Prediction versus Explanation in the Measurement of Values

Michael Hechter
Hyojoung Kim
Department of Sociology
University of Washington

&

Justin Baer
American Institutes for Research

Forthcoming, European Sociological Review
Abstract

Although economists and sociologists have often concluded that values and other internal states have little causal weight in determining behavior, there is some evidence that the tide is turning in this respect. This paper contributes to the recent revival of interest in subjectivity by comparing two different kinds of survey methods that can be used to measure values in general, and the value of civic-mindedness in particular. The explicit approach -- widely used in current empirical research -- derives value measures on the basis of direct questions asked in nationally representative surveys such as the General Social Survey and the National Election Studies. The factorial approach imputes values indirectly from answers to vignettes. In this study, the explicit approach is revealed to have greater predictive validity but substantially lower construct validity than the factorial approach. This finding highlights the distinction between prediction and explanation in social research, and casts doubt on the adequacy of revealed preference theory.
1. INTRODUCTION

Rational action theory is among the most venerable of social scientific research programs. At the individual level, its fundamental premise is that agents select a course of action from a feasible set on the basis of its expected value – that is, its expected net benefit -- to them. Although this premise seems mundane, in fact it can be quite provocative. For, in combination with knowledge about the opportunity cost of alternative courses of action, rational action theory suggests that an individual’s values can explain his subsequent behavior.

To produce useful empirical results, action theorists must be able to specify ex ante the values of the expected outcomes of the set of feasible courses of action. When these courses of action only have implications for pecuniary outcomes, this requirement is easily satisfied. This is so because there is good reason to believe that most people always prefer more money to less (Hechter 1994). In situations of this kind researchers can accurately specify agents’ values a priori – they have no need to query agents about their values.

To the degree that agents’ courses of action have implications for non-pecuniary outcomes, however, then matters become increasingly complex. It is one thing to hold -- as rational action theorists do -- that agents will choose that single course of action providing them with maximum expected net benefit. But it is quite another to specify just what the benefit in question consists of. Once we leave the monetary terrain, it is often difficult to determine ex ante just what a ‘benefit’ or ‘cost’ consists of. The reason, course, is that the very terms ‘benefit’ and ‘cost’ are themselves laden with values.
Values are one of a class of hypothetical constructs of psychological tendencies, or dispositions, to act in particular (that is, biased) ways. Preferences are also a member of this class of constructs. These psychological tendencies are of varying duration (Eagly and Chaiken 1993: 2). Whereas preferences are both highly specific and relatively evanescent (for example, they respond quickly to satiation effects), values are more general and durable internal criteria for evaluation (Hechter 1992; Hechter, Nadel and Michod 1993). Social norms are another matter entirely, for they provide external criteria for evaluation. Unlike preferences and values, therefore, they require sanctions to be effective (Hechter and Opp 2001).

To illustrate, assume that someone is facing a choice between two job situations that are identical in all respects but one. The first job is in a city with many overcast days (think Seattle) and the second one is in a city with many sunny ones (think Tucson). A person who prefers sunny skies to overcast ones might regard the prospect of living in Seattle as relatively costly; whereas another person (say, someone who is highly light-sensitive) might regard the same prospect as relatively beneficial. The point is that without knowing the value of sunny days to each person, it is quite impossible to predict their relative ranking of the two jobs. The phrase de gustibus non est disputandum has a very important implication: it suggests that what people take to be costs and benefits is hardly ever fixed in stone. If so, then much of the promise of rational action theory as a research program in social science would appear to be vitiated unless valid and reliable measures of individual values can be found.

Naturally, a great deal of effort has been expended in the search for such measures. For many of the value measures employed in sociological research, the results have been modest. As typically measured in survey research, values do a poor job of predicting behavior (Hechter et al. 1999: 405-406). Behavior indeed can be well predicted by highly detailed behavioral intentions.
that are elicited just prior to an individual’s action (Ajzen 1988; Ajzen and Fishbein 1980; Fazio 1986). But these intentions are both too specific and too proximal to the predicted behavior to be of abiding interest to social scientists.⁷

As a result, pessimism has grown about the causal status of individual values in current social science. Whereas values remain an abiding concern among many social psychologists, some cognitive psychologists doubt the utility of the concept entirely (Kahneman and Snell 1993). Moreover, interest in values has waned in most of the other social sciences – so much so, that complaints are increasingly being voiced about their absence in current social research (Campbell 1996; Hechter 2000; Hitlin and Piliavin 2004). Since values undergird all theories of action, much would have to be sacrificed if they were found to be superfluous concepts (Hechter et al. 1993; Campbell 1996; Kim and Bearman 1997; Boudon 2001; Joas 2000; Kanazawa 2001).⁸

In the present climate, we seem to be faced with two polar choices. Either values are empirically useless concepts that should be abandoned as causal factors in social explanations, or they are useful concepts that nonetheless require better measurement than they frequently have received.

Although a number of promising new approaches to value measurement have been advanced in several fields,⁹ these innovations are often costly and time-intensive to administer. Many researchers therefore rely on standard datasets – like the General Social Survey (GSS) and the National Election Studies (NES) -- to provide the value measures for their studies. Typically, these datasets employ value indicators based on self reports. In what this paper terms the explicit survey approach, researchers pose direct questions taken from standard datasets about
Doubts have been raised about the adequacy of the resulting measures, however. People may be unaware of some of their values; they may conceal true values from interviewers for strategic purposes; there is no baseline by which to compare the results of different subjects; and subjects in explicit surveys typically face little cost in misrepresenting their values (Fischhoff 1993).

In an attempt to provide a superior measure of values, Hechter and his associates (Hechter et al. 1999) employed an alternative implicit factorial survey method (based on Rossi 1951, 1979; Rossi et al. 1974; Rossi and Anderson 1982; Rossi and Berk 1985). In this method values are revealed not from direct questions, but from respondents’ choices in a series of hypothetical vignettes. The subjects in their study were asked to indicate their treatment preferences about their health states and prognoses in forty different vignettes, each of which constituted a mock advance directive, or living will. Since the measured values predicted individual treatment preferences well, the article suggested that factorial surveys offer hope for improving value measurement.

However, their approach left a number of questions unanswered. The method’s ability to predict behavior at a temporal remove from the value measurement is unclear. Previous research (some of it cited above) has already demonstrated that behavioral intentions can predict subsequent behavior with reasonable accuracy, but this is only a modest achievement. Since the predicted behavior in Hechter et al. (1999) – treatment preferences in a mock advance directive – occurred only a few instants after the measurement of the values, and since the values themselves
were closely related to the predicted behavior, questions can be raised about the adequacy of these value measures.  

The factorial survey method was designed to improve the validity and reliability of measures of individual values. Presumably, the method is more costly than its explicit counterpart, because respondents are presented with a large number of vignettes. Researchers interested in measuring values therefore face a distinct choice between explicit and factorial methods. To make an informed choice, they must appreciate the relative advantages and disadvantages of each method. Little is known in this regard.

