# CONVERGENT AND DIVERGENT VALIDITY OF THREE MEASURES OF CONSERVATION BEHAVIOR
## The Multitrait-Multimethod Approach

**VICTOR CORRAL-VERDUGO** *is a professor of psychology at the University of Sonora in Hermosillo, Mexico, and a member of Mexico's National Researchers System. He obtained a Ph.D. from the University of Arizona. His research interests include the study of determinants of proenvironmental behavior (PEB) and methods of research in PEB.*

**AURELIO JOSÉ FIGUEREDO** *is an associate professor and director of the Ethology and Evolutionary Psychology Program at the University of Arizona–Tucson. He obtained a Ph.D. from the University of California–Riverside. His research interests include evolutionary psychology of domestic violence, quantitative ethology, and the use of latent variable modeling applied to social problems.*

ABSTRACT: Observations of the amount of reuse of glass, clothing, and metal were conducted at households of 130 individuals, and those direct observations were contrasted with the self-report of reuse of the same products. Two kinds of self-report were obtained: frequency of reuse self-reports and quantity of reuse self-reports. Thus, patterns of reuse of each particular type of material were assessed using the three methods of measurement. A multitrait-multimethod (MTMM) matrix of correlations between the reuse of these three materials, using these three methods of measurement, was analyzed. This matrix revealed the convergent and discriminant validities for the assessments of reuse. Higher correlations between direct observations and quantity self-reports were obtained than between observations and frequency reports. A confirmatory factor analysis of the MTMM matrix confirmed those results, adding significance testing to the validity assessment and to the partitioning of trait and method variance, modeled as latent factors.

**The validity of measures** of proenvironmental behavior (PEB) is a topic of concern for most investigators within the field of conservation behavior

AUTHORS' NOTE: *This study was made possible thanks to grant L0069-H from Mexico's National Council of Science and Technology (CONACyT). Reprint requests*

ENVIRONMENT AND BEHAVIOR, Vol. 31 No. 6, November 1999 805-820

research. Without the assessment of validity, no researcher could claim that what is purportedly being measured is what is actually being measured.

PEB is manifested in the form of a variety of interrelated activities, yet it also shows multiple facets. These conceptually related behaviors do not necessarily constitute a unitary latent construct involving every kind of proconservation activity (Berger, 1997; Daneshvary, Daneshvary, & Schwer, 1998; Lee, De Young, & Marans, 1995). For example, not only are recycling and reuse different, and potentially competing, behaviors, but also each particular conservation action seems to consist of independent types of activity. Being a recycler of aluminum cans does not make someone a recycler of paper (Corral-Verdugo, 1996), and general recyling does not predict the recycling of a particular object (Lee et al., 1995). This multifaceted nature of conservation behavior results in a more complex picture, which requires precise and valid measurement when investigating this kind of activity.

Construct validation includes the following principles: (a) independent verification, using related measures of the same studied characteristic, and (b) discrimination of results, produced by measures of different characteristics (Campbell & Fiske, 1959). The first notion relates to the concept of convergent validity, whereas the second has to do with discriminant or divergent validity.

Typically, a measure of PEB is validated by contrasting results of several assessments of a given variable, for instance, paper recycling. People are asked to report whether they recycle sheets, packing paper, and non-packing paper. A high or significant correlation of responses between these three items indicates that paper recycling is being measured in a (convergently) valid manner. To assess the divergent validity, one has to contrast the results of paper recycling against one or more independent constructs, for example, paper reuse. If the correlation between these two activities is low, it would indicate a discrimination between reuse and recycling assessments.

In assessing validity, the emphasis is generally placed on traits—features of behavior that are the interest of PEB researchers. Examples of traits are conscious consumption, newspaper recycling, and proecological attitude. It is assumed that these traits are important in the explanation of the total partitioning of variance of any assessment. Total variance means the entire variation of a given behavioral measure, whereas total partitioning of variance refers to the components of such total variation. Classically, total partitioning of variance is thought to be composed from explained variance and

---

*should be sent to Victor Corral-Verdugo, Provincia Albacete 98, Fracc. Los Portales, Hermosillo, Sonora, 83240, Mexico; e-mail: vcorral@kunkaak.psicom.uson.mx.*

unexplained variance. When estimating the variance of a behavioral measure, the focus is on trait variance (i.e., the portion of total variance in a given variable that is explained by a trait). The underlying assumption is that the response of a participant to a research question is explained by a latent trait, plus the unexplained component, called error.

