decision-making processes. Increasingly, economists create "contingent markets" (Cummings, Brookshire & Schulze, 1986; Mitchell & Carson, in press; Smith & Desvousges, 1986), offering possible transactions that do not exist in the real world (e.g., being able to pay to preserve atmospheric visibility, or being able to agree to compensation for loss of visibility). Most commonly called contingent valuation studies, these experiments ask participants to respond as if a market existed with those transactions. The remainder of the present analysis is devoted to developing a comprehensive framework for designing and interpreting contingent valuation studies, with a special focus on those involving changes in atmospheric visibility (Rowe & Chestnut, 1982; Tolley et al., 1986). This framework is form-

ulated, however, so as to allow generalization to other research methods that use transactions to reveal values. In this sense, it is a step toward a general conceptualization of value measurement.

1. Background

1.1 Contingent valuation as an archetype

Although contingent valuation (CV) is a novel (and somewhat controversial) enterprise for economics, it is part of a long tradition of valuation methodologies, spread over many disciplines. In order for CV to achieve internal consistency and paradigm status, its practitioners must agree on which factors must be considered when creating and interpreting the transactions in their studies. In order to achieve credibility with other disciplines in the evaluation business, CV must consider all the factors deemed critical by those disciplines. Otherwise, CV researchers must demonstrate, either empirically or rhetorically, that the factors that they ignore have no effect on the transactions that they pose.

In this sense, any valuation methodology could serve as a vehicle for developing a comprehensive framework, since all must consider the same factors. CV provides a particularly useful focus for several reasons. One is that, after 15 years of intense and productively uncoordinated activity, it is ripe for some codification (Smith, 1986). A second reason is that most CV studies have been relatively uninformed by the relevant research literatures in other disciplines. Now that CV researchers have a feeling for the special nature of their tasks, they may be ready for the rapid progress that comes with exploiting the hard-earned lessons of related disciplines. A third reason is that CV studies pose tasks that are especially suited to systematic analysis. They attempt to create very specific markets offering very specific goods. As a result, it is possible, in principle, to determine analytically whether the staged transaction is the correct one (i.e., does it create the appropriate market and offer the intended good?). The point of reference for such analyses is some contemplated action. An example might be a proposal to increase utility bills in order to pay for scrubbers on fossil fuel power plants, thereby improving atmospheric visibility and reducing health risks; or a proposal to increase park entrance fees in order to repair upland trails; or a proposal to lower taxes as compensation for accepting a hazardous waste processing facility in one's neighborhood. Such concrete policy proposals can provide much more detailed guidance for creating tasks than, for example, the psychological theories that constrain experimental gambles or the general desire to sample the public mood which motivates much opinion polling.

Taking a comprehensive perspective should provide an opportunity both to coordinate CV research and to integrate it with the experience of other disciplines. However, this may not be an opportunity that CV researchers are eager to exploit.

Other disciplines have passed up similar chances. For example, psychologists often feel little discomfort when relying on data that sociologists find quite unconvincing, and vice versa. Indeed, if any field gets sufficiently large, it can ignore the competition, at least as long as its members talk only to one another. For sub-disciplines in a formative stage, as contingent valuation has been, merely achieving legitimacy within a larger discipline may be difficult enough, without worrying about the judgments of outsiders. CV researchers have faced the particular problems of having a parent discipline, economics, that is hostile to asking people hypothetical questions, the essence of CV transactions.

In any field, the main intrinsic reason for undertaking a comprehensive perspective is its intellectual challenge. The main extrinsic reason is convincing outsiders of the validity of one's results. CV studies often have the specific goal of setting social policy—by pricing unmarketed goods so that they will be considered in cost-benefit analyses. In such circumstances, one cannot be satisfied with intradisciplinary standards, but must be ready to handle criticism from all comers.

1.2 Components of transactions

Any proposed transaction has three constituents: something being received, something being given in exchange, and a social context within which the exchange would be enacted. In an economic transaction, these might be called the *good*, the *payment*, and the *marketplace*. Each can affect whether a proposed transaction is conducted. For example, other things being equal, a transaction should be more attractive (and more likely to be accepted) if the good is enhanced, the payment is reduced, and the transaction has no adverse effects on nonparticipants (i.e., externalities).

For a transaction to be satisfactory, each of these three constituents must be well defined and well understood by all participants involved. For routine transactions in actual markets, many of these details are often provided as a matter of course. For example, when buying toothpaste, the nature of the product is spelled out on the package (e.g., ingredients, net weight, and test results); the price is clearly displayed; government agencies assure quality and safety; the applicable social norms are well understood. Indeed, because the marketplace for this type of transaction is so familiar, the consumer can focus on the good and whether it is worth the payment demanded. Moreover, an observer can feel fairly confident that the consumer thinks of the good quite narrowly, without worrying, say, about what precedents are being set or how basic political and property rights are being affected.

The task of market researchers is often simplified by invoking these routine settings. Researchers do not have to describe the ambience of a supermarket, or the kinds of consumer protection that it offers, or the experience of paying a cashier for a purchase. Therefore, it is plausible that the participants in market research

studies can imagine a proposed good being offered and estimate what they would pay for it in an actual market. There are, however, limits to people's imaginations, especially with more novel goods and settings.¹

Novelty complicates the staging of a transaction both quantitatively and qualitatively. In general, the more novel a transaction, the more details of it will need to be explained and the more difficult ensuring that those details are understood will be. For example, the prospectus for a mutual fund or an initial stock offering shows the amount of detail considered necessary to describe the risks and possible benefits associated with one novel good, using a familiar form of payment (i.e., money) and offered in a reasonably familiar social context (at least for experienced investors). Even for readers accustomed to such material, the sheer quantity of information in a prospectus can be a barrier to understanding. For unaccustomed readers, the technical details in a prospectus may not be comprehensible even in isolation. Without such understanding, a transaction cannot be considered satisfactory. As we shall see, daunting amounts of detail might be needed to specify such unesoteric goods as visibility.

In the supermarket or stock market, various authorities (e.g., regulators, the courts) try to ensure that transactions are satisfactory. In staged transactions, like CV studies, that responsibility falls to investigators (and, indirectly, to those relying on their work). For every feature of a transaction that could affect behavior, the investigator must first determine how it should be defined and then establish that participants have understood it as intended. Specifically, participants should understand the experimental situation well enough to identify the course of action that is in their own best interests. That means being able to see the situation-specific examples of their own general values and predispositions. Whether such understanding has been attained is an empirical question requiring direct observation. For example, simply telling people everything provides no guarantee that they have understood everything. Such a strategy might even impede understanding if attention to critical features of the contingent market is diverted by a deluge of details about features that could have gone without saying because they have little practical effect on decisions.

1.3 *The framework*

As mentioned, the present analysis offers a comprehensive framework for specifying the features that define any transaction, using, as a case in point, contingent valuation studies involving changes in visibility. These features are specified for each constituent of a transaction: the *good*, the *value measure* (or *payment*), and the *social context* (or *marketplace*) of the exchange.

With actual transactions, investigators must take whatever observations they can get, and then try to discern what features the transactions had and how satisfactory they were for participants. With staged transactions, the first task of investigators is to determine what transactions they want to create in order to serve their

theoretical or practical purposes. To this end, our analysis discusses in general terms how to choose among alternative specifications. It then considers practical obstacles to staging an intended transaction, either through omitting essential features, so that participants do not know exactly what the transaction is, or by miscommunicating those features, so that participants consider the wrong transaction.

In staged transactions, investigators typically perform various operations on the raw data and summarize the results in some aggregate form. Just as much care is required in analyzing the data as in formulating the questions. Although we do not discuss data analysis in any detail here, a worked example considering alternative data analyses can be found in Fischhoff and Furby (1987).

2. The good

Although they are transferred as wholes, goods may be thought of as bundles of *attributes*, representing outcomes of accepting the transaction that might be valued either positively or negatively. The first step in defining a good is identifying its potentially valued attributes. This is also the first step in the orderly evaluation of a good. In the terms of decision theory, for individuals with completely articulated preferences, mere mention of these attributes would evoke a multiattribute utility function over a space containing all goods having these attributes. Actually locating a good within that space means specifying it in terms of each attribute. This second step, too, must be undertaken both by those who offer goods and by those who evaluate them. The first of these steps determines *why* someone might value a good (considering the kinds of attributes it has), whereas the second determines *how much* they value it (once they know how it rates on each attribute). These steps are discussed below as the *substantive* and *formal* components of a good's definition, respectively. For investigators interpreting actual transactions, producing these definitions is a descriptive enterprise. They must determine how participants perceived the goods they considered. For investigators creating hypothetical transactions, the definition must be derived from the (policy or theoretical) question that the study is intended to answer. Consider, for example, an attempt to determine the value of visibility as one input to a cost-benefit analysis of a pollution abatement program. Atmospheric haze and the prevalence of plumes are two possible attributes in a substantive definition of visibility (Malm et al., 1981). If haze alone is affected, then the substantive definition should make that clear. If the program is intended to reduce haze by 50%, but has only a 75% chance of succeeding (two aspects of the formal definition), then both those aspects of the good's quantity need to be specified.

The purest transaction for valuing visibility would offer a change in visibility and no accompanying effects on other valued attributes (e.g., respiratory well-being). Such a transaction could provide better visibility, asking respondents how much money they would pay to secure it; or, it could provide worse visibility, ask-

ing respondents how much they would want in order to accept it. The transaction could even result in current visibility, if respondents paid to prevent a threatened deterioration. Indeed, people could be asked to pay for reduced visibility, where the benefit is slowing the rate of decline.

People respond to the offer they perceive, which may differ from the actual offer. Ensuring that the good being evaluated matches the intended definition requires an analysis of evaluators' psychology. What aspects of the definition need to be explained? What misinterpretations are possible? What misconceptions need to be undone? Whenever a good has any novelty (as it does in all CV transactions), the evaluation process involves inference, as evaluators try to figure out what the novel good is worth on the basis of their basic values and the worth of related, more familiar goods (Rokeach, 1973). Making such inferences requires a very clear idea of what the good is. Achieving that clarity in study design is a craft, but one that can be aided by the scientific study of potential pitfalls. That study has been part of the research methodology of every social science that asks people to answer unfamiliar evaluation questions (Turner & Martin, 1985). The following discussion raises representative issues.