Comparing measurement methods is inherently difficult because different methods employ different variables and different epistemological assumptions. Despite these difficulties, this paper explores the relative advantages of the explicit and factorial survey methods for measuring values. Specifically, we assess the relative validity and reliability of the measures of one particular value, civic-mindedness, generated by each method. The values measured by each method are then deployed to account for actual participation in community service.

The paper has three sections. The following section presents the research design. Next, we discuss the general analytic strategies for the measurement comparison. Finally, we report the statistical outcomes and conclude by discussing the implications for future research.

2. RESEARCH DESIGN

2.1. The Focal Value: Civic Mindedness
Although many different values might have been selected as candidates for this test of measurement methods, *civic mindedness* was chosen for its substantive import. Civic mindedness has become a topic of abiding interest among students of governance, economic development and civic society (Minkoff 1997; Putnam 1994; Putnam 1995; Putnam 2000; Skocpol and Fiorina 1999; Skocpol, Ganz and Munson 2000). Presumably, the salience of civic mindedness is positively associated with contributions to one’s community. In this study, civic mindedness is defined as *an individual’s disposition to sacrifice time and energy to provide some public service*.

2.2. The Survey Instruments

*The explicit survey.* Because values are both unobservable and general, one common way of measuring them in explicit surveys is by presenting respondents with a relatively broad spectrum of specific questions. This shotgun approach has a distinct advantage: Whereas some of the questions may miss the targeted value, there is a good chance that one or more will hit the mark. This hope is not blind; rather, it is buttressed by the popularity of the question items used in explicit surveys such as GSS and NES among others. These surveys do feature numerous questions that probe respondents’ views and behavior concerning community service. On the face of it, these questions are plausible indicators to civic-mindedness. From the GSS and NES, we select sixty-three question items on a variety of subjects that appear to be related to civic-mindedness. Most of these questions use a five-point Likert scale (see Appendix 1).

Note that the survey instruments used in GSS and NES commonly ask respondents to rate the salience of a single value at a given time. This rating method is used in other surveys as well,
most notably the World Value Survey (Inglehart 1990) and the Schwartz Value Survey (Schwartz 1994). Alternatively, one may pit one value against others and force respondents to rank order them. Examples of this ranking approach includes Kohn’s (1959, 1969) “top 3” method, the Rokeach Value Survey (Rokeach 1967, 1973). Each approach is built on a distinct conceptualization of values and thus defies easy comparison (see Hitlin and Piliavin [2004] for a review of the debate). One limitation of the ranking approach is that its validity hinges on a set of values that is preselected by researchers. Yet, survey researchers often cannot resort to any such accepted set of values. The rating approach does not entail this requirement. In practice, the value instruments in GSS, NES, and WVS are more widely utilized in secondary research. Accordingly, we focus on this rating implementation of the explicit survey as the most typical example of explicit survey methods.

The factorial survey. Whereas the explicit survey asks subjects direct questions about their values and behavior, the factorial survey approaches the measurement task indirectly. It never asks subjects to report on their internal states, but merely probes their behavioral intentions under varying conditions. This method has a particular advantage. It conditions behavioral intentions on the varying costliness of parameters, thereby providing an estimate of the opportunity cost of the given value.

The factorial survey consists of forty short vignettes describing a fictional community organization (Renew Seattle). Each vignette includes four dimensions assumed to influence the likelihood of participation in the group’s activities:

- the amount of time required to participate
- the degree to which participation would improve the subjects’ resumé
- the number of people who would be helped by the organization, and
- the amount of *fun* that subjects would derive from participation

These dimensions vary randomly across vignettes and are assessed in quantifiable intervals. For example, time commitment varies in increments of 20 from 10% to 90% of the amount of free time that participation in the community organization would require. Subjects report their willingness to participate by placing a mark on a horizontal line that appears beneath each question. The far left of the line corresponds to “Never Willing,” the center to “Uncertain,” and the far right to “Always Willing” (see Appendix 2).\(^\text{12}\)

### 2.3. Evaluation Criteria for the Value Measures

In assessing the relative merits of each method, we focus on construct and predictive validity of the respective indicators (Singleton, Jr. et. al. 1988). As the most popular and widely used criterion, *construct validity* assesses the fit between the indicators and the theoretical construct they purports to measure. A measure obtains construct validity to the degree that multiple indicators of a focal concept “converge on the same meaning, namely that conveyed by the underlying concept” (Singleton et al. 1988: 121).

In addition, we also assess the predictive validity of the indicators of civic-mindedness. A value measure has *predictive validity* to the degree that it predicts future behavior that is ostensibly motivated by the focal value. The criterion of predictive validity is of particular importance in the study of values. Due to their unobservability, values have been a contested concept in social research. Skeptics are only likely to countenance values if they are shown to affect behavior.
In the best of all possible worlds, the same method will yield measures with high construct and predictive validity; hence it will enable us to predict and explain the behavior of interest. Yet to the degree that any given behavior has multiple determinants, the predictive and construct validities of a value measure are likely to diverge.

2.4. Participation in Service Learning

Predictive validation of the measures of civic-mindedness requires identification of a behavioral pattern that is unambiguously motivated by this particular value. We focus on undergraduate student participation in service learning as the behavior of interest. Service learning is a federally-sponsored program in American colleges and universities that combines service to the community with student learning. According to the National and Community Trust Act of 1993, service learning helps students learn and develop through active participation in “thoughtfully organized service that is conducted in and meets the needs of communities.”

Our focus on service learning has two important advantages. On the one hand, participation in service learning is an instantiation of civic-minded behavior. This is especially the case where such participation is not a requirement but a choice for students. On the other, since records of service learning participation are centrally archived, we can obtain objective rather than self-reported measures of behavior.

2.5. The Samples

Both surveys were distributed to University of Washington undergraduates in three
classes offering the option of service learning in the Fall of 2000. Two of these were large introductory courses primarily intended for freshmen and sophomores (one in Women’s Studies, the other in the Comparative History of Ideas). The third was a mid-level course on development cross-listed in the Geography and International Studies departments. Each type of survey was randomly distributed on the first day of the quarter (hence, also the first day of class) before instructors mentioned anything at all about the option of service learning. No reference was made to service learning in either survey. This avoided the possibility that subjects’ responses would be influenced by framing effects, such as the instructor’s discussion of the benefits of service learning. Student decisions about participation in service learning were made subsequent to the survey at the end of that week.

After eliminating invalid surveys and missing values, we collected 118 factorial and 115 explicit surveys for analysis. Each subject in the sample was assigned a unique tracking identification number to preserve respondent anonymity. At the end of the quarter, we received a list of the tracking numbers of the students who participated in service learning. By matching this list to the tracking numbers taken from the surveys, we were able to obtain information about participation for these respondents.

3. MEASURES OF CIVIC MINDEDNESS

Since the two survey instruments employ different indicators of civic mindedness, we develop separate measurement strategies for each method.