Campbell and Fiske (1959) acknowledged additional influences in the total partitioning of variance of any measured variable. One of these is method variance. Method variance implies that the use of a particular way of collecting data can influence the response of the individuals studied. As the pertinent literature shows, PEB results may vary if they are obtained from observations, verbal self-reports, or traces of behavior (Corral-Verdugo, 1997; Corral-Verdugo, Bernache, Encinas, & Garibaldi, 1995; Cote, 1984; Geller, 1981; Rathje, 1989). Therefore, total variance in a particular assessment would be formed from trait variance, plus method variance, plus error (Marutza, 1977).

Because methods play a role in explaining variability of measures, they also should be considered in the assessment of validity. Campbell and Fiske (1959) suggested a multitrait-multimethod (MTMM) matrix of correlations as a criterion of validity. A trait might be a behavior, a facet, or a characteristic of something studied, whereas a method is a way of measuring a trait. According to Campbell and Fiske, independent measures obtained from a single method share variance (due to the method). For example, Corral-Verdugo (1997) found that the self-reporting of proenvironmental beliefs, motives, and behavior resulted in significant correlations among these three variables. However, the correlations of those self-reported beliefs and motives with behavior significantly decreased when such behavior was observationally assessed. Thus, the former correlations seemed to be partially explained by the shared method variance, which was caused by the use of a single method. Also, according to Campbell and Fiske, different measures of a trait could be interrelated due to the shared trait variance. For example, three independent assessments of PEB, measured with three different methods (i.e., the multimethod component), could be significantly interrelated because they refer to the same behavior.

In an MTMM matrix of correlations, each variable represents a trait assessed by a particular method. In such a matrix, combinations of these variables include (a) correlations of the same trait measured with different methods; (b) correlations of different traits measured with the same method; and (c) correlations of different traits measured with different methods. Correlations of class (a) represent the convergent validity of traits, whereas those of class (b) are one set of indicators of discriminant validity for traits and also represent the effects of using a single method on the measure of those traits, whereas those of class (c) are evidence of divergent validity for both traits and methods. Thus, one would expect to find the highest correlations in (a), followed by lower correlations in (b), and the lowest possible correlations in (c) (Campbell & Fiske, 1959; Ferketich, Figueredo, & Knapp, 1991).

Although Campbell and Fiske's approach represented an advance in the assessment of construct validity, it exhibited some limitations. Among them, Widaman (1985) mentioned its lacking significance testing when evaluating correlations. That is, one had to guess when a correlation was salient enough to be considered as relevant or high. In addition, each "trait-method unit" of Campbell and Fiske's matrix was composed of trait, method, and random variance. However, using the bivariate MTMM matrix, it was not possible to analyze and separately estimate these components (Ferketich et al., 1991). Also, although method and trait were supposed to be constructs explaining variable intercorrelations, they were not modeled as latent factors in an MTMM matrix. Thus, Widaman (1985) proposed using confirmatory factor analysis of the MTMM matrix. In doing that, he produced latent traits and methods from correlations among observed variables and assessed convergent and divergent validity based on significance testing. According to this procedure, convergent validity should be obtained from high and significant factor loadings between each latent variable (trait, method) and its corresponding indicators (observed variables), whereas discriminant validity should be indicated by correlations between the latent factors, which should be lower than the factor loadings in order to indicate divergence.

Finally, some critical issues should be considered in the use of an MTMM approach. Ferketich et al. (1991) proposed the following: First, the methods used must really be different. Direct observations and self-reports are different, but true-false and multiple-choice questionnaires are not. Second, the facets should discriminate the focus trait (i.e., the specificity of the studied behavior). Third, criteria for evaluating the magnitude of similarities and differences should be specified (i.e., hypothesis testing of correlations). Fourth, correlations in the MTMM matrix must be a result of its intentional design. If a test is indeed a trait-method unit, then correlations should be attributable to the shared method variance, shared trait variance, or a combination of both.