One general design question involves default assumptions—i.e., what participants infer about a good's definition in the absence of explicit instruction. For example, when a good is provided, do people assume that it is certain to be provided unless they are told otherwise? Or, do they treat it as having about the same probability of delivering promised benefits as do other sociotechnical programs? Do they naturally assume that a changed state of nature is caused by human intervention? Or, do they think that it could be the result of natural causes? When people make the right assumption, that detail can be left unsaid in the design of the study, simplifying the task. Where they make the wrong assumption, the correct definition might have to be mentioned pointedly and repeatedly.

When the good that respondents perceive differs from the good that was intended, then extrapolating from the obtained evaluations to the needed ones requires some bridging theory. Plausible extrapolations may be fairly available for some aspects of the formal definition. For example, if respondents believe that a visibility change is guaranteed when there is only a 50% chance of it occurring, then halving their evaluations might seem reasonable. However, it could also be the case that sure things are evaluated especially highly, whereas 50% chances are treated as having a somewhat indeterminate probability. Extrapolating from one substantive feature to another is much more complex. For example, evaluators who report being irritated by a distinctive plume could feel quite differently about ambient haze. There would be little way of knowing this without actually asking them. Moreover, eliminating both intense haze and a distinct brown plume might be worth more than eliminating either one alone, but less than the sum of eliminating each separately—reflecting the decreased marginal utility of additional improvements. Or, the combined change might be worth more than the sum of the individual changes—if an evaluator felt that the combined changes created “air worth looking at,” while neither change alone did.

2.1 Substantive definition

2.1.1 Attribute(s). Although a good may have many attributes, not all need affect its value. Those that do must be part of its definition. Unless an investigator somehow knows in advance which attributes affect a good's value, its definition should include all reasonable possibilities, starting with those attributes whose effects on evaluations interest the investigator. Adopting the investigator's perspective, these might be called *focal attribute(s)*. Thus, visibility might be the focal attribute of an environmental intervention (e.g., building and operating a fossil fuel plant) which has many other effects as well. In addition, visibility itself has several faces, each of which might be valued separately (Rowe & Chestnut, 1982). Table 1 lists some visibility-related attributes. Identifying the attributes to specify should be straightforward when a CV study is conducted with a particular environmental intervention in mind. Such identification should be more difficult for a study motivated by a general desire to find out what visibility is worth.

Visibility provides a good example of the difficulties of presenting a good's attributes clearly enough to allow people to understand why they might care about it. Although the general notion of visibility is common enough and some level of visibility is seen most days, few people have thought deliberately about what its possible components are worth (e.g., visual range, haze of different hues, plumes), in isolation or combination. In order to clarify these attributes, CV investigators have invested great efforts in depicting visibility states realistically (e.g., through photos or on-site interviews).

There are two opposed philosophies for presenting the attributes affected by a transaction. Attributes could either be deliberately pointed out, in order to ensure respondents' attention, or deliberately ignored, in order to avoid implicitly heightening their importance. The latter (no-mention) option is possible only when respondents can easily see for themselves what attributes are involved. For example, programs that affect haze intensity may also affect visual range (two attributes of visibility), as well as clothes soiling, building corrosion, and ecosystem stability (to mention three corollary effects that CV studies have sometimes bundled with visibility).² Not mentioning all such attributes means counting on respondents to realize their existence spontaneously. If mentioning an attribute directly does increase its perceived importance, then achieving balance might require mentioning all attributes. With complex interventions, there might be many such mentionable attributes, each requiring the sort of detailed explication described below. One can only hope that the attributes fit together in some natural way, giving respondents a coherent feeling for the overall intervention.

An alternative response to complexity is having respondents concentrate on the focal attributes and ignore the rest. For example, in a study focused on visibility, respondents might be told to disregard how a change in air pollution affected their health risks. However, such selective forgetting may not always be possible. If it is natural to think of an intervention's impacts as a whole, there may be no way to segregate mentally its individual effects (Tolley et al., 1986).

Table 1. Components for Defining Transactions (With Examples from Visibility Valuation)**The Good**

(e.g., visibility)

Substantive Definition**Attribute(s)**

Haze intensity

Visual range

Plume (color)

Light extinction

Context

Natural or built

Judged uniqueness

Associated activities (e.g., hiking, viewing, playing)

Significance (e.g., religious, culture, historical)

Source of Change

Predominantly natural (e.g., vegetation, forest fires, dust storms, humidity)

Predominantly human (e.g., power plant, other factories, field burning, slash burning, motor vehicles)

Formal Definition**Reference and Target Levels**

Magnitude and direction of change

Statistical summary

Representation (mode, richness, organization)

Extent of Change

Geographical

Temporal (existence, direct enjoyment)

Timing of Change**Certainty of Provision****The Value Measure**

(e.g., money, time, discomfort, effort)

Substantive Definition**Attribute(s)**

Leisure, work (for time)

Physical, emotional (for discomfort)

Context

Electric bill, sales tax, income tax, park entry fee, environmental fund (for money)

When convenient, when demanded (for time)

Extended existing periods, additional new periods (for discomfort)

When rested, when exhausted (for effort)

Constituency**Formal Definition****Reference and Target Levels**

Magnitude and direction of change,

Statistical summary

Elicitation (response mode, response format, cues, feedback)

Extent

Frequency

Duration

Timing of Payment**Certainty of Payment***(continued)*

Table 1. (Continued)

The Social Context
Other People Involved
Provider of the good
Others present
Resolution Mechanism
Determining parties
Iterations
Constraints
Other Stakes
Externalities
Precedents
Legitimacy of process

Having identified the relevant attributes, respondents can begin considering why they might care about each. Spontaneously producing those reasons may be quite difficult in the short amount of time offered by most CV elicitation sessions. Because CV goods are not traded, respondents cannot have thought about them previously in the sort of comprehensive, focused way needed to produce a unitary judgment of value. Nor can they have had the sort of trial-and-error experience needed to show them how satisfying it has been to pay various prices. Nor can they seek guidance in the behavior or advice of experienced others. At best, they might be able to consult people who have thought generally about why one might care about such goods. In the case of visibility, such sources might include outdoors people, poets, landscape architects, geographers—and the investigators. Their thoughts might be brought to respondents' attention if respondents cannot get to the thinkers. In any case, helping subjects understand their own best interests regarding novel goods poses a severe challenge for CV investigators (Fischhoff, Slovic & Lichtenstein, 1980).

2.1.2 Context. Evaluations of visibility may depend greatly on where they are measured geographically. For example, dirty air may hardly be noticed over an ugly industrial landscape, but may draw concerned attention over a national landmark. On the other hand, residents of an industrial area may be particularly sensitized to (additional) environmental insults, but may hardly notice the air in areas of uncommon natural beauty. Or, the evaluative difference between 50 and 60 miles of visibility may be modest in most cases, but enormous when a valued sight is 55 miles away from a vantage point.

As a result, any feature of context that affects why people care about visibility is a necessary part of the definition. Stated more generally, the definition of any good must include any conditioning factors that influence the importance of its attributes. This imparts a certain circularity to the design of studies, insofar as some knowledge of people's values is needed in order to characterize adequately the stimuli used to reveal those values. Here, investigators may seek guidance from ex-

perts in what people notice and like in environments (Stokols & Altman, 1986), although without specific transactions or precise evaluations in mind. Table 1 lists some possible contextual features of a good's substantive definition. As elsewhere, those actually used should depend on the particular changes in visibility being contemplated.

If the value of visibility depends on where it is obtained, describing the physical site of the intervention should help clarify the reasons for valuing it. For example, even hearing "Grand Canyon" may remind people of their value for preserving scenic vistas, just as "Denver" or "Mexico City" may evoke the importance of clear air for daily lives.

When there are multiple reasons for valuing an attribute at a site, then each reason must be evoked. For example, the Grand Canyon's air may be evaluated differently when depicted in ways that highlight concern for hikers, as compared to motoring tourists, or RV enthusiasts, or Native Americans, or the site itself (devoid of people). Special efforts may be needed to remind viewers of historical, cultural, religious, or symbolic values that are not readily apparent in physical stimuli. For example, casual viewers cannot be expected to know that 60 miles is the critical visual range for seeing the San Francisco peaks (sacred to the Hopi) from Second Mesa.

When a good includes multiple contexts, each may require separate mention. For example, representing the Eastern United States with a picture of Niagara Falls (e.g., Tolley et al., 1986) may cause visibility in urban areas to be overlooked or even downgraded (because the investigator found only natural beauty worth mentioning).

2.1.3. Source of change. People often care not only about the end states of the goods that they (can) receive, but also about the processes creating those states. With environmental changes, a critical aspect of the process may be whether responsibility lies with people or nature. A scene might be evaluated quite differently if seen as reflecting the power of a natural process or the triumph of an unnatural one. For example, evaluations can shift dramatically when individuals realize that a vivid sunset largely reflects atmospheric pollution (or that a summer haze is due to humidity, rather than smog).

Table 1 lists some natural and human sources of visibility change, each with potentially different effects on evaluations. For example, the smoke from forest fires might be a reminder of the associated destruction, while dust storms causing the same reduction in visual range might seem transitory and benign. Or, smoke from field burning might evoke images of the verdant fields that follow, whereas smoke from slash burning evokes images of charred clear-cuts. As elsewhere, this feature should be specified as precisely as warranted by respondents' information needs. *Natural* versus *human origin* may be enough information for some respondents, whereas others need to know just who or what is responsible.

The adjective *predominantly* (in table 1) reflects the blurred attribution of some changes. For example, the proximal cause of flood damage is typically unusual

precipitation; however, its scope depends on what things people put in the water's path and what they have done to reduce the absorptive capacity of watersheds (Burton, Kates, & White, 1978); similarly, the health risks of logging reflect a combination of natural conditions (e.g., terrain, weather) and industrial practices. Often, current conditions are multiply determined, complicating the attribution of responsibility. A potential advantage of contingent valuation studies is that they typically propose transactions in which conditions are changed by identifiable (usually human) sources.³

Explaining the source of the change represented in a good should be straightforward. It should be well-known to those posing the transaction and readily communicated. As elsewhere, how much needs to be said depends on what respondents assume (by default) without being told. This is an empirical question requiring empirical assessment. People may be surprised to hear that humidity causes haze, or that conifers emit pollutants ("killer trees"), or that white plumes are primarily water vapour.