3.1. Explicit Survey Indicators of Civic Mindedness
One of the advantages of the explicit survey method is its flexibility: it affords researchers a number of distinct indicators of values. Researchers can either cast a broad net by selecting a large array of value indicators, or opt for a narrow net, by selecting a short list of indicators. Narrow netters, in turn, can employ two different strategies to select value indicators. Indicators can be selected either empirically on the basis of association with the dependent variable, or on a priori grounds regardless of their actual association.

The strategy we employ for the explicit survey lies between these two extremes. Following the empiricist’s strategy, we first identify a subset of the sixty-three survey items in the explicit survey that have stronger empirical association with service learning participation. To do that, we ran a multiple logistic regression analysis of all 63 survey items in which an independent variable is recursively removed if it does not add to the explanatory power and a variable is included if it does. The procedure eventually identified a total of 19 survey variables (see Appendix 1 for a complete list of their survey questions).

From this set of the variables, we then select a final set of indicators on the basis of their substantive closeness to our definition of civic mindedness. On a priori grounds, a measure of civic mindedness should meet the following criteria: (1) A survey item should be general enough to capture a respondent’s overall disposition toward civic mindedness. (2) A survey item should not, on the face of it, incur other value orientations. Thus we reject the following variables because they are too specific to capture the general disposition of a respondent: the survey items that probe social service systems (VAR15; VAR64), charitable activities (VAR13, VAR56), and issues of inequality, human rights, environment, or youth work (VAR66-VAR69). Other survey items were rejected because they tend to reflect values other than civic-mindedness. These
include items that probe self-responsibility (VAR12), social change (VAR41), efficiency of community work (VAR40), and altruism (VAR53, VAR54, VAR56, VAR63). The remaining five survey items nicely capture the salience of civic mindedness in evaluation about the self, the general others, and friendship, in the perception of a good person, and in past behaviors.

These items are

- **Good Person (VAR17)**. Measured as a five category ordinal scale (from ‘strongly disagree’ to ‘strongly agree’) to the statement “The good person must be deeply involved in the problems and activities of the community.” This variable captures the degree to which respondents regard civic mindedness as a crucial criterion for judging the character of others, with a mean of 2.80 and standard deviation of 1.09.

- **Community Work (VAR55)**. Measured as a five category ordinal scale (from ‘never’ to ‘often’) to the question, “How often have you worked with others in helping to improve your community?” This attempts to capture the subject’s civic mindedness by providing a self-report of past community service, with a mean of 3.54 and standard deviation of 1.04.

- **Own Community Obligation (VAR58)**. Measured as a five category ordinal scale (from ‘no obligation’ to ‘strong obligation’) to the question, “How strong an obligation do you feel you have to help improve your community?” This variable reflects subjects’ sense of community obligation (e.g. it measures what some regard as a personal norm), with a mean of 3.25 and standard deviation of 0.94.

- **Others’ Community Obligation (VAR60)**. Measured as a five category ordinal scale (from ‘no obligation’ to ‘strong obligation’) to the question, “How strong an obligation do you feel people have to help improve your community?” This indicates civic mindedness by probing
the degree to which subjects consider communal obligation to be normative, with a mean of 3.17 and standard deviation of 0.91.

- *Friendship (VAR62)*. Measured as a five category ordinal scale (from ‘not at all desirable’ to ‘very desirable’) to the question “How desirable do you think it is for your friends to be community minded?” This question assesses subjects’ level of civic mindedness as reflected in its desirability for evaluating friendship, with a mean of 3.46 and standard deviation of 0.97.

The Community Work variable appears to violate the second criterion because it reflects not only one’s civic mindedness, but the effects of structural constraints (such as time limits) and opportunities (such as paid community work) as well. Despite this, we include it because behavioral indicators are frequently adopted in the literature.16

3.2. Factorial Survey Indicators of Civic Mindedness

In contrast to the explicit survey method, the factorial survey enables researchers to construct measurement strategies tailored to the definition of a given value. Yet, difficulties arise in a conventional factorial survey analysis in which a behavioral intention (such as willingness in our example) is regressed on vignette characteristics (such as time in our example) and the resultant regression coefficients are used as measures of values (Hechter et al. 1999). This strategy is valid if one is interested in measuring how much value a respondent assigns to the vignette characteristics such as fun, time, resources, or social capital. Like many other values, however, civic-mindedness does not allow for easy manipulation as a vignette characteristic. Neither are the residuals from such regressions helpful, for they include not only civic-
mindedness but many other values such as religiosity.

The theoretical definition of civic mindedness provides a solution. Since civic mindedness is defined as a disposition to sacrifice time and energy to provide some public service, one is civic minded if and only if one is willing to provide public service even if it is costly to do so. In other words, civic mindedness can be operationalized as unconditional willingness to contribute toward public good. To capture this unconditionality, each factorial survey vignette projects varying costs (the expected amount of time required for the service) and benefits (the expected contribution to one’s resumé, the expected amount of fun stemming from the service, and the expected degree of psychological satisfaction stemming from the awareness that one’s act helps a large number of other people) resulting from participation in the hypothetical Renew Seattle project (for a description of the vignette, see Appendix 1).

The vignette asks how willing a person is to participate in the project given these varying net costs. Willingness to participate is measured continuously to range between –3 (never willing) and +3 (always willing). Each of the net cost factors is pre-determined to vary across vignettes, from low (Time=10, 30 ; People=50, 100; Resumé=70, 90; Fun=70, 90), to medium (Time=50 ; People=25 ; Resumé=50; Fun=50), to high net cost (Time=70, 90; People=5, 10; Resumé=10, 30; Fun=10, 30). To ensure comparability across the four factors, all three benefit factors are subtracted from 100. Thus, the higher values associated with any of these factors indicates increasing net cost expected to stem from participating in the Renew Seattle project. The willingness variable is also recoded to range between 0 (never willing) and 6 (always willing).
Thus, the factorial survey indicators of civic mindedness are measured by positing the respondent’s expressed willingness to participate against each of the four net cost factors. Focusing on an individual vignette,

- *time*\(^{\text{willingness}}\) measures how willing a respondent is to participate in *Renew Seattle* even when such participation is time intensive.

- *resumé*\(^{\text{willingness}}\) measures how willing a respondent is to participate in *Renew Seattle* even when such participation is not expected to contribute to one’s job or career prospects.

- *people*\(^{\text{willingness}}\) measures how willing a respondent is to participate in *Renew Seattle* even when such participation is expected to help fewer people, and therefore be less effective.

- *fun*\(^{\text{willingness}}\) (the product of *fun* and *willingness*) measures how willing a respondent is to participate in *Renew Seattle* even if doing so isn’t enjoyable.