## METHODS OF PEB RECORDING

For most researchers, self-report is a preferred method for collecting PEB data. It consists of obtaining verbal descriptions of individuals' own behavior through questionnaires, surveys, or interviews. Self-reports are widely used because of their advantage in collecting information related to varied and diverse aspects of behavior, their low cost, and their being easily obtained (Bechtel, 1990). Some self-reports provide information regarding the subjective frequency of people activities. An example of this kind of reporting includes most of the scales used in research of conservation behavior. In these, individuals point out if they *always*, *often*, *sometimes*, or *never* conserve products, instead of wasting them (see, e.g., De Young, 1990-1991; Lee et al., 1995; Margai, 1997) . Self-reporting of subjective frequency is an easy way of producing information regarding responsible environmental behavior. However, one common disadvantage is its subjectivity, because nobody knows exactly what "often" or "sometimes" means. "Often" for someone could be "sometimes" for another, or vice versa. An additional and major problem is reactivity of subjective frequency reporting, which leads people to overreport their proecological effort because this is socially desired (Geller, 1981). Also, the effect of time and memory may lead to incorrect self-reports of subjective frequency of PEB (Warriner, McDougal, & Claxton, 1984). Thus, self-reporting of subjective frequency is sometimes an inaccurate description of a person's own behavior.

An alternative method is self-report of quantity, which consists of asking people to report numerically how many times they engaged in proecological behavior or how many items they conserved. By using numbers, instead of words, sometimes it is possible to enhance memory retrieval. This kind of reporting has not been used frequently in research of PEB (exceptions are found in Margai, 1997; Oskamp, Burkhardt, Schultz, Hurin, & Zelezny, 1998). In other areas, it has been found that quantification produces a decrease in the accuracy of verbal reporting of behavior (Hartley, 1977). Corral-Verdugo, Zaragoza, & Guillén (in press) found that a less frequent reuse and recycling behavior was associated with inaccurate quantitative self-reports. However, when such behavior is salient (i.e., frequent), quantitative self-reports increase the accuracy of verbal reports.

A third method is direct observation. It consists of recording the frequency of a behavior or trace of conduct, which has been previously selected as the focus of interest. If reliable, observational data can be a valid assessment of PEB and other forms of behavior as well (Oskamp et al., 1998). Additional advantages of this method have been recognized. For example, some observations, such as the recording of physical traces, are unobtrusive and easy to use (Zeisel, 1993). A potential limitation is that possible ethical problems could result from the use of observations because they could violate people's privacy when recording traces of their behavior or their actual overt conduct. Other biases may affect observation. People's willingness to have their PEB observed may vary according to the socially desirable or undesirable connotations of that behavior. For example, although an environmentally conscious researcher may put a positive evaluation on reuse of materials, certain subjects may perceive the same behavior as a sign of poverty and therefore less socially desirable.

Therefore, every method has pros and cons, advantages and biases, that could affect the reliability and validity of measures. Thus, the use of multiple methods in combination should improve the chances that the bias of one method is mitigated by the others (Zeisel, 1993).

#### REUSE BEHAVIOR

Reuse is one of the options people have to deal with the problems of controlling and managing garbage in their communities. In this regard, the U.S. Environmental Protection Agency (1989) suggested a hierarchy of priorities to address solid waste problems. This hierarchy lists management alternatives in the following order of preference: (a) source reduction, (b) recycling or waste reduction, (c) incineration with energy recovery, (d) incineration without energy recovery, and (e) landfilling. These solutions and hierarchy could also apply to most poorer nations, such as Mexico. Source reduction and waste reduction are preferred alternatives because the solution offered by incinerators and landfills is not enough to counteract their negative side effect: more pollution. Besides, no one wants incinerators or landfills in their community (Linn, Vining, & Feeley 1994). Source reduction and waste reduction include conscious consuming, reuse, and recycling (De Young, 1990-1991; Linn et al., 1994).