2.2 Formal definition

2.2.1 Reference and target levels. Any transaction involves a change of state. For a consumer good, that change is typically from having less of the good to having more of it (and to having less of the money used to pay for it). For a public good, like visibility, the change is across worlds having different levels of each affected attribute (and of the resources needed to make that change happen). We use *reference level* for the state obtained if the transaction is not enacted and *target level* for the state obtained if it is. In a transaction offering improved visibility through a pollution abatement program, the good's reference level would be current visibility and its target level some greater visibility. In a transaction that would halt a decline in visibility, the reference level is some reduced visibility (which would occur without the program) and the target level is present visibility (which the program would preserve).

Reference and target levels must be specified for every (potentially) valued attribute affected by a transaction. In some cases, specifying even one of these levels for a single attribute may be quite complex. For example, an intervention may affect visual range differently at different times of the day (e.g., daylight hours, nighttime, rush hour) or year (e.g., summer, rainy season). If the importance of visual range depends on when it occurs, then some statistical summary is needed for both the reference and the target levels. That summary must be faithful to both the data and the respondents, describing accurately and understandably those features of the data needed for evaluation. For example, if people are more sensitive to losses than to gains (Kahneman & Tversky, 1979), then pooling changes in both directions (to get a statistical average) would obscure vital information. Or, a two-mile improvement in mean visual range might seem quite different if achieved by a four-mile improvement half the time (and no change the rest) than if

achieved by equal amounts of a seven-mile improvement and three-mile deterioration (especially if that resulted in more days of "really bad air").

One could simplify matters by describing just the change in a given attribute (e.g., a five-mile improvement) rather than both its reference and target levels. That description would be formally equivalent to the two-state description if respondents knew the reference level so well that they could compute the target level by adding the change to it. Even if respondents did not know the reference level, the change itself would be sufficient if respondents had linear utility functions for all attributes (e.g., if a five-mile improvement in mean annual visual range always had the same values, whether it took one from one to six miles or from 41 to 46 miles).

Once chosen, a summary must be presented. Displaying quantitative statistics helps to ensure that the right things are said, but not necessarily that they are heard. How complex a display people can understand is an empirical question. How people's range of understanding can be expanded is a design question.⁴

One alternative to statistical summaries is showing raw data. For example, respondents could view pictures (or slides) for each day in an average year under reference and target conditions. Here, too, being comprehensive need not ensure being comprehensible. A mass of concrete details could be as bewildering as a small number of abstract ones. An intermediate solution might be to show pictures covering the range of visibility variation with a statistic showing the probability of each outcome. If respondents infer that the attention devoted to an attribute reflects its importance, then each attribute may need comparable detail.

Understanding statistical summaries means understanding the units in which they are expressed. Thus, a unit like *miles of visual range* might be formally precise yet still convey a vague impression psychologically. Direct observation seems like the obvious way to concretize a unit, with pictures (photos or slides) the next best thing. Obviously, it is best if the observation accompanies evaluation (to avoid relying on memory) and involves a variety of scenes and contexts. These should present views that people can actually frequent, rather than, say, aerial shots, which are relatively easy to photograph, but lack a ready translation to actual experiences. Using several photographs for each scene might reduce the effects of transient features (e.g., clouds, trucks, construction). Although all stimuli may merit equal attention, they may not get it. An extensive experimental literature has documented the effects of spatial and temporal organization on how people process multiple stimuli (Atkinson et al., in press; Fitts & Posner, 1965). By judicious design, it may be possible to counterbalance such effects.

At times, the effects of an intervention can be predicted only in terms of a scientific measure that is indirectly related to lay people's experiences. For example, light extinction can be measured accurately and, to some extent, can be predicted from knowledge of environmental interventions. It cannot, however, exactly predict lay estimates of visibility (Malm et al., 1981; Middleton et al., 1985). The effect of an intervention on the distribution of particulate emissions might be predicted even more readily; however, it is even further removed from people's experience.

Good photographs are closer to actual experience, but they may not provide enough information to compute either objective measure. In such cases, there may be a gap between what is measured and what is evaluated.

2.2.2. Extent of change. Transactions involve securing the enjoyment of a good at a specified target level. The extent of that enjoyment must be defined in both space and time.

The temporal limits of a change may be defined in terms of how long the target state will be maintained and how long it will be directly enjoyed. The former definition is needed to assess a good's (a) *existence value*: the benefits of knowing that the target state exists, independent of whether it will ever be experienced directly; and (b) *option value*: the benefits of knowing that the target state could be experienced directly if one so chose (Cummings et al., 1986). The units for expressing how long a target state will be maintained are typically those of time (e.g., days, years), although they could be event-related (e.g., until a referendum revokes the program, until the PUC's composition changes). Defining how long the target state will be enjoyed directly is essential for assessing its actual *use value*: the benefits of experiencing the target state. Its units are typically those of experiences (e.g., park or viewpoint visits), but can also be those of time.

The geographical limits of a change might be as narrow as a single vista (e.g., for a program that eliminated an unsightly plume visible from a valued scenic outlook) or as broad as a continent (e.g., for a program that reduced emissions from all motor vehicles). Those limits, like temporal limits, determine the quantity of change offered in the transaction. If they are left unspecified, then evaluators must guess how much of the target state they are getting. If they are misspecified, then investigators must extrapolate from the transaction they presented to the one that interests them. Unfortunately, there is no necessary relation between the value associated with one areal extent and that associated with another. Value might increase linearly with area, if one felt that every acre counted equally. Or, additional area may be of decreasing marginal utility—with the decline perhaps beginning after some substantial area has been affected.

Because they use familiar units, geographical and temporal extent should be relatively easy to describe. People should have little difficulty understanding, say, that a visibility change will be provided for one year over a radius of 50 miles from an urban center or within sight of the downwind plume from a fossil fuel plant. For geographical extent, a map might clarify just what area is covered, perhaps indicating the sites covered as a reminder of the implied uses. Similarly, the meaning of the temporal extent might be clarified by mentioning the visibility events occurring within that time period (e.g., seasons, rush hours, vacations).

The challenge for respondents is to develop a value estimate for precisely the amount of the good being offered. This seems like a tall order, requiring a fine feeling for value and considerable facility for mental arithmetic (e.g., assigning a 56% higher value for a 50-mile radius than for a 40-mile one—assuming that visibility is valued equally over each square mile). Empirical studies (Tolley et al., 1986)

varying the extent of visibility have sometimes found people failing rudimentary tests of internal consistency.⁵

When extent is left unspecified, respondents must rely on their default assumptions. For geographic extent, they might take their cue from the apparent source of visibility change (if that is provided). For temporal extent, respondents might look to the payment scheme, assuming, for example, that a monthly payment (e.g., through increased utility bills) means buying a month's worth of change or that a lump sum payment means a change in perpetuity. Even when default assumptions are correct, however, respondents may overlook some aspects of extent that would be important if mentioned explicitly (e.g., nighttime visibility).

2.2.3. Timing of change. Any change begins at some point in time. Any waiting may affect a good's value. On the one hand, having to delay gratification may make the good less valuable. On the other hand, being able to savor a good before receiving it may add to its value. Where a payment can be invested with interest and there is a delay between making the payment and receiving the good, respondents should consider the interest lost during that delay.

Where nothing is said, respondents' natural default assumption may be that the good will be received when the payment is made. However, that need not be the case. For example, up-front investment may be needed for capital improvements whose effects will be felt only several years hence, either because construction takes time or because natural systems respond slowly to changes in the burden placed on them.

As elsewhere, respondents may hear technical details but miss their meaning. For example, the idea of lost interest might not occur to them.⁶ Or, respondents may inappropriately decide that changes slated for the future are less likely to take place at all—thereby letting timing affect estimates of certainty (see following section).

2.2.4. Certainty of provision. Agreeing to a proposed transaction involves an exchange of promises: to deliver a good and to make a payment. As a result, two defining aspects of any agreement are the probabilities of each promise being kept. The less likely a good is to be received, the less an offer to provide it should be valued. According to utility theory, people should value a transaction in direct proportion to the probability of receiving the offered good. A more psychological account (Kahneman & Tversky, 1979) predicts a special premium being given to goods obtained with certainty and, conversely, a sizeable discount for any uncertainty. By this account, people are relatively insensitive to differences in intermediate levels of probability (e.g., so that the difference between 95% and 100%, or even between 90% and 95%, is considerably more significant than that between 40% and 45%). People may also treat intermediate probabilities as being somewhat indeterminate, or imprecisely estimated, feeling that once a promise is not

iron-clad, its execution is inherently hard to predict. How people respond to ambiguous probabilities is an area of active current research.

The formal expression of uncertainty is the probability of the promised good being provided. Unless that probability is provided explicitly, respondents must guess at the cumulative effect of whatever physical and social processes might interfere with their receiving the good. In CV experiments, respondents are likely to search for the same kinds of clues they use when estimating the probability that real-life promises will be kept. Those cues may be unintended and misinterpreted.

When provided, probabilities should be given numerically, in order to avoid the ambiguity associated with verbal probability expressions (e.g., “very likely”) (Beyth-Marom, 1982). If uncertainty comes from several sources, then explicit summary probabilities should be provided in order to avoid the errors that arise when people try to combine intuitively the probabilities for sequences of events (Bar Hillel, 1973).