Since each respondent is given forty vignettes with varying net costs, an individual’s civic mindedness is measured by averaging the product of their expressed willingness to participate and the degree of net cost portrayed in each vignette over all forty vignettes. Formally, respondent \(i\)’s level of civic-mindedness for the \(k\)th indicator \((C_i^k)\) is

\[
C_i^k = \frac{1}{40} \sum_{v=1}^{40} \left[ W_{iv} X_{iv}^k \right],
\]  

where \(W_{iv}\) refers to \(i\)’s expressed willingness to participate in *Renew Seattle* for the \(v\)th vignette, and \(X_{iv}^k\) the corresponding net cost of participation. It follows that there are four measured factorial survey indicators of a respondent’s civic mindedness: *time*\(^{\text{willingness}}, *people*^{\text{willingness}}, *resumé*^{\text{willingness}, \text{ and }} *fun*^{\text{willingness}}.\caret

To illustrate how these variables help indicate civic mindedness, examine the following hypothetical examples focusing on the *time*\(^{\text{willingness}}\) variable. Suppose that respondent A’s
expressed willingness to participate is 6 (always willing) when participation is either expected to consume 10% of A’s free time (thus, time=10), or 90% of it (time=90). In contrast, respondent B’s expressed willingness is 0 (never willing) in both circumstances. Finally, suppose a third respondent C is always willing to participate (willingness=6) when it costs little time (time=10), but never willing (willingness=0) when it is too costly (time=90). According to our definition of the value, the relationships between one’s willingness to participate in a civic project and the net cost entailed in participation reveals an individual’s true level of civic mindedness. A’s willingness to participate in public service despite its high net cost reveals strong civic mindedness, B’s unwillingness at all levels of cost reveals weak civic mindedness at best, and C’s sensitivity to net costs reveals an intermediate amount of civic mindedness. The measurement equation (1) captures these differences: A has a \( \text{time} \times \text{willing} \) score of 300, B a score of 0, and C a score of 30. Note that we divide the resultant factorial survey indicators by 100 so that the scales for explicit and factorial survey indicators are comparable. As a result, the mean of \( \text{time} \times \text{willing} \) is 1.27 with standard deviation of 0.37. The corresponding statistics for \( \text{people} \times \text{willingness} \) are 1.93 and 0.35, 1.47 and 0.59 for \( \text{resumé} \times \text{willing} \), and 1.43 and 0.53 for \( \text{fun} \times \text{willing} \).

4. STRUCTURAL EQUATIONS MODELLING FOR VALIDATION

4.1. Model Specifications

In validating the construct and predictive validity of each method, we employ structural equations modeling strategy separately for each method. A structural equations model consists of
two parts. The *measurement* part assesses the construct validity of indicators of civic mindedness, whereas the *causal* part reveals the predictive validity of latent civic-mindedness for service learning participation. Figure 1 Panel A illustrates a baseline structural equations model for the explicit survey method. The rectangular boxes signify observed variables while the ellipses indicate unobserved latent variables. The upper left corner of the figure represents the measurement model for civic-mindedness. The lambda coefficient (λ₁ to λ₅) reports the factor loading of civic-mindedness in each indicator and assesses its construct validity. By contrast, the error variance (δ₁ to δ₅) and R-square statistic associated with each of the indicators assess its reliability. For scaling purposes, the lambda coefficient for Friendship (λ₅) is fixed at 1.0.

[Insert Figure 1 about here]

The remaining part of the path diagram assesses the predictive validation of civic-mindedness with respect to service learning participation. Here, the key information comes from the gamma coefficient of civic mindedness (γ₁).

In assessing predictive validity, it is important to control for structural factors that may also induce participation. Our survey probes students’ gender, year in college, the course in which a student is taking the survey, whether the course is a major requirement, membership in sorority or fraternity, past experience with service learning, and grade point average (GPA).¹⁸ Only two variables have consistent and statistically significant effects on service learning participation: whether a student is enrolled in Women 200 (WMN200), and logged hours worked per week (LNEMP). Hence, we control these two variables throughout the remaining analyses.¹⁹

Note that no study exists that helps identify the sub-dimensions of civic mindedness. Accordingly, the baseline model in Figure 1 Panel A is predicated on two simplifying assumptions: that civic-mindedness is uni-dimensional, and that measurement errors are not
intercorrelated. In the event that these assumptions are unjustifiable, we modify the baseline model according to various diagnostics. We follow the same modeling procedure for the factorial survey data.

4.2. Estimation

Figure 1 includes ordinal indicators of civic mindedness and service learning participation, but this poses a problem for estimating structural equations models. With ordinal indicators, the usual Maximum Likelihood estimates may be inefficient and inconsistent (Bollen 1989). However, Weighted Least Squares (WLS) estimates are asymptotically consistent and efficient. They are obtained by weighting polychoric correlations with an inverted asymptotic covariance matrix (Jöreskog and Sörbom 1996; Jöreskog 2001). Yet, this estimator has a drawback. Inverting an asymptotic covariance matrix is demanding on sample size; therefore the WLS estimator does not work well on a small sample. A minimum of 400 observations is suggested for a model with six indicators and a single dimension (Jöreskog 2000). As an alternative, one may use a Maximum Likelihood Estimator that corrects for excess multivariate kurtosis by an asymptotic covariance matrix (hereafter, Adjusted MLE; see Jöreskog 2001). This procedure does not involve inverting the asymptotic covariance matrix and hence tends to work better with as small a sample as 200 (Kelloway 1998, Jöreskog 2001).

But our samples are even smaller than this. The explicit survey has only 115 respondents and the factorial survey only 118. Consequently, we estimate both Adjusted MLE and WLS for our structural equations model. The results are robust. We focus on Adjusted MLE estimates.
5. FINDINGS

5.1. The Explicit Survey

5.1.1. The Baseline Model

Table 1 Panel A reports the Adjusted MLE estimates of the baseline structural equations model in Figure 1 Panel A. The numbers in parentheses are standard errors. Those in brackets are completely standardized regression coefficients. The overall fit statistics suggest that the model fits the data reasonably well.  

The left half of Panel A shows the measurement model for civic mindedness. The lambda coefficients measure the construct validity of the indicators of civic mindedness. All the explicit survey indicators have statistically significant lambda coefficients at the 0.05 level. This means that each is a valid indicator of civic-mindedness. The relative magnitudes of the completely standardized coefficients reveal that Own Community Obligation is the most valid indicator of civic-mindedness. In contrast, Good Person and Community Work are the least valid indicators.

[Table 1 about here]

The theta-delta (the error variance unexplained by civic-mindedness) and the R-square statistics measure the reliability of the indicators. The statistically insignificant theta-delta and large R-square statistic for Own Community Obligation indicate that it is the most reliable of the indicators. In contrast, Good Person and Community Work have the largest theta-deltas and smallest R-square statistics, demonstrating that they are the least reliable indicators of civic-mindedness.
The right half of Table 1 Panel A reports the predictive validity of Civic-mindedness, Course and Work. Together they explain 22% of the variance present in Service Learning Participation. However, only Course (enrollment in WMN200) and Work (LNEMP) are statistically significant at the 0.05 level. This suggests that the explicit survey measure of civic-mindedness lacks predictive validity.