Conscious consuming implies assuming a frugal lifestyle (De Young, 1990-1991) or the increase of consumer awareness directed toward avoiding purchases of non-environmentally friendly products (Linn et al., 1994). This latter option does not necessarily oppose the dominant consumerist lifestyle of industrialized countries because people would not decrease their total consumption. However, they would avoid consuming toxic or polluting products. In Third World countries, the problem of waste reduction would seem not so different from that of industrialized countries. Economic constraints and chronic financial crisis make inaccessible a lot of products that in richer nations are cheap for most families. Nonetheless, it is a fact that certain sectors of poor societies produce significant amounts of solid waste. Moreover, poorer societies are experiencing the highest rates of population growth and frequently have ineffective garbage disposition and collection services, such that Third World countries have large and growing amounts of garbage, both in their cities and in their rural regions.

Reuse, as a form of source reduction, attempts to prevent pollution by avoiding generation of waste at its source (De Young, 1990-1991).

Sometimes, reuse involves using a potentially discardable object in a different way than was originally assumed when the object was acquired/purchased (e.g., reusing a wine bottle as a flowerpot). However, sometimes reuse involves using again the reused object for its original function (e.g., a glass juice container could be used to again serve juice made for concentrate). In either case, in reuse objects are neither discarded nor reprocessed (as in recycling). Recycling, in turn, refers to treating or processing discarded objects so as to make them available for use in the original or some other form. In recycling, a special treatment is required to reconvert the recycled object. Thus, energy could be required, and although recycling saves resources it may produce pollution due to the reconversion process.

According to De Young (1990-1991), reuse and recyling practices refer to waste reduction and source reduction, respectively. Waste reduction refers to any technique that reduces the quantity of discards reaching landfills—including recycling and composting—whereas source reduction attempts to prevent pollution by avoiding the generation of waste at its source. Thus, reuse practices are source-reduction procedures and a preferred alternative for handling garbage problems.

In Mexico—where recycling programs or campaigns are scarce—reuse is a more suitable alternative to deal with garbage problems. In rural Mexican communities, housewives reuse cans, cardboard, glass, wood, and the like in order to save their scarce economic resources (Corral-Verdugo, 1996). In cities, where products are more accessible, reuse is not as common as in rural regions, however, people reuse in different ways. They use empty bottles as flower containers or store cereals or different foods in them. Reuse of clothing is a broadly accepted practice within families, which are usually larger than in the United States, so that the same clothes might be used by three or four children. Other products that urban Mexican families usually reuse are steel components of cars, bicycles, or furniture and paper and cardboard.

## STUDY GOAL

The purpose of this article is to present a study in which different kinds of reuse behavior, assessed by different methods, are contrasted. Because multimethod studies of PEB are rare, this research represents an opportunity for modeling the effect of trait and method on the variability of subjects' responses regarding their proecological behavior. It also permits us to contrast the contribution of diverse methods assessing the variance of conservation behavior, which in previous studies has been validated by convergence of independent (yet related) variables representing a trait but not a method.

Observations of reused materials (glass, clothing, and metal) as well as self-reports of the frequency and quantity of the reuse of those materials were used as methods of data collection. The three reuse behaviors were considered traits in this study. In daily use, the term *behavior* is certainly more common than *trait*, however, the latter is the conventional name identifying behaviors, facets, or characteristics to be assessed and represented in an MTMM matrix. Results of the use of the three methods were included in an MTMM matrix, and a confirmatory factor analysis of results was conducted.

## METHOD

*Subjects*. Individuals studied were 130 persons from Hermosillo, a medium-sized city (population = 600,000) in northwestern Mexico. Forty-three families were randomly selected from three stratified subsamples, representative of low ($n$ = 70), medium ($n$ = 50), and high ($n$ = 10) social class neighborhoods in the city, and three members of each family were interviewed (two adults and a juvenile between 12 and 18 years). Subjects were 40 males and 90 females; their age mean was 35.2 years ($SD$ = 14.5), and their average monthly family income was US$780.00 ($SD$ = US$582.00). The mean family size was 4.5 ($SD$ = 0.6) members. Both age and family size were normally distributed within the sample, whereas income presented a positive skewness. In Mexico, a distribution of higher income in a limited number of families is a population characteristic (Corral-Verdugo, 1996).