Often, the intended probability will be 1.00. That is, the experimenter wants respondents to treat receipt of the good as a certainty—if the transaction is accepted. Understanding such certain prospects should be easy. Theoretically, it should also be straightforward to extrapolate from respondents’ value for a certain good to the value that they would assign if it were promised only with probability .XX (using either economic or psychological theories). However, saying 1.00 may not ensure that respondents hear 1.00. They may not feel that there can be any certainties when it comes to sociotechnical programs, particularly the novel ones often offered in CV studies. Unless effective assurances can be made, it may be necessary to extrapolate backward from the value assigned by respondents inferring a probability of .XX to the value that they would have assigned had they believed the claimed probability of 1.00.⁷

3. The value measure

In return for the good in a transaction, the evaluator must surrender something, the payment. It represents an overt expression of the value attached to the good. As a matter of analytical convenience, many studies prescribe payments in dollars. However, a payment could also be made in other ways, such as time, or effort, or alternative material goods (Freeman, 1979). As an expository convention, we focus on evaluators who consider surrendering a dollarlike payment in return for receiving a valued good. In paying for something, people might be thought of as trading one good for another. Often, however, one of the goods is a physical thing (e.g., changed air quality) with intrinsic value, while the other good is valued principally as a medium of exchange (e.g., money). From this perspective, an evaluation process begins with the good, to which a payment is attached either by the offerer (i.e., “this is how much I want for it”) or the possible purchaser (i.e., “this is how much I am willing to pay for it”).⁸

Defining a value measure requires as much attention to detail as defining a good—with the choice of definition potentially affecting the values expressed. For example, the chosen value measure can accentuate some attributes of the good and deemphasize others, or it can require a resource (e.g., time, money) that is particularly scarce (or plentiful) for particular evaluators. Once the definition is set, respondents must be able to translate their personal value for the good into those terms. That process can be frustrated by a value measure that is complex, imprecise, or unfamiliar.

3.1 Substantive definition

3.1.1. Attribute(s). Where *payment* connotes money, that becomes the value measure's focal attribute. As mentioned, however, other forms of payment are also possible. Time, like money, is continuous, typically in short supply, and with largely interchangeable units. Much of working life involves investing time in order to secure money, which is then used to secure goods. However, people also spend time directly as an expression of what goods are worth to them. In some cases, time is a substitute for money. For example, people might canvass for a referendum promising improved visibility when they lack funds to contribute to the cause. Either investment means that the cause has competed successfully for their scarce resources.

Other pleasures (or even necessities) of life may also be surrendered to achieve valued ends. For example, people pay with decreased social acceptability in order to support political causes (expending what one social psychologist (Hollander, 1969) has called their "idiosyncrasy credit"). By adjusting their thermostats, they pay with discomfort and risk of illness for increased environmental preservation and national energy independence. People expend mental effort (what might be called quality time) to achieve a wide variety of ends.

As with the good, the first step in defining a value measure is identifying those attributes that depend on the outcome of the transaction. For example, lowering one's thermostat to pay for improved visibility may affect immediate comfort, risk of illness, and length of friends' visits. Unless all relevant attributes are identified in a completed transaction, then investigators will underestimate how much has been paid. Unless the attributes are clearly specified in proposed transactions, then evaluators may underestimate how much they are being asked to pay (leading investigators to overestimate evaluators' willingness to pay). For example, people might lower their thermostat less if told about the effect of this action on their social life. As in defining the good, knowing which attributes are relevant requires an understanding of what people value in the value measure.

With single attribute measures, such as money, the only definitional complication might be which budgetary account provides the payment. People may feel differently, say, about gambling, investment, retirement, rent, and entertainment money. Time, too, is often apportioned into accounts (e.g., leisure time, family

time, work time), with different rules for each and some reluctance to make transfers between them. Such rigidity may seem unreasonable, since it leaves people unwilling to consider some opportunities (e.g., spending leisure time on a particularly lucrative work activity). However, partitioning time and creating monetary accounts may provide benefits by helping people to organize their lives, which are much more important than being able to optimize each specific transaction. Whatever the justification for this accounting, the same size of payment may be valued differently when it comes from different accounts.

Assessing a good's value means observing all possible transactions involving it. When people can substitute alternative payment methods, an investigator risks missing one or more of an individual's scattered payments (e.g., in the case of energy conservation, seeing the time lost through reduced motor vehicle speeds, but not seeing the changed thermostat settings or the lobbying for an oil import fee which are also paid). When people cannot make substitutions, an investigator may miss the value that a good has for people unable to make a focal payment (e.g., because they have money, but not time to pay). In such cases, people may have little of the focal resource overall or just not enough for the somewhat large units in which the good is offered. For example, looking at house prices to reveal a visibility premium will miss the value that visibility has for people unable to buy a house at all; looking at travel costs to reveal the value of places with superior visibility will miss the values of people unable to accumulate enough time for a proper trip.

Generally speaking, using a constrained measure permits those with more of the resource to express higher values—adding a clear political–ethical aspect to the transaction's definition.⁹ Participants should evaluate transactions within their personal resource constraints, lest they agree to spend too much. However, if competing demands are stressed too heavily, the result may be implicit pressure to pay little (i.e., “you can't really afford it”).

Thus, constrained measures force respondents to make hard choices—at the price of giving less weight to resource-poor individuals, and to goods especially valued by such individuals. Like other value measures, constrained ones can reduce the valuations given to goods whose worth cannot be translated into those particular terms, as well as focus respondents' attention on particular attributes of goods. For example, monetary measures may emphasize monetary aspects of a good (e.g., the increased tourism revenue from improved visibility), while deemphasizing its nonmonetary benefits (e.g., aesthetic pleasure from clearer views). Nonmonetary measures (e.g., time spent enjoying scenery) may do just the opposite.

When a CV study is conducted to evaluate a specific policy proposal, the attributes of the value measure should follow directly. Without such direction, the investigator may have latitude in choosing a measure. For example, by offering a variety of alternative payments, investigators might provide policy makers a broad perspective on public involvement with a good; nonmonetary payments might help predict political resistance to a policy measure, which might not be captured

by willingness to pay money. In the absence of policy constraints, the best attributes might be those permitting respondents to express their value for the good most easily and accurately.

3.1.2. Context. A payment must be extracted in some concrete way, generally called the *payment vehicle* in the contingent valuation literature (Cummings et. al., 1986). In principle, all forms of payment involving a common set of attributes should be equivalent. Why should it matter, for example, how the money (or time) goes out? In practice, however, people often do care.

The context may matter where the means of payment has value beyond its ability to achieve the end of securing the good. For example, some people view paying taxes as submission to coercive government oppression; others see it as affirming public commitment to the programs being supported. For some people, preserving a landscape through increased park entry fees seems like a reasonable way to pay for their personal enjoyment of that resource; others find such surcharges irritating, preferring to pay general taxes and let government worry about the details of allocation. In these respects, it matters how a dollar is paid.¹⁰

In a contingent valuation study, the payment context should follow directly from the motivating policy question (e.g., use increased park entry fees if that is the only proposed method to pay for better visibility). In such cases, even though the payment may influence respondents' expressed values, those values are a legitimate representation of how people feel about that specific policy question. However, when any payment context is possible, then the value associated with the one that happens to be used can obscure the good's value. In order to avoid such influences, one could leave the context unspecified. However, doing so need not keep respondents from trying to guess how the payment will be extracted. Thus, the varying dollar values obtained with different payment vehicles may represent legitimate, thoughtful responses whose appropriateness depends on the match between the context used in the study and the context found in the policy question that motivated it.

If context matters when spending money, then how time is spent should matter even more—not to mention “spending” discomfort or effort. Because they come out of a person's flesh, so to speak, these other means of payment have not gone through the neutralizing processes that give dollars the appearance of fungibility.

The payment context may also influence expressed values by suggesting the size of payment. For example, \$100 seems like an enormous increase in a park entrance fee, but only a modest increase in federal income tax. A contribution jar at a delicatessen cash register may suggest a much smaller contribution than a mail solicitation. In a CV experiment, respondents might reasonably take the experimenter's choice of context as a cue to the size of payment expected.

3.1.3. Constituency. Often, people represent more than just themselves when they evaluate transactions. For example, in voting on taxes, parents might be ex-

pected to vote the interests of their children (and even of subsequent generations), in addition to their own. In a CV questionnaire, respondents are often asked to give a value for their entire household. In such cases, respondents should assess and then aggregate the good's benefits to all members of their constituency. One obvious difficulty is knowing constituents' values; there are limits to empathy even for important issues in close-knit social groups. A second difficulty is deciding how to aggregate those inferred values; one may wish to give more weight to some constituents than to others, perhaps out of personal favoritism, perhaps reflecting some general principle (e.g., the welfare of the young—or old—is particularly important). Thus, asking people to respond for a constituency means accepting the possibility that not all constituents' welfare will be valued equally or accurately.

Responses on behalf of a constituency may also be influenced by what constituents might contribute to the payment. That contribution could be in the same form and at the same time as the respondent's payment (e.g., extra taxes coming out of a spouse's income). It could be in the same form but at a different time (e.g., a mother may pay more for goods benefitting her children if she anticipates their future financial support). Or, it could be in quite a different form (e.g., children foregoing treats so that their parents will contribute to UNICEF). How much each constituent *can* pay may vary with the payment measure. How much each actually will pay may vary with the collection method.

3.2. *Formal definition*

3.2.1. Reference and target levels. Prior to the transaction, a participant has some amount of the resource demanded as payment for the good. If the transaction is consummated, the participant will have a different and lesser amount. These are, respectively, the reference and target levels of the payment resource. A good's value is the largest acceptable difference between those levels.

Both target and reference levels need to be specified because, as with the good, the difference between them need not determine a transaction's acceptability. The worth of a given absolute change may depend greatly on how much one already has of that resource. Because those who offer transactions seldom know respondents' reference levels, most transactions, both actual and staged, are stated in terms of the difference between levels. As a result, respondents must compute their personal target levels, complicating the ostensibly simple task of understanding a good's price tag.¹¹

As mentioned, we have adopted the expository convention of focusing on payments from evaluators, meaning that they prefer the reference (or prepayment) level of the payment resource to its target (or postpayment) level. However, in each such transaction, someone receives the payment in return for surrendering the good. Transfers in that direction are, for example, the focus of CV studies that ask respondents how much compensation they demand to accept a degraded environ-

ment (e.g., Bishop & Heberlein, 1986). In a mutually acceptable and satisfactory transaction, both parties should be relatively indifferent between the good and the payment. Nonetheless, a transaction can have a very different character depending on whether one pays or is paid.

One difference is the upper limit of possible demands: what people can ask to receive is unlimited, whereas what they can realistically be asked to give is limited by their assets. In the heat of negotiating a transfer, those maximum demands can take on a life of their own, shaping expectations and anchoring estimates (e.g., so that what one might conceivably get influences perceptions of what one might realistically—or should fairly—get).¹² The direction of payment may also cue particular values. For example, being asked to pay to improve the environment may imply that one ought to take such initiatives. By contrast, being asked what payment would compensate for environmental degradation may suggest that whoever threatens to cause the deterioration has the initiative and is, almost as a courtesy, offering a last-minute chance to change well-set plans. Or, the question may be interpreted as implying a right to be consulted and perhaps even a right to a stable environment.