5.1.2. The Revised Model

Despite the reasonably good fit statistics, there is room for improvement in model specification. The LISREL modification indices strongly suggest that Good Person directly affects the latent dependent variable, and that the error variances of Others’ Community Obligation and WMN200 are correlated. The correlation matrix of the five indicators also suggests that Good Person and Community Work are distinct from the other three indicators. Community Work taps into past behaviors. By probing the role of civic-mindedness in character assessment, Good Person appears to capture a more general dimension than the other three indicators. In addition, the gamma coefficient for Work is positive and statistically significant at the 0.05 level. This finding is counter-intuitive because it suggests that students who worked more were more likely to participate in service learning.

Given these results we revise the baseline model (see Figure 1 Panel B). The revised model differs from the previous one in two respects. First, civic-mindedness is now modeled as a three-dimensional concept. Second, the error variances of Others’ Community Obligation and WMN200 are now allowed to be correlated. Table 1, Panel B reports the Adjusted MLE estimates for this revised model. All the fit statistics indicate that the revised model is superior.
The revised model also resolves some of the difficulties encountered in the previous model. As expected, the error variances of Others’ Community Obligation and WMN200 are inter-correlated with a statistically significant theta-delta of -0.22. And Work has a statistically insignificant gamma coefficient.

Since Civic-mindedness 1 and 2 each have only one indicator, we cannot assess their construct validity and reliability. Therefore we focus on Civic-mindedness 3. The lambda coefficients of Own Community Obligation and Others’ Community Obligation are statistically significant at 0.05 level, demonstrating the construct validity of the explicit survey indicators. The validity and reliability of these two indicators vary, however. Own Community Obligation has the largest completely standardized lambda coefficient, indicating the highest level of construct validity. It also has a statistically insignificant theta-delta, and the largest R-square statistic, and thus attains high reliability. In contrast, Others’ Community Obligation and Friendship have sizable and statistically significant error variances with corresponding R-square statistics being less than 0.50. These indicators are valid but unreliable measures of civic-mindedness.

Turning to predictive validity, the causal component of the model, the gamma coefficients for Civic-mindedness 2 and 3 are statistically insignificant. However, the latent variable of Civic-mindedness 1 does have a large and statistically significant effect on the latent dependent variable. Thus Civic-mindedness 1 does have predictive validity.

In sum, the explicit survey indicators do obtain construct validity and reliability, but in varying degrees. Own Community Obligation is the only explicit survey indicator that is both valid and reliable. In contrast, Good Person is the only predictively valid indicator.
5.2. The Factorial Survey

5.2.1. The Baseline Model

Figure 2, Panel A graphically illustrates the factorial survey counterpart of the baseline model. As in the explicit survey, this model assumes a single dimension for civic-mindedness and no correlation across error variances. Table 2, Panel A reports the Adjusted MLE estimates of this baseline model. The lambda coefficients for the factorial survey indicators of civic-mindedness are all statistically significant at 0.05 level. And their completely standardized coefficients are highly comparable. Not only do the factorial survey indicators obtain construct validity, but each provides as good a measure of civic-mindedness as any other. Moreover, the theta-delta of each indicator is statistically insignificant at 0.05 level; the corresponding R-square statistic reveals that most of the variance in the observed indicators is explained by civic-mindedness. Evidently, the factorial survey indicators – and each of them – are valid and reliable measures of civic-mindedness.\(^{22}\)

[Figure 2 about here]

[Table 2 about here]

However, the factorial survey measure of civic-mindedness fails to obtain predictive validity. The gamma coefficient of Civic-mindedness is positive but statistically insignificant at the 0.05 level. Apparently, service learning participation is determined not by students’ value orientations, but by external structural conditions such as exposure to their instructor’s advocacy of service learning (Course) and time constraints resulting from employment (Work).
5.2.2. The Revised Model

However, the fit statistics of the baseline model -- especially the Satorra-Bentler Scaled Chi-Square and RMSEA -- strongly suggest that it does not fit the data well. The LISREL diagnostics (modification indices) show that Fun*Willing behaves somewhat differently from the other three indicators and that the error variances of Resumé*Willing and WMN200 should be allowed to be correlated. Accordingly, we revise the baseline model (see Figure 2, Panel B). The revised model differs from the previous one in two respects. First, civic-mindedness is now indicated only by three observed independent variables (Fun*Willing has been removed). Second, the error variances of Resumé*Willing and WMN200 are now allowed to be correlated.

Table 2, Panel B reports the Adjusted MLE estimates for this revised model. The validation outcomes are essentially the same as in the baseline model. The factorial survey indicators are valid and reliable in measuring civic-mindedness. But once more they fail to obtain predictive validity. The major difference lies in the overall fit statistics. The model now fits the data reasonably well. Thus the factorial survey indicators are highly valid and reliable in measuring the underlying theoretical construct of civic-mindedness, but they fail to predict service learning participation.

5.2.3. Comparison of the Methods

How do the two different survey methods fare in measuring civic mindedness? No clear-cut winner emerges. Both methods produce valid indicators. The factorial survey is superior in achieving construct validity. Unlike the explicit survey, its indicators perform uniformly well.
Likewise, the factorial survey indicators had no statistically significant random measurement errors, in contrast to all but one of the explicit survey indicators. Despite this, the explicit survey method has superior predictive validity. The factorial survey measure of civic-mindedness fails to predict service learning participation.

On the surface, these findings seem to suggest that the explicit survey method should be preferred over its factorial counterpart. It appears to do a good job in obtaining both construct and predictive validity. However, there is a striking discrepancy between the indicator with the highest degree of construct validity and reliability (Own Community Obligation) and the indicator with predictive validity (Good Person). The most valid and reliable indicator of the underlying value fails to predict the target behavior, whereas the least valid and reliable indicator is the best predictor of it. Although the explicit survey indicators are better able to predict service learning participation, at the end of the day we have little inkling about the mechanism responsible for this relationship.26

6. CONCLUSION

There is a growing appreciation that the study of values merits renewed attention in sociology.27 The hope is that better measures of individual values will yield superior explanations of individual behavior and social outcomes. Although value measurements of increasing sophistication are currently being developed in several disciplines, many of these are too cumbersome or expensive to be useful in large-scale social research. Our study compares the relative efficacy of two survey methods used to measure values that have no such limitation -- the explicit survey based on questions from large-scale datasets, and its factorial counterpart. In
the explicit survey respondents typically are asked direct questions about behaviors and attitudes in hopes of capturing the value of interest. In the factorial survey, respondents are asked to state a behavioral intention under a very small set of varying conditions. Values then are inferred indirectly from these behavioral intentions. To borrow a metaphor from the world of fashion, the explicit survey is off-the-rack, but the factorial survey is custom-made.