*Instruments*. A questionnaire investigating the reuse of clothing, glass, and metallic products practices was used. According to previous research, these materials are frequently reused in the studied community (Corral-Verdugo, 1996). Two ways of verbally referring to these practices were requested of participants. In the first way, individuals self-reported the amount of their reuse with four response options: *never, sometimes*, *often*, and *always*. This section was the subjective frequency of reuse self-report. The second included questions asking each individual to point out how many objects he or she (and no one else) had reused in his or her household. Subjects had to indicate a number as their response, thus this section was the quantity of reuse self-report. The instrument also included a section where the researchers noted the number of reused objects that they observed at the household, which corresponded to the products conserved by the individual investigated, constituting the direct nonreactive observations of reuse.

*Procedure*. Consent to participate in the study was given by an adult, even if the interviewed individual was a minor. They were told this study's objective was to investigate conservation practices and that the investigator would require their permission to observe, throughout the household, the kind and amount of conserved products they had. The amount of time invested for doing the observation depended on the number of observed objects but, on average, it took about 15 minutes. This observation was done at one point in time. Telling people in advance that their reuse products would be observed had the objective of decreasing the reactivity of self-reports, thereby enhancing the correspondence between verbal reporting and observations. Everybody asked agreed to participate in this study. Self-reports of subjective frequency for every conservation practice were obtained first, then the quantity of reuse self-reports, and finally the direct nonreactive observations of reused materials were made. Every participant was given a card where the definition of reuse was written: "Reuse means using an object, which has lost its original utility, in a similar or different way than that it previously had. For example, some people reuse empty glass bottles to store beans, rice or sugar."

*Data analysis*. Data from the qualitative self-reports were recorded for analysis: 1 = *always*, 2 = *often*, 3 = *sometimes*, and 4 = *never* were the options for responding to the question, "How often do you reuse (name of the object)?" This coding assigns lower numeric ranks to higher levels of reuse and recycling behavior. Data from quantity of reuse self-reports and direct observations were entered in the analyzed database as they were recorded, and the means of both the directly observed and self-reported quantity of reuse were computed. An MTMM matrix of correlations between results of the three methods was obtained. In addition, this matrix was used as data input for an MTMM confirmatory factor analysis.

As it is known, in a confirmatory factor analysis a number of directly measured items or variables (called manifest variables or indicators) are related to a smaller set of hypothetical constructs (called latent variables or factors) presumed to be underlying the correlations between them. For purposes of this study, this procedure is superior to traditional exploratory factor analysis in that the latter derives the multivariate constructs empirically from the correlations between manifest indicators and consequently runs the risk of capitalization on chance associations (alpha slippage) and of equivocal post hoc interpretation of factors. Instead, confirmatory factor analysis permits the theoretical specification of the latent constructs as a priori hypotheses to be tested against the correlational data. This exclusive prior assignment of each item to theoretically specified constructs reduces the number of factor loadings needed and enhances the efficiency of parameter estimation.

Confirmatory factor analysis of an MTMM matrix is an application of the above-mentioned principles in which two categories of latent variables are modeled: (a) traits, resulting from correlations of measurements of a single facet using different methods, and (b) methods, which result from correlations between measurements of different traits using the same method.

In this analysis, three traits (reuse of glass, reuse of clothing, and reuse of metal) and three methods (direct observation, self-report of subjective frequency, and self-report of quantity) were specified and tested as latent variables emerging from correlations among their corresponding observed variables. For example, the latent trait “reuse of glass” was built from interrelations among all measurements in which glass reuse was assessed: “observed reuse of glass,” “reported frequency of glass reuse,” and “reported quantity of glass reuse.” The method observation, as a latent factor, was constructed from interrelations among all measurements assessed by observation: “observed glass reuse,” “observed clothing reuse,” and “observed metal reuse.”

The convergent and divergent validity of each latent factor was assessed through significance testing of factor loadings and factor covariances ($p < .05$) in addition to the salience of such statistics (i.e., the value of the factor loading). Between-methods covariances and between-traits covariances were specified and tested, but the correlation between methods and traits was fixed to zero.