In experimental tests using contingent markets, direction of payment has proven to be a potent design feature, sometimes effecting a five-fold difference in expressed values (Knetsch & Sinden, 1985). As a result, the price people will pay to receive a good cannot be taken as a substitute for what they will demand to surrender it. Thus, a staged transaction must use the payment direction of whatever actual transaction (e.g., environmental intervention) is envisioned.

For a payment to capture a good's value accurately, some statistical summary is needed for the reference and target levels of every valued attribute of the payment. Ensuring that these (largely quantitative) levels are understood requires special sensitivity to respondents' capacity for mental arithmetic. For example, respondents asked about monthly installments may not realize the annual payment involved, nor may respondents realize the impact on their pocketbooks of a 1% increase in sales tax. Without understanding these implications, respondents will not really know what they are saying. Critical to the evaluation process is the elicitation procedure used. When the price tag on a proposed transaction is fixed, agreeing to it determines only a lower bound for the good's value (i.e., it is worth at least that amount), while refusing the transaction only determines an upper bound. When iterations are possible, a more exact value can be determined, by offering the good at successively higher or lower prices until the upper and lower bounds converge (Coursey et al., 1984). Iterations may also give respondents the opportunity to reflect repeatedly on their values—at the risk of being somewhat tedious and influencing valuations by the choice of starting points and incremental changes (Boyle, Bishop & Welsh, 1985).¹³

An alternative procedure asks respondents directly for the maximum they will pay for the good, leaving the price entirely to the respondent. Although less time-consuming, this procedure is also more demanding, forcing respondents to im-

agine various prices and determine which is just below the lowest at which they would refuse the transaction.¹⁴

Respondents' ability to express their values may also depend on the response mode in which those values are elicited. For example, speaking may be easier, but less precise, than writing. Within any given response mode, some formats may be more complicated than others. For example, comparing pairs of objects may be simpler than ranking a large set. Any comparative response format may be easier than providing specific numbers (e.g., dollar bids): On the other hand, the latter may be easier when the number of alternatives (and, hence, comparisons) becomes very large.¹⁵

In order to help respondents, CV investigators often give them deliberate cues about how to answer, such as starting points in iterative bidding procedures, scale endpoints, information on other respondents' values, or reminders about expenditures on other goods (Cummings et al., 1986). Such cues can produce more consistent values—at the risk of imposing the investigators' beliefs about what responses are appropriate. Especially when precise responses are desired, the risk arises of forcing more out of people than they have to give.

One form of forcing is to discourage expressions of ignorance (e.g., by not offering a “don't know” response option). Insisting that respondents provide values may signal that any nascent feelings should be trusted and expressed. Or, it may leave respondents confused, searching for what to say. An elaborate literature in survey research treats alternative techniques for probing the existence and intensity of respondents' views (Smith, 1985). It shows, for example, that screening individuals for whether they have answers to a question can produce different response distributions than just allowing a “don't know” response.

In staged transactions, another form of forcing is feedback, such as being told that an offered payment is too small. A more obscure message is having a response rejected for methodological reasons. For example, offers to pay \$0 may be interpreted as protests against the procedure, rather than as bonafide valuations. If they are not disqualified outright, then such responses may be met by encouragement to pay something. For example, the claim that “someone else should pay” may be rejected by the investigator (“we'll deal with that later, but for now imagine it is you who are paying”) (e.g., Tolley et al., 1986). Respondents who “don't know” may be pressed “at least to guess.” The impact of such feedback may depend on the social relations between respondent and investigator, which depend, in turn, on personal characteristics that are hard to control or predict.

3.2.2. Extent. Like the good, the payment is bound in time.¹⁶ It may be made once or on multiple occasions and, if so, over varying periods. For example, a sales tax is exacted clearly and repeatedly; an income tax is taken subtly at every withholding period and more blatantly every spring; a utility bill is paid monthly; an endowment may be given all at once. Whenever payment involves more than a single lump sum, the payment schedule and its temporal limits must be specified (e.g.,

monthly for five years). The acceptability of a given total payment may depend on that schedule. For example, it may be easier to pay an amount through monthly installments (like utilities' level-payment plan), or it may be a nuisance to be bothered repeatedly.

3.2.3. Timing of payment. Just as people might have to wait some time after paying for a good before they can receive it, so might they be able to wait before paying for it. When their payment could be invested in the interim, then they are, in effect, paying less than the stated amount. As economic beings, they should think about how much they are willing to pay in terms of the *discounted present value* of a future payment (i.e., the amount that, if invested today, would grow to equal the payment by the time it is needed). As psychological beings, they may find future payments less real (and, hence, less painful) than present payments (Thaler & Shefrin, 1981). People who respond to "buy now/pay later" appeals may discount future payments at a much steeper rate than that justified by their investment opportunities. Promises to spend time in the future (e.g., writing an article) are often made much more casually than promises to perform equivalent tasks sooner. On the other hand, people may prefer doing a task immediately, rather than having it hang over their heads.

Timing may also affect people's ability to make payments. People may expect to have different amounts of time or money to spend, or a different capacity for discomfort, at different points of time. In some cases, intertemporal transfers might reduce these differences (e.g., doing extra work now, so as to free up time later). Such transfers should be easiest with monetary payments (e.g., borrowing on the strength of future income in order to make earlier payments). These possibilities, however, must be tempered by the (financial and psychological) transaction costs of borrowing.

Determining when the payment is required should be relatively straightforward for both actual and staged transactions. So should describing that date. Rather more difficult is ensuring that people incorporate that timing fully in their evaluations. They may not understand the principles of timing (e.g., the possibilities and problems of intertemporal transfers). Or, they may not be able to apply those principles under the conditions of an evaluation exercise (e.g., performing the mental arithmetic needed to compute discount rates).

3.2.4. Certainty of payment. Unless a payment is extracted immediately, there may be some uncertainty about whether it will be extracted at all. Any uncertainty should make the payment seem less onerous. As with uncertain goods, respondents' value for a payment may be proportionate to its probability, or merely monotonic with it. Unless provided quantitative probabilities of payment, people are left to read other details of an experiment for cues.¹⁷

4. Social context

All human actions occur in a social context. Even when acting in private, people bear with them the values they have acquired from their society and the need to face society as someone who has undertaken (or declined) a particular action. That social context both constrains people's actions and gives them meaning. In order to understand their own values, people need to see the social context of their decisions and the world that they are creating for themselves through their decisions.

For marketplace transactions, much of the social context is readily apparent. Various mechanisms determine what goods are offered, how they are described, what can be done if the description proves inaccurate, how strong the bargaining positions are, and so on. Other mechanisms affect how others might perceive one's response to a proposed transaction (and whether one cares), what precedents a decision might set, whether it seems consistent with preceding decisions (and whether one cares), and so on.

Whichever contextual features are important must be noticed and understood if people are to evaluate transactions appropriately. The risk of missing or misinterpreting vital features seems particularly great in novel settings (e.g., staged transactions) simply because it is hard to see through complex ramifications in any short period of time. However, oversights can also arise, or even be encouraged, in familiar settings. For example, merchants who recognize this social construction of reality often attempt to evoke social cues that favor acceptance of their offers and to suppress cues that do not. That may mean isolating people from social contact, but flooding them with social stimuli (e.g., ads showing attractive people accepting a transaction). Or, it may mean placing people in coercive social settings (e.g., high-pressure, used-car salespeople; Tupperware parties).

In staged transactions, the entire social context must somehow be provided. For the sake of apparent simplicity, investigators might be tempted to skip these details, hoping that respondents will concentrate on the good's worth to them personally and assume that the transaction is, in all other ways, ordinary. However, respondents may assume these missing details, just as they may assume other aspects of a transaction's definition. As elsewhere, it seems best to specify the intended context explicitly, unless it fortuitously matches respondents' default assumptions. In that specification, three basic questions need to be answered: (a) Who else is in a position to influence the evaluation process? (b) What mechanism determines whether the transaction will be consummated? and (c) What else is at stake, beyond the possible exchange of a payment for a good?

4.1. Other people involved

In the simplest of transactions, a single individual acquires a good from nature for private consumption in return for a personal payment, such as the time needed to

visit a scenic sight or the labor needed to reclaim a garden plot from weeds. Individuals evaluating such transactions need not care about what nature thinks,¹⁸ nor worry about deliberate interference in their decision-making process. Nature will not hamper their attempts to understand the terms of the transaction, nor pressure them to make particular choices, nor reprove (or approve) their choice. Evaluators may still choose suboptimally, but will have only themselves to blame. When other people are involved, evaluation becomes more complicated. Those others may be divided into those providing the good and everyone else.

4.1.1. Provider of the good. When a good is offered by an individual (or group or institution), several social dimensions must be considered. One is the possibility of the offerer manipulating the evaluation in order to achieve some end. That end may be securing the highest possible payment or it may be ensuring that some transaction is completed in a timely manner (with payment above some minimal level). Manipulation might involve outright misrepresentation or, more mildly, directed representation, highlighting certain features and downplaying others. In real life, evaluators must consider offerers' motivation and opportunities for manipulation (as well, perhaps, as their own).

Those opportunities will depend, in part, on the social and legal contracts among the actors. In marketplace decisions, those contracts reflect such ambient considerations as consumer protection laws, avenues for suit in case of fraud, voluntary norms of business behavior, and concern for reputation. There may also be transaction-specific agreements, such as explicit promises and an implicit conferral of informed consent to any associated risks.

Although contracts, both implicit and explicit, can reduce the opportunities for unfair manipulation, they cannot eliminate the manipulation that comes from legitimate disparities in the parties' bargaining positions. When analyzing a transaction, it is sensible to ask not only "What is this good worth to me?" but also "What is it worth to those offering it?" If the good is dear to the offerer, then the probability increases of having to pay full value for it. It is also sensible to ask "What do the offerers know about my position—and how well can they exploit that information?" Better negotiating skills may compensate somewhat for inferior bargaining positions. Presumably, people will be willing to pay less than their full value for a good if they can get away with it.

At times, goods are offered by agents acting on the owner's behalf. Dealing with these agents raises additional questions for evaluators. What are the agents' reward systems, and how do these differ from the owner's? Will the owner and the agents act together to manipulate me? A familiar arena for such complex social interactions is dealing with a salesperson, the credit manager, and the boss at a car dealership.