Our study indicates that neither method is unequivocally superior. The explicit survey does a better job of predicting participation in community service, but the factorial survey attains greater construct validity. Whereas we are accustomed to thinking that an adequate value measure will enable us to both predict and explain, this finding foregrounds the distinction between prediction and explanation (Elster 1989: 8; Lieberson and Lynn 2002: 10-11). The explicit survey does better at predicting service learning behavior because it measures a wider range of values. Some of these values – like altruism -- may be distinct from civic-mindedness but somewhat related to it. Although the explicit survey gives us a greater ability to predict which students will perform community service, it doesn’t tell us why they participate.

In contrast, the factorial survey is specifically constructed to measure civic-mindedness; in addition, it attempts to gauge the opportunity cost of participation in community service. For these reasons it is not surprising that it provides a better measure of the given value. However, knowing which students are most civic minded does not enable us to predict who actually participates in community service. Clearly participation in community service – like most other behaviors – is not uniquely determined by a single value. This finding casts doubt on the adequacy of revealed preference theory, the economists’ favorite means of evading value measurement. On the contrary, behavior can be motivated by a number of different values as well as other kinds of factors. As our analysis demonstrates, structural factors such as the
exposure to instructor advocacy and time constraints play a large role in determining participation in service learning.

This study therefore suggests that the optimal choice of these two methods to value measurement depends on the goals of the intended research. For research aiming to predict social outcomes (like elections), the explicit survey based on general datasets has much to commend it. For research aiming to understand the reasons why people behave as they do – critical for the development of successful public policy interventions -- the factorial survey should be given serious consideration.
References


### Table 1. Adjusted Maximum Likelihood Estimates: The Explicit Survey

**A. The Baseline Model**

<table>
<thead>
<tr>
<th>Measurement Model</th>
<th>Lambda-X</th>
<th>Theta Delta</th>
<th>R-sq</th>
<th>Causal Model</th>
<th>Gamma</th>
<th>Psi</th>
<th>R-sq</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Good Person</strong></td>
<td>0.54 (0.21) [0.35]</td>
<td>0.87 (0.13)</td>
<td>0.13</td>
<td>Civic-mindedness</td>
<td>0.20 (0.28) [0.13]</td>
<td>0.78</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Community Work</strong></td>
<td>0.66 (0.18) [0.43]</td>
<td>0.82 (0.12)</td>
<td>0.18</td>
<td>Course</td>
<td>0.41 (0.10) [0.41]</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td><strong>Own Community Obligation</strong></td>
<td>1.38 (0.29) [0.90]</td>
<td>0.19 (0.16)</td>
<td>0.81</td>
<td>Work</td>
<td>0.15 (0.05) [0.15]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Others’ Community Obligation</strong></td>
<td>1.04 (0.22) [0.68]</td>
<td>0.54 (0.16)</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Friendship</strong></td>
<td>1.00 -- [0.65]</td>
<td>0.58 (0.17)</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fit Statistics:** $\chi^2=25.6 \ (df=17; \ p=0.08)$ RMSEA=0.07 (p=0.28) CFI=0.88 IFI=0.89 AGFI=0.84

**B. The Revised Model**

<table>
<thead>
<tr>
<th>Measurement Model</th>
<th>Lambda-X</th>
<th>Theta Delta</th>
<th>R-sq</th>
<th>Causal Model</th>
<th>Gamma</th>
<th>Psi</th>
<th>R-sq</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Good Person</strong></td>
<td>1.00 --</td>
<td>-- --</td>
<td></td>
<td>Civic-mindedness 1</td>
<td>0.36 (0.14) [0.36]</td>
<td>0.69</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Community Work</strong></td>
<td>1.00 --</td>
<td>-- --</td>
<td></td>
<td>Civic-mindedness 2</td>
<td>0.05 (0.14) [0.05]</td>
<td>(0.15)</td>
<td></td>
</tr>
<tr>
<td><strong>Own Community Obligation</strong></td>
<td>1.41 (0.30) [0.90]</td>
<td>0.18 (0.16)</td>
<td>0.82</td>
<td>Civic-mindedness 3</td>
<td>-0.08 (0.31) [-0.05]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Others’ Community Obligation</strong></td>
<td>1.04 (0.22) [0.67]</td>
<td>0.55 (0.16)</td>
<td>0.45</td>
<td>Course</td>
<td>0.41 (0.10) [0.41]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Friendship</strong></td>
<td>1.00 -- [0.64]</td>
<td>0.59 (0.17)</td>
<td>0.41</td>
<td>Work</td>
<td>0.11 (0.06) [0.11]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fit Statistics:** $\chi^2=4.77 \ (df=9; \ p=0.85)$ RMSEA=0.00 (p=0.93) CFI=1.00 IFI=1.01 AGFI=0.94

1) Numbers in parentheses are standard errors, whereas those in brackets are completely standardized coefficients.

2) The Adjusted MLE estimate of the theta-delta between Others’ Community Obligation and WMN200 is -0.22 (0.08).
Table 2. Adjusted Maximum Likelihood Estimates: The Factorial Survey

A. The Baseline Model

<table>
<thead>
<tr>
<th>Measurement Model</th>
<th>Lambda-X</th>
<th>Theta Delta</th>
<th>R-sq</th>
<th>Causal Model</th>
<th>Gamma</th>
<th>Psi</th>
<th>R-sq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fun*Willing</td>
<td>1.00 (0.02)</td>
<td>0.11 (0.10)</td>
<td>0.89</td>
<td>Civic-mindedness</td>
<td>0.05 (0.07)</td>
<td>0.76</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>[0.94]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resumé*Willing</td>
<td>1.01 (0.02)</td>
<td>0.09 (0.11)</td>
<td>0.91</td>
<td>Course</td>
<td>0.42 (0.09)</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.95]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People*Willing</td>
<td>0.99 (0.02)</td>
<td>0.13 (0.11)</td>
<td>0.87</td>
<td>Work</td>
<td>-0.18 (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.94]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*Willing</td>
<td>1.00 --</td>
<td>0.12 (0.10)</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.94]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fit Statistics: \( \chi^2 = 49.1 \ (df=11; \ p=0.00) \) \( \text{RMSEA} = 0.17 \ (p=0.00) \) CFI=0.98 IFI=0.98 AGFI=0.86

B. The Revised Model

<table>
<thead>
<tr>
<th>Measurement Model</th>
<th>Lambda-X</th>
<th>Theta Delta</th>
<th>R-sq</th>
<th>Causal Model</th>
<th>Gamma</th>
<th>Psi</th>
<th>R-sq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resumé*Willing</td>
<td>0.99 (0.02)</td>
<td>0.11 (0.11)</td>
<td>0.89</td>
<td>Civic-mindedness</td>
<td>0.10 (0.07)</td>
<td>0.74</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>[0.94]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People*Willing</td>
<td>0.97 (0.02)</td>
<td>0.14 (0.11)</td>
<td>0.86</td>
<td>Course</td>
<td>0.42 (0.09)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.93]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*Willing</td>
<td>1.00 --</td>
<td>0.09 (0.10)</td>
<td>0.91</td>
<td>Work</td>
<td>-0.17 (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.95]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fit Statistics: \( \chi^2 = 9.97 \ (df=5; \ p=0.08) \) \( \text{RMSEA} = 0.09 \ (p=0.17) \) CFI=0.99 IFI=0.99 AGFI=0.92

1) Numbers in parentheses are standard errors, whereas those in brackets are completely standardized coefficients.
2) The Adjusted MLE estimate of the theta-delta between Resource*Willing and WMN200 is 0.08 (0.01).
Appendix 1: Survey Examples

A. The Explicit Survey

Please indicate the degree to which you agree or disagree with each of the following statements by circling the appropriate number. If you do not know (“DK”), please circle the number 9. For example, if you strongly disagree with a statement, you would circle the number 1.