The confirmatory factor analysis was performed by using the EQS statistical software (Bentler, 1993), which provides goodness-of-fit indicators for a model of relations between variables. These indicators are chi-square, the Bentler-Bonett Comparative Fit Index (CFI ), the Bentler-Bonett Normed Fit Index (BBNFI), and the Bentler-Bonett Non-Normed Fit Index (BBNNFI). Chi-square measures the statistical goodness of fit of the correlation matrix observed to that reproduced by the factor model. A significant chi-square is therefore grounds for rejection of the factor model specified, and a nonsignificant chi-square is grounds for its tentative acceptance. With large samples, a small effect may result in a statistically significant lack of fit. The CFI, BBNFI, and BBNNFI are measures of practical goodness of fit for large sample sizes. Values greater than .90 for practical goodness-of-fit indicators are considered evidence of goodness of fit, even if significant chi-square values are obtained (Bentler, 1993; Byrne, 1994).

**TABLE 1**
**Multitrait-Multimethod (MTMM) Correlation Matrix, Contrasting Results From Assessing Three Different Kinds of Reuse With Three Different Methods**

| | *Observations* | | | *Frequency* | | | *Quantity* | | |
|---|---|---|---|---|---|---|---|---|---|
| | *OG* | *OC* | *OM* | *FG* | *FC* | *FM* | *QG* | *QC* | *QM* |
| OG | | | | | | | | | |
| OC | *.17* | | | | | | | | |
| OM | *.28* | *.38* | | | | | | | |
| FG | **.32** | .12 | .21 | | | | | | |
| FC | .19 | **.35** | .05 | *.29* | | | | | |
| FM | .13 | .11 | **.47** | *.21* | *.19* | | | | |
| QG | **.56** | .09 | .37 | **.45** | .14 | .21 | | | |
| QC | .03 | **.67** | .11 | .10 | **.41** | .03 | *.20* | | |
| QM | .15 | .09 | **.66** | .14 | .05 | **.54** | *.38* | *.07* | |

NOTE: Methods—O = Observation, F = Frequency of self-report, Q = Quantity of self-report. Traits—G = Reuse of glass, C = Reuse of clothing, M = Reuse of metal. Italicized numbers indicate that they are different traits, measured with a single method. Bolded numbers indicate that they are the same trait, measured with different methods.

## RESULTS

The amounts of observed and self-reported quantity of reuse revealed that clothing was the most conserved product at households (means of 8.1 and 8.6 reused items, respectively), followed by glass (5.2 and 4.0), and steel objects (2.8 and 2.0). Correlations among all three measures are presented in Table 1, which shows the MTMM matrix obtained in the study. The bolded diagonals in the table represent heteromethod diagonals. These contain the correlations among each of the traits measured by two different pairs of the methods. As expected, the heteromethod diagonals produced the highest correlations in the MMTM matrix, yet differences were found among the three methods: The correlations among observations of the reuse of glass, clothing, and metal where higher than the correlations among the quantity self-reports of reuse, which, in turn, were higher than those found among the subjective frequency self-reports of such conservation activities.

Table 1 also exhibits three italicized triangles representing monomethod blocks, which are formed by correlations between different traits measured by the same method. These correlations were not as high as the ones found in the heteromethod diagonals, however, they were higher than the remaining correlations representing heteromethod-heterotrait blocks.

Because this original approach lacks significance testing when evaluating correlations (Widaman, 1985), a confirmatory factor analysis was conducted in order to obtain statistical indicators of difference between traits and methods.

Figure 1 shows the results of this confirmatory analysis. In general, high and significant lambda weights (i.e., factor loadings) were found between each latent method and trait and its corresponding observed variable (the only exception was the nonsignificant factor loading between the factor “quantity self-report” and the observed variable “quantity of reported metal reuse”). The covariances among the three methods and among the three traits were lower than the values of the lambda weights (among factors and their corresponding indicators). Yet, one factor loading was slightly higher than the covariance between the observation and quantity self-report methods, and two lambda weights exceeded the value of the “reuse of glass”–“reuse of metal” covariance.

## DISCUSSION

Evidence of construct validity for the measurement of reuse practices was