In staged transactions, the experimenter may either provide the good or represent whoever is interested in the results of the transaction. In some cases, this role may be clear; in others, participants may be left guessing. One likely clue in either case is the experimenter's apparent desire to have some transaction be completed.

Perceived pressure to provide some nonzero evaluation for the good might produce compliance or resistance (Brehm, 1969).¹⁹ Clues to a study's sponsor may suggest what values are wanted and expected, as well as the strength of the respective bargaining positions.

The novelty of a contingent market should increase participants' search for clues to its social dynamics and possibly increase the chances of reading in things that are not there—or at least were not intended to be. Those dynamics may not even be understood thoroughly by those who stage such markets. The history of psychological experimentation has been described as the continual discovery of unintended influences of staging on behavior (McGuire, 1969). Thus, despite a commitment to creating satisfactory transactions (in the sense defined in the second paragraph of this article), investigators may inadvertently manipulate respondents' expressed values. Under the rubric of *demand characteristics*, psychology has studied extensively the power of the cues and pressures that ostensibly neutral experiments can create, especially when the stakes of the experimental task itself are small relative to the importance of having a satisfactory relationship with the experimenter (Rosenthal & Rosnow, 1969). The research also shows how imaginative people can be at discerning cues (both intended and unintended, both correct and incorrect) in even the simplest of settings.

4.1.2. Others present. Normally, whoever offers a transaction also observes to it. So might friends, relatives, bystanders, and, in the case of some staged transactions, other respondents. Whoever observes a transaction may also influence its outcome. To some extent, these influences are legitimate. People care what others think about their decisions. They might be chagrined later to have made a decision without considering what their spouses, children, parents, or drinking buddies might think and say. Of course, they cannot entirely ignore the ways in which a lifetime of interaction with others has shaped their values. However, people who are out of sight may also tend to be out of mind, whereas those who are present may not be in the evaluators' natural reference group. Although those outsiders' opinions might enrich the evaluation process by suggesting new perspectives, they may also distort it by imposing a transitory set of concerns. Whatever their apparent identity, observers by their very presence may accentuate socially defined values (e.g., altruism, shared experiences, collective action), relative to hedonic personal values.

The power of such influences has been a (or, perhaps, the) major topic of research in social psychology—in part, because those influences are a (or, perhaps, the) major topic of social life. Clearly, people have some ability to fend off others' opinions. However, the conditions for exercising that ability are complex. For example, it may be easier to manage the opinions of unfamiliar others (because their opinions matter less), but also more difficult (because it is harder to put their opinions into perspective). Resistance should be easiest for people who understand the transaction well enough to know their own opinions—and care enough to want to express them.

In staged transactions, respondents might reasonably assume that everything is there for a reason. If a group is present, respondents would wonder why they have been brought together. The reason should be made clear to them. Unless the intent is to manipulate, groups should be composed to offer alternative perspectives, and conducted to share those opinions uncoercively. Making evaluations privately should reduce any pressure to make particular evaluations, or any evaluation at all.

4.2. Resolution mechanism

Transactions represent agreements between individuals. Whether they are actual or staged, some mechanism is needed to determine which possible agreements will be consummated. The resolution mechanism may be as simple as a consumer deciding (independently) to pay a fixed amount for a fixed good (e.g., to buy a newspaper) or as complex as an iterative bidding procedure in which potential buyers compete for a good that is continually redesigned to make it more attractive. A transaction may even be resolved collectively, as in a referendum on increasing taxes to pay for environmental improvements.

In order to express their values accurately, people must understand the resolution mechanism. Doing so means mastering three essential features of that mechanism: (a) which parties determine its outcome; (b) how many iterations it provides for considering the transaction proposal; and (c) how it constrains possible agreements.

4.2.1. Determining parties. In transactions with nature, the individual decision maker is sovereign in deciding whether the transaction will be consummated and may have considerable flexibility in determining what transactions are possible. When a transaction requires agreement between all providers of the good and of the payment, individual buyers typically have less say in shaping alternative options and thus in determining whether a transaction will be consummated. Even when the terms are mutually agreeable, completing a transaction means coordinating the actions of the individuals involved, a process that may fail for reasons of personality or communication.

The next level of complexity has several individuals offering the good together. Although still a transaction between two parties, dealing with a collective means recognizing its internal resolution mechanism. For example, it may seem possible to play members of the offering team off against one another, thereby reducing one's payment. On the other hand, well-orchestrated teams may extract higher payments by manipulating their internal roles (e.g., the old nice guy-bad guy routine, the cast at an auto dealership). Doubts about the offerers' ability to agree on a transaction may discourage taking a transaction seriously. Being part of a buyers' collective reverses these considerations.

Alternatively, each of several parties may offer an individual transaction, only

one of which can be chosen. Although each such transaction creates a separate dyadic relationship, to be resolved on its own, the psychology of the evaluation process may be quite different. On the negative side, considering several transactions may reduce the attention paid to each. On the positive side, seeing different possibilities may enhance understanding of each. Any competition among transactions may affect their offering prices and evaluators' offered prices. In such situations, the resolution mechanism must contain a procedure for sequencing the evaluations. Does the evaluator deal with all until agreement is reached on one? Can each be evaluated and then the best selected? Or, must each be considered in turn and irrevocably accepted or rejected?²⁰

When independent evaluators compete for a single good, these roles are reversed. Each individual has a dyadic relation with the offerer and some procedure is needed to sequence the solicitation and resolution of their offers.²¹ These are, in effect, bidding procedures, conferring goods on those promising the most in some chosen currency. For staged transactions, quite sophisticated procedures have been developed to encourage competitors to express their values accurately (Groves & Ledyard, 1977; Vickrey, 1976). These incentive schemes reward individuals for bidding what a thing is really worth to them, assuming that they understand the rules well enough to identify that optimal response.

The outcome of competitive bidding depends on participants' ability to pay as well as on their willingness to pay. Therefore, the values of individuals having little of the focal resource (e.g., time, money) will have little effect on eventual transaction prices.²² That is, even if a bidding scheme can help individuals compare alternative evaluations, it cannot reduce to comparable terms the evaluations of different individuals.

In order to avoid losing information on the values of those who cannot, or will not, bid very much, investigators can solicit bids from everyone, not just those likely to win. Doing so assumes that resource-poor individuals will take the bidding seriously, even though their likelihood of losing means that their bid is inconsequential (to them, although not to the investigator). Differences in some resources (e.g., money) might be reduced by giving participants resources to use in the bidding—hoping that they will not view these dedicated resources as simply an addition to their overall stock. One further threat to bidding procedures is having them take on their own dynamic, confounding the value of the good with the value of winning the bidding.

Quite different resolution mechanisms are needed when the relationships among evaluators are cooperative, rather than competitive (Dawes, in press). When securing a good requires the concerted efforts of multiple individuals, the resolution mechanism might be interactive, such as a community organizing itself to curb backyard burning. Or, it might be noninteractive, such as a community voting on taxes for air pollution controls. Somewhere in the middle might be a campaign to raise funds for an environmental education program. As elsewhere, any public discussion before the resolution mechanism is exercised may both inform and pressure participants.

More generally, whenever a transaction involves other people, there may be (or at least may appear to be) opportunities to manipulate those people to personal advantage. For example, one classic strategem is offering less than a good is worth to oneself, hoping that others will pick up the burden. Much is known about the theoretical possibilities for gaming in various pricing systems (Cummings et al., 1986). Whether those possibilities are realized depends on whether participants recognize their existence, are willing to exploit them, and succeed in doing so.

4.2.2. Iterations. Some transactions are offered in a fixed form that evaluators can either accept or reject. Goods on supermarket shelves, ballot referenda, and entrance fees for national parks are such take-it-or-leave-it offers. A more dynamic resolution mechanism has one party offer the good, the second party offer a payment (thereby completing the transaction's definition), and the first party resolve the proceeding by accepting or rejecting that payment.

Alternatively, once offerers have seen the bids (and perhaps the bidders), they may adjust their minimum selling price, the good, or some aspect of the transaction's social context. These roles could also be reversed. Whichever party has the initiative may be able to define most aspects of the transaction. For example, initiators may make strategic offers in order to see how small a good they can get accepted or how large a payment they can extract—knowing that they can replace inappropriate offers and hoping not to increase the other party's strategizing.

When iterations are possible, a resolution mechanism is needed to bring them to an end. They could continue until all parties are satisfied, until one essential party resigns, or until some predetermined limit is reached (e.g., time, number of iterations). As elsewhere, whether iterations converge on participants' actual values depends on the fairness and comprehensibility of the resolution mechanism. Thus, while participants might use the iterations to reflect on the good's value, their attention may be focused on understanding the evaluation process better. For example, the process of negotiating for a new car may prove less enlightening about the car than about the process itself and about its participants. The outcome of such a process may depend heavily on design details, such as who opens, what the response time is, and how experienced the parties are (Bazerman & Lewicki, 1983).

As elsewhere, in staged transactions, these details should be dictated by the policy question motivating the research. For example, multiple iterations might imitate best the give and take of an interactive political process. By contrast, iterations may have irrelevant effects on participants who would normally evaluate the good in private.

4.2.3. Constraints. In actual transactions, not everything is possible. People cannot offer goods they do not control, pay more than they have, nor ensure attention to all their comments and proposals. Laws or social norms may prohibit yet other transactions, such as grossly unsatisfactory ones (e.g., where goods are misrepresented) or ones affecting third parties adversely (discussed below in section

4.3.1.). As a result, a defining characteristic of the resolution mechanism is how it constrains transactions. These constraints may be imposed through formal review (e.g., by lawyers or government inspectors) or informal checks (e.g., Has the heat of the bidding led to evaluators accepting unmanageable payments?).

Although these constraints might be intuited in a familiar setting, in a staged transaction they need to be described. Respondents need to know, for example, who vouches for a good's description, who has screened proposals for violations of legal rights and political sensibilities, what happens if the deal comes unglued, and how far they can challenge the terms of a proposal. Without that information, they may diffuse their energies in needless worrying, erroneously assume customary protections, or mistrust the entire procedure.