VAR 12. In our society everyone must look out for him or herself. It is of little use to work with others to improve your community.

VAR 13. Charitable acts are self-defeating because they increase the recipient’s dependence on external help rather than assisting them to become self-reliant.

VAR15. Our system of social services goes much too far. It takes too much care of people and deprives them of too much individual responsibility.

VAR17. The good person must be deeply involved in the problems and activities of the community.

How important or unimportant is each of the following factors in determining whether you would join a charitable or community organization? (1=unimportant; 5=important)

VAR40. If you could make a contribution to your local community

VAR41. If you could help to bring about social or political change

Here is a list of various changes in our society that might take place in the near future. For each one, please indicate whether you think it would be a good thing, a bad thing, or somewhere in between.

VAR53. How important is it for a child to learn to help others when they need help?

VAR54. How much do you really care about what happens to other people?

VAR55. How often have you worked with others in helping to improve your community?

VAR56. How often do members of your family or household give money, food, or clothing to the less fortunate?

VAR58. How strong an obligation do you feel you have to help improve your community?

VAR60. How strong an obligation do you feel people have to help improve their community?

VAR62. How desirable do you think it is for your friends to be community-minded?

VAR63. How desirable do you think it is for your friends to be generous toward the needy?

Please circle the number of hours you estimate you have spent participating in each of the following activities and organizations since you entered college: (0-3; 4-10; 11-24; 25+)

VAR64. Social welfare services for elderly, handicapped or deprived people

VAR66. Local community action on issues like poverty, employment, housing, or racial equality

VAR67. Third world development or human rights

VAR68. Conservation, the environment, ecology

VAR69. Youth work (e.g. scouts, guides, youth clubs, etc.)

B. The Factorial Survey

1. Renew Seattle is a new project that builds community in the Seattle area.

According to faculty who have had experience with this project, by participating in Renew Seattle, the strength of your résumé would increase by 50%, you’d be engaged in a task that 10% of students have found to be fun, and you’d be helping 50 people in your community. Your participation would take up 70% of your free time.

How willing would you be to participate in the project?

Never Willing | Uncertain | Always Willing
--------------|-----------|-----------
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgements

The first and second authors contributed equally to the conception and completion of this project. This research was made possible by a grant from the Center for Statistics and the Social Sciences at the University of Washington. Earlier versions of the paper were presented at colloquia at the Instituto Juan March, Madrid, the Department of Sociology at Cornell University and the Center for Statistics and Social Sciences at the University of Washington. We acknowledge helpful comments and suggestions from two anonymous reviewers, and from Ross Matsueda, Peter Hoff, and Hartmut Esser. Direct all correspondence to Michael Hechter (hechter@u.washington.edu) or Hyojoung Kim (hyojoung@u.washington.edu), Department of Sociology, University of Washington, Seattle, WA 98195.

Endnotes

1 Action theories (such as rational choice theory) hold that individual behavior is importantly affected by agents’ intentions; action therefore is regarded at least in part as the product of \textit{ex ante} deliberation. By contrast, behavioral theories (such as learning theory) usually discount the causal role of intentionality in accounting for individual behavior.

2 Whereas it is child’s play to determine benefits and costs \textit{ex post}, this kind of exercise is merely tautological.

3 To say that values are more durable than preferences is not to say how much more durable they must be to qualify as values. At this point, there can be no hard and fast answer to this question. Clearly, values change over the life course: we would be surprised if some of children’s values persisted into adulthood. Moreover, the rate of value change is likely to vary across individuals. Similar caveats hold for the generality of values relative to that of preferences.
For the purposes of this discussion ‘internalized norms’ are considered to be values. There is little consensus about the definition of values in the scholarly literature. Whereas some theorists use the definition employed in this paper, others do not. Thus Joas (2000: 16) argues that values exist over and above desires; they enable us to reflect upon and thus evaluate our desires. Note that civic mindedness, the concept of concern in the present study, constitutes a value according to both of these otherwise different definitions.

Here ‘job situation’ refers to the qualities of the job and all aspects of its setting – e.g. the city, school systems, cultural resources and so forth. Obviously, two cities with such different climates are likely to differ in many other respects as well, but for the purposes of this illustration the only relevant difference is in the amount of sunlight afforded by each job situation.

The most extensive discussions are to be found in the literature relating values (often referred to as ‘attitudes’) to behavior in social psychology (Eagly and Chaiken 1993: Ch 4).

Hence, Ajzen and Fishbein’s theory of reasoned action rings hollow to the sociologically-attuned ear.

For a recent review of the related (and voluminous) literature in behavioral economics on preferences and preference formation, see Rabin (1998: Part 2).

For example, the implicit association test has been used in psychology to measure a variety of internal states (Greenwald and Nosek 2001); economists have employed the ultimatum game to detect variations in the value of fairness (Henrich et al. 2001); political scientists have incorporated experimental methods into their surveys (Zaller 1992; for a useful review, see Sniderman and Grob 1996).

See fn. 3 above.
Ironically, although this literature is ostensibly about civic mindedness, none of the contributors to it attempt to measure the value. Instead, they are content to regard a behavioral indicator – participation in civic associations – as an unproblematic proxy for the value itself. As this paper will demonstrate, the mere existence of such participation cannot be attributed to any particular value. The degree to which civic mindedness leads to civic participation is unknown and deserves empirical study.

The large number of vignettes with random permutations of vignette specifications of fun, resumé, people and time may induce boredom, loss of concentration and difficulty in recognizing varying vignette conditions among respondents. If these problems occur systematically they will compromise the validity of our indicators; if they occur randomly, it will challenge their reliability. A focus group study conducted among undergraduates before the surveys were administered, however, reveals that there is little cause for concern on this score. Students in the focus group quickly recognized the permutational pattern. None expressed any boredom or difficulty in recognizing the varying conditions. Our confidence in the method is strengthened by the finding that all the factorial survey indicators achieve a high level of measurement validity and reliability (see Table 1). Yet, further research is needed to examine this issue -- especially when the factorial survey method is applied to less educated respondents who are not as skilled as college students in pattern recognition.

Prior to distributing the surveys, both instruments were completed by two focus groups comprised of undergraduates, our sample population. We refined the surveys based on their comments, striving for greater clarity in tapping civic mindedness.
Surveys were invalidated if any of the following conditions were met: 1) the subject was under the age of 18, 2) the subject completed less than half of a survey, or 3) the subject’s name was missing or illegible.