An implicit constraint in many contingent valuation studies is that all respondents must participate, seeing it through to some resolution. Refusals are often dismissed as *protest responses*. Voting with one's feet represents the equivalent behavior in actual transactions. Such protests might be ignored, like the opinions of those who sit out an election. Or, they might be taken seriously, as when political campaigns are shaped to arouse nonvoters. The protesters themselves may also wish to be heard or to be ignored depending on whether they have defiantly rejected the procedure or have just been insufficiently involved to think it through. CV studies often seem like social settings that strongly encourage compliance. If respondents care enough not just to go along, then protest responses may be harbingers of political protests against the policy motivating the study.

4.3. Other stakes

People shape their world through their transactions. Those transactions establish and abrogate rights, engender experiences that shape tastes, establish precedents for personal behavior, and determine relationships among people (which are valued both intrinsically and for their role in future transactions). To the extent that they define themselves by their actions, thoughtful actors may see more at stake than just the exchange of a good for a payment.

4.3.1. Externalities. A transaction affects individuals beyond those directly involved in its resolution. In economic parlance, these bystander effects are *externalized* benefits and costs.

There are many reasons why those with standing in a transaction might care about how it affects others. These include altruism (e.g., "Making my kids' environment safer is a high priority for me"), enlightened self-interest (e.g., "If this measure makes other people safer as well, it will help reduce my insurance costs"), status seeking (e.g., "When everyone shares a good, having it myself is worth less"), community feeling (e.g., "I like the thought of others sharing this experience"), exclusivity (e.g., "I like enjoying this alone"), or perceived injustice (e.g., "It makes me mad that they're enjoying it when I'm making it happen"). These emotions may

not all seem nice, even to those who have them. However, they can affect evaluations.

As with other features, evaluators' understanding of externalized effects should depend on a transaction's complexity and familiarity. Each overlooked externality means some overlooked value. Any overlooked externality means emphasizing personal effects to the detriment of social effects. Externalities can evoke such varied emotions that evaluators need particular help here to avoid incomplete or ephemeral responses. In staged transactions, that help might include time to ruminate, peers to consult, or prior analysis of possible perspectives.

4.3.2. *Precedents.* Just as transactions can affect other people, so can the aspects of participants' lives that are not directly at stake—insofar as they set liberating or constraining precedents. Some precedents are on a personal, psychological level. What people do shapes their mental self-image, even when the consequences of their actions are unanticipated (e.g., “I am now someone who has been in (or caused) a serious traffic accident”). The value of setting precedents may be seen in people's special pride in bringing about the first transaction effecting a positive change (e.g., a reduction in risk or improvement in visibility) and special reluctance to accede to a novel bad (e.g., allowing environmental degradation from a new source).

Other precedents are set interpersonally. Accepting one transaction may be construed as constituting consent to other, related transactions (Gibson, 1983).²³ Such claims may be primarily rhetorical (e.g., “You brought it for me once”). Or, they may involve a real surrender of rights (e.g., paying to prevent one environmental threat may weaken claims of the right to a clean environment or to compensation for degradation). When the parties have a continuing relationship (e.g., consumers judging the products offered by merchants, voters judging the referenda created by politicians), each evaluation sets precedents for its successors. Indeed, the response to one offer may be designed to affect future offers on other (seemingly unrelated) topics. Similarly, the expectation of future relations may help keep offers and responses honest and reasonable.²⁴

In the novel social setting of a staged transaction, participants might be able to guess how their response will affect their self-image. The legal and social ramifications of their choices should be much more difficult to predict. For example, respondents might assume, if they agree to pay for one environmental amenity, that they will be expected—or will earn the right—to pay separately for others. In that case, a single study with modest stakes becomes the harbinger of a major change in environmental management. Or, respondents may treat the study as a hypothetical exercise, designed to satisfy some investigator's curiosity, with little else at stake.

4.3.3. *Legitimacy of process.* People often care how decisions are made. They want not just to get a good deal but also to deal and be dealt with fairly. Other things being equal, they should be more willing to accept transactions embedded

in legitimate social processes (Furby, 1986; Tribe, 1972). One obvious threat to legitimacy is anything that makes a transaction less satisfactory. For example, the same exchange should seem less attractive if either party negotiates from a position of inferiority, hence feels coerced to accept a particular transaction, or even just any transaction.²⁵ Unsatisfactory transactions may also come from denial of necessary information, of the opportunity to shape options, or of the ability to evaluate transactions. In such cases, the opportunity to evaluate transactions may seem more like exploitation than enfranchisement, and perhaps even like an attempt to make evaluators responsible for the outcome of the transaction without actually giving them effective control.²⁶ Seeing other evaluators in similarly disadvantaged positions may make matters worse—because the rights of an entire public are being denied—or better—by making those arrangements seem more widely accepted.

Potential transactions must come from somewhere. They may reflect the chance meeting of consumer and producer in the marketplace. They may be placed on the ballot by legislators or citizens' petitions. They may be generated by individuals contemplating how to invest their leisure time. People may feel—and look—better accepting transactions from legitimate sources. Such decisions are easier to justify (to oneself and to others). There may even be some satisfaction associated with supporting the institution that created the transactions.

Legitimacy may be general or specific. For example, one may feel comfortable with *any* transaction between informed, consenting parties in a free marketplace bolstered by appropriate legal protections; or, one may want to exclude those parties from proposing transactions involving parenthood (e.g., surrogate mothers, free market adoption). Cost-benefit analysis may seem like an acceptable process for evaluating transactions that a society might undertake—except those including human lives (which should not be monetized). One may favor referenda as a way to involve and respect the public—except for those tough decisions that legislators should resolve, rather than deferring them to an electorate ill-equipped to identify (and vote) its own best interests. One may endorse actual referenda—because the associated political campaigns mobilize and educate the public—or experimental referenda—because they focus people's attention in the presence of a well-informed experimenter. In the end, people may accept a relatively unattractive transaction simply because of how they were asked. Or, they may reject an attractive one because it does not seem like a legitimate way of doing business.

5. Conclusions

Transactions are complex social and psychological phenomena. Each has many features that participants may consider in deciding whether to accept it. Knowing how people have interpreted these features is essential in order to know what transaction those people thought they were facing. Specifying all relevant features, and ensuring that they have been understood, is essential to staging transactions.

Unless a feature is specified explicitly (and comprehensively), evaluators must guess its value and, hence, what the offer really is. If they guess wrong, then they risk misrepresenting their values.

The preceding analysis defines the set of potentially relevant features. Those identifying the good and the payment were divided into *substantive* and *formal*. Substantive features represent reasons why an evaluator might care about a transaction. Their importance is a matter of personal preference. Thus, insensitivity to manipulations of these features could mean misunderstanding the task or just not caring.

Formal features specify how much of the good is being offered, or how much of the payment is being demanded. If the good or payment has any value at all, then evaluations should vary with the quantity involved. Insensitivity to changes in quantity indicates either a failure of the presentation (to make the quantity clear) or a failure of the evaluator (to have sufficiently articulated preferences).

Ensuring understanding is the responsibility of those who pose transactions (actual or staged)—if they wish to claim that participants were sufficiently well-informed to act in their own best interests. A straightforward way to seek understanding is to describe all features of a transaction. This way may not be effective, however, if the quantity and quality of information defy comprehension. One way to simplify descriptions is to omit features that match participants' default assumptions—what they believe in the absence of evidence to the contrary. The presentation can then focus on features that would not or could not be inferred—provided that special attention is not interpreted as implying that the mentioned features are particularly important.

The problems people have in comprehending transactions and the strategies which can minimize those problems are both empirical considerations. The test of success is enabling evaluators to identify courses of action in their own best interests. Passing the test means that the evaluator both understands the content of transactions and can determine its personal relevance. Such determination may require both time and help. Where those resources are lacking, expressed evaluations must be suspect.

Understanding the meaning of alternative formulations is essential for those who pose transactions, as well as for those who evaluate them. The proper formulation is the one that most closely matches the research or policy question being investigated. Determining that formulation will be hard for those motivated only by a vague desire to know what something is worth. A sharper understanding of what questions to ask is as important to the science of measuring values as is a sharper understanding of how to ask them.

Notes

1. Market researchers recognize these limits when they use *focus group* discussions to generate ideas that people might otherwise ignore or think of only after the interview was over.

2. Every attribute of a good poses specification and explication problems analogous to those associated with visibility attributes (and discussed in the text). Increasing the number of nonvisibility attributes increases the complexity of the judgments that respondents must make, and also complicates the investigator's task of discerning the value attributed to visibility within the overall evaluation of the good. In extreme cases, CV studies could resemble hedonic pricing studies, forcing investigators to detect the effect of one modest attribute (e.g., visibility) in a large transaction (e.g., house purchase) having many other attributes (e.g., purchase prices, mortgage rates, school districts), each of which is understood by respondents to some degree (Cox, 1985; Fischhoff & Cox, 1985).

According to this framework, defining a good effectively cannot eliminate the need to evaluate complex goods. It can, however, ensure that the goods that are evaluated are ones that really matter, namely those relevant to the motivating policy (or theoretical) question.

3. The evaluator will have some social relation to any human source of change, if only one of powerlessness. That relationship could affect the deal being struck (e.g., by determining the evaluator's perceived right to compensation or negotiating strength). These issues are considered in the discussion of the social context of a transaction. Here, we treat the source's impact on the good's value.

4. This issue is addressed by Tufte (1983), Mosteller & Tukey (1976), Peterson & Beach (1967), Wright (1987), and others.

5. More rigorous tests of consistency are often hard to apply because people could have steeply declining marginal utility for additional amounts of utility beyond any break point. Thus, suspicions could be raised, but not confirmed, if the value of one year's improved visibility were only slightly larger than the value for ten days.

6. Indeed, the general notion of opportunity costs seems to have limited psychological reality (Thaler, 1981).

7. Conceivably, one might induce respondents to treat the good as a certainty by promising to refund their payments if the good is not delivered.

8. A reversal of these customary roles occurs with design-to-price contracts, in which an evaluator tells vendors, "This is what I am willing to pay. These are the attributes that I will use in evaluating candidate goods. What can you offer me?"

9. Resource constraint issues can be avoided with unconstrained measures such as ratings of desirability or satisfaction. However, one then cannot compare or aggregate responses across respondents.

10. In addition to its intrinsic value, the payment vehicle can be a potent cue to the kind of transaction being proposed. For example, does it exact payments from (those whom respondents consider) the right people? Does it set important precedents for other transactions? These topics are considered in the section analyzing transactions' social context.

11. By focusing people on changes in level rather than on their personal reference and target states, presenting prices also increases people's susceptibility to the framing effects of prospect theory (Kahneman & Tversky, 1979).

12. In principle, this asymmetry also applies to receiving and forfeiting the good. In practice, though, there may be a physical limit to how much of a good could be produced or consumed. Those constraints are much less acute with fungible payments, which people can absorb in seemingly limitless amounts.

13. For example, evaluators asked to pay \$1.50 after first agreeing to pay \$1.00, and then \$2.00 if they agree again, might feel safe in assuming that whoever offers the transaction does not expect the bidding to go to \$100.

14. These and alternative procedures are discussed in later sections in the context of *resolution mechanisms* for determining whether a transaction will be conducted.

15. Coombs' (1964) *Theory of Data* provides a compendium of alternative elicitation approaches, tailored to the kind of information that people can best supply.

16. The payment schedule may also provide a source of cues regarding the good's temporal extent, if the latter has been left unspecified. In some cases, evaluators may assume that the good will continue only as long as the payments continue. In others, it may seem as though a single payment should ensure a lasting good.

17. As discussed in the section on social context, the fate of some transactions depends on the joint action of several evaluators (e.g., voting on a referendum). Until the actions of these others are known, some uncertainty surrounds the fate of the transaction. That uncertainty is in addition to any uncertainty about whether the good will be provided and whether the evaluators will have to pay—even if the transaction is approved.

18. Except perhaps for deists.

19. Participants might be quite surprised, however, to find that \$0 responses are discarded, while very high evaluations are adjusted downward to (what investigators deem) more reasonable levels (e.g., Tolley et al., 1986).

20. Roughly the same alternative sequencing options are relevant when a subset of the offers can be accepted.

21. Any individual for whom the good had no value would simply drop from the competition.

22. Unless, for some reason, the good appeals primarily to less prosperous bidders.

23. In cases where the precedent is legally binding, the initial transaction should have been defined more broadly. For example, the transaction in joining a record club is not paying a dollar for three (3) records, but agreeing to buy nine (9) records over a year.

24. An alternative conceptualization is to view each such evaluation as one round in an iterative evaluation process involving a simple transaction with diverse and sequentially revealed goods.

25. Even those doing the coercing might be pleased with the deals that they get, but still unhappy with how they got them.

26. The opportunity to choose among competing phone services seems to have been a false freedom of this sort for many people. They found themselves neither willing nor able to develop the street smarts needed to unravel the conflicting claims of competing providers.

References

- Atkinson et al. S.S. *Stevens' Handbook of Experimental Psychology*. New York; Wiley, in press.
- Bar Hillel, M. The Subjective Probability of Compound Events. *Organizational Behavior and Human Performances* (Vol. 9, 1973), pp 396-406.
- Bazerman, M.H. & Lewicki, R.J., eds. *Negotiating in Organizations*. Beverly Hills, CA: Sage, 1983.
- Beyth-Marom, R. How Probable is 'Probable'? *Journal of Forecasting* (Vol. 1, 1982), pp 257-269.
- Bishop, R.C. & Heberlein, T.A. Does Contingent Valuation Work? In: R.G. Cummings, D.S. Brookshire, and W.D. Schulze, eds., *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*. Totowa, NJ: Rowman & Allanheld, 1986, pp 123-147.
- Brehm, J. *A Theory of Reactance*. New York: Academic Press, 1969.
- Boyle, K.J., Bishop, R.C., & Welsh, M.P. Starting Point Bias in Contingent Valuation Bidding Games. *Land Economics* (Vol. 61, 1969), pp 188-194.
- Burton, I., Kates, R.W., & White, C.F. *The Environment as Hazard*. New York: Oxford University Press, 1978.
- Coombs, C.H. *A Theory of Data*. New York: Wiley, 1964.
- Coursey, D.L., et al. *Experimental Methods for Assessing Environmental Benefits: Volume II*. Draft report, U.S. Environmental Protection Agency, 1984.
- Coursey, D.L. & Schulze, W.D. The Application of Experimental Economics to the Contingent Valuation of Public Goods. *Public Choice* (Vol. 49, No. 1, 1986), pp 47-68.
- Cox, L.A., Jr. Theory of Regulatory Benefits Assessment: Econometric and Expressed Preference Approaches. In: J.D. Bentkover, V.T. Covelio, and J. Mumpower, eds., *Benefits Assessment*. Dordrecht: D. Reidel, 1985.
- Cummings, R.G., Brookshire, D.S., & Schulze, W.D., eds. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*. Totowa, NJ: Rowman & Allanheld, 1986.
- Dawes, R.M. *Rational Choice in an Uncertain World*. New York: Holt, Rinehart & Winston, in press.

- Fischhoff, B. & Cox, L.A., Jr. Conceptual Framework for Regulatory Benefits Assessment. In: J.D. Bentkover, V.T. Covello, and J. Mumpower, eds., *Benefits Assessment*. Dordrecht: D. Reidel, 1985.
- Fischhoff, B. & Furby, L. *A Review and Critique of Tolley, Randall et al., "Establishing and Valuing the Effects of Improved Visibility in the Eastern United States."* ERI Technical Report 87-6. Eugene, OR: Eugene Research Institute, 1987.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. Measuring Labile Values. In: T.S. Wallsten, ed., *Cognitive Processes in Judgment and Choice Behavior*. Hillsdale, NJ: Erlbaum, 1980.
- Fitts, P. & Posner, M. *Human Performance*. Belmont, CA: Brooks-Cole, 1965.
- Freeman, A.M. *The Benefits of Environmental Improvement: Theory and Practice*. Baltimore: John Hopkins Press, 1979.
- Furby, L. Psychological Studies of Justice. In: R.C. Cohen, ed., *Justice: Views from the Social Sciences*. New York: Plenum, 1986.
- Gibson, M., ed., *Risk Consent and Air*. Totowa, NJ: Rowman & Allenheld, 1983.
- Groves, T. & Ledyard, J. Optimal Allocation of Public Goods: A solution to the Free Rider Problem. *Econometrica* (Vol. 45, 1977), pp 783-809.
- Jones-Lee, M.W. *The Value of Life: An Economic Analysis*. Chicago: University of Chicago Press, 1976.
- Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* (Vol. 47, 1979), pp 263-291.
- Knetsch, J.L. & Sinden, J.A. Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value. *Quarterly Journal of Economics* (Vol. 100, 1984), pp 507-521.
- Maler, J. A Note on the Use of Property Values in Estimating Marginal Willingness to Pay for Environmental Quality. *Journal of Environmental Economics and Management* (Vol. 4, 1977), pp 355-369.
- Malm, W., et al. Human Perception of Visual Air Quality (Uniform Haze). *Atmospheric Environment* (Vol. 15, 1981), pp 1875-1890.
- McGuire, W.J. Suspiciousness of Experimenters' Intent. In: R. Rosenthal and R.L. Rosnow, eds., *Artifacts in Behavioral Research*. New York: Academic Press, 1969.
- Middleton, P., Stewart, T.R., & Leary, J. On the Use of Human Judgment and Physical/Chemical Measurements in Visual Air Quality Management. *Journal of the Air Pollution Control Association* (Vol. 35, No. 1, 1985), pp 11-18.
- Mitchell, R.C. & Carson, R.T. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. In press.
- Mosteller, F.A. & Tukey, J. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1976.
- National Research Council. *Survey Measures of Subjective Phenomena*. Washington, DC: Author, 1982.
- Peterson, C.R. & Beach, L.R. Man as an Intuitive Statistician. *Psychological Bulletin* (Vol. 68, 1967), pp 29-46.
- Rokeach, M. *The Nature of Human Values*. New York: Free Press, 1973.
- Rosenthal, R. & Rosnow, R.L., eds. *Artifacts in Behavioral Research*. New York: Academic Press, 1969.
- Rowe, R.D. & Chestnut, L.G. *The Value of Visibility: Theory and Application*. Cambridge: Abt Books, 1982.
- Smith, T.W. Non Attitudes: A Review and Evaluation. In: C.F. Turner & E. Martin, eds., *Surveying Subjective Phenomena*. Beverly Hills, CA: Sage, 1985.
- Smith, V.K. To Keep or Toss the Contingent Valuation Method. In: R.G. Cummings, D.S. Brookshire, & W.D. Schulze, eds., *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*. Totowa, NJ: Rowman & Allanheld, 1986, pp 162-179.
- Smith, V.K. & Desvousges, W.H. *Measuring Water Quality Benefits*. Boston: Kluwer-Nijhoff Publishing, 1986.
- Stokols, D. & Altman, I., eds., *Handbook of Environmental Psychology*. New York: Wiley-Interscience, 1986.

- Thaler, R. Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization* (Vol. 1, 1980), pp 39-60.
- Thaler, R. & Rosen, S. The Value of Saving a Life: Evidence from the Labor Market. In: N. Terleckyj, ed., *Household Production and Consumption*. New York: Columbia University Press, 1976.
- Thaler, R.H. & Shefrin, H.M. An Economic Theory of Self Control. *Journal of Political Economy* (Vol. 89, 1981), pp 392-406.
- Tolley, G. et al. *Establishing and Valuing the Effects of Improved Visibility in Eastern United States* (USEPA Grant #807768-01-0). Washington, D.C.: U.S. Environmental Protection Agency, 1986.
- Tribe, L.H. Policy Science: Analysis or Ideology? *Philosophy and Public Affairs* (Vol. 2, 1972), pp 66-110.
- Tufte, E. *Visual Display of Quantitative Data*. Cheshire, CT: Graphics Press, 1983.
- Turner, C. & Martin, E. *Measuring Subjective Phenomena* (2 vols). Beverly Hills, CA: Sage, 1985.
- Vickrey, W. Auctions, Markets, and Optimal Allocation. In: Y. Amihud, ed., *Bidding and Auctioning for Procurement and Allocation*. New York: New York University Press, 1976.
- Viscusi, W.K. *Risk by Choice: Regulating Health Safety in the Work Place*. Cambridge, MA: Harvard University Press, 1983.
- Wright, P. Writing Technical Information. *Review of Research in Education*. 1987.