The complete survey that includes all 63 value items is available upon contact with the authors.

We also conducted the structural equations analysis reported in the remaining the paper without this behavioral indicator. However, its removal had no effect on the conclusion.

One may suspect that our measurement strategy of including the same willingness in each of the factorial survey indicators may artificially increase the correlations among the value indicators. However, the respondent-vignette level correlations among the four indicators remain low (they are no higher than 0.37). See footnote 22 for further discussion of the implication of this measurement strategy on the findings.

While all other control variables are obtained from student self-report, the GPA scores are directly obtained from the University Registrar.

They are included in the model without any measurement error (i.e., $\lambda_6=\lambda_7=1.0$ and fixed error variance for each). Similarly, we also construct a latent variable (S/L Latent) for the observed service learning participation (S/L Observed) without any measurement error ($\lambda_8=1$ with a fixed error variance).

We report Satorra-Bentler Scaled Chi-squares that correct for non-normality by using an asymptotic covariance matrix but without inverting it (Jöreskog 2001). Note also that despite the oft-noted concern about RMSEA in small samples, RMSEA has been found to be overestimated in a small sample (Curran et al. 2003). Therefore, RMSEA provides a conservative test of the model fit.
Each of the latent variables “Civic-mindedness 1” and “Civic-mindedness 2” is indicated by a single observed independent variable. Accordingly, we fix the corresponding lambda coefficient at 1.0 with fixed error variance.

The highly comparable lambda coefficients for the factorial survey indicators may be an artifact of the computation of the indicators, rather than from a more valid means of asking questions. As discussed above, each factorial survey indicator is computed by averaging over the 40 vignettes the product of a single willingness score to each of the four vignette-specific Time, Resumé, People, and Fun. To this end, we examined the correlational patterns among the indicators at the respondent-vignette level, before they are averaged out per respondent. The obtained correlations do not exceed 0.37, indicating that less than 14% of the variance in each indicator is explained by another indicator. Despite the multiplication entailed in the measurement procedure, the resultant vignette-level indicators are sufficiently different from each other. This confirms our interpretation that the lambda coefficients in Table 1 indeed reflect the validity of asking questions in factorial survey method.

We attempted to implement two separate dimensions of civic-mindedness, one for Fun*Willing and the other for the remaining three indicators. This strategy resulted in a high level of multicollinearity between these two dimensions.

Might other vignette specifications affect the outcomes? Table 2 suggests otherwise. The highly comparable factor loadings across factorial survey indicators suggest that they are to a large extent redundant. This makes sense because civic mindedness is operationalized here as the unconditional willingness to engage in public service.

The completely standardized lambda coefficients range between 0.93-0.95 for the factorial survey, whereas they range between 0.35 to 0.90 for the explicit survey.
There are other examples of the disjunction between prediction and explanation in the literature. Consider Inglehart’s well-known research program on post-material value shift based on the European Values Survey. Although the values in this research do appear to have some predictive ability (Inglehart and Baker 2000), the value scales have little construct validity (Haller 2002 is the latest in a long string of critiques of this research program).

As discussed above, our explicit survey method only implements the rating-based approach. Hence, the following conclusions are not directly applicable to other explicit survey methods such as those based on ranking.

Another criterion for comparing the *practical* merit of each method is the relative efficiency of data generation. Our study suggests that the explicit survey method is apparently more efficient than the factorial survey because it relies on only four indicators instead of forty vignette questions. Yet, any such conclusion is premature. No research has specifically examined the minimal number of vignettes required for adequate imputing value measures. The apparent redundancy among the factorial survey indicators in our study suggests that fewer vignettes may be required. This question can only be answered on the basis of further research.
Figure 1. Path Diagrams of the Structural Equations Models: The Explicit Survey

A. The Baseline Model

\[ \delta_1 \rightarrow \text{Good Person} \]
\[ \delta_2 \rightarrow \text{Community Work} \]
\[ \delta_3 \rightarrow \text{Own Community Obligation} \]
\[ \delta_4 \rightarrow \text{Others’ Community Obligation} \]
\[ \delta_5 \rightarrow \text{Friendship} \]
\[ \text{WMN200} \]
\[ \text{LNEMP} \]

\[ \lambda_1 \]
\[ \lambda_2 \]
\[ \lambda_3 \]
\[ \lambda_4 \]
\[ \lambda_5 = 1 \]
\[ \gamma_1 \]

\[ \epsilon \]

\[ \lambda_6 = 1 \]
\[ \gamma_2 \]
\[ \gamma_3 \]

\[ \lambda_7 = 1 \]

\[ \delta_1 \rightarrow \text{Civic-mindedness} \]
\[ \delta_2 \rightarrow \text{Civic-mindedness} \]
\[ \delta_3 \rightarrow \text{Civic-mindedness} \]
\[ \delta_4 \rightarrow \text{Civic-mindedness} \]
\[ \delta_5 \rightarrow \text{Civic-mindedness} \]

\[ \lambda_8 = 1 \]

B. The Revised Model

\[ \delta_1 \rightarrow \text{Good Person} \]
\[ \delta_2 \rightarrow \text{Community Work} \]
\[ \delta_3 \rightarrow \text{Own Community Obligation} \]
\[ \delta_4 \rightarrow \text{Others’ Community Obligation} \]
\[ \delta_5 \rightarrow \text{Friendship} \]
\[ \text{WMN200} \]
\[ \text{LNEMP} \]

\[ \lambda_1 = 1 \]
\[ \lambda_2 = 1 \]
\[ \lambda_3 \]
\[ \lambda_4 \]
\[ \lambda_5 = 1 \]
\[ \gamma_1 \]
\[ \gamma_2 \]
\[ \gamma_3 \]
\[ \gamma_4 \]
\[ \gamma_5 \]

\[ \epsilon \]

\[ \lambda_6 = 1 \]
\[ \lambda_7 = 1 \]

\[ \lambda_8 = 1 \]
Figure 2. Path Diagrams of the Structural Equations Models: The Factorial Survey

A. The Baseline Model

δ₁ → Fun*Willingness
δ₂ → Resource*Willing
δ₃ → People*Willing
δ₄ → Time*Willing

Civic-mindedness

WMN200 ← λ₅=1
LNEMP ← λ₆=1

Course
Work

S/L Latent ← λ₇=1

ε

δ₁ → Fun*Willingness
δ₂ → Resource*Willing
δ₃ → People*Willing
δ₄ → Time*Willing

Civic-mindedness

WMN200 ← λ₅=1
LNEMP ← λ₆=1

Course
Work

S/L Latent ← λ₇=1

ε

B. The Revised Model

δ₁ → Resource*Willing
δ₂ → People*Willing
δ₃ → Time*Willing

Civic-mindedness

WMN200 ← λ₅=1
LNEMP ← λ₆=1

Course
Work

S/L Latent ← λ₇=1

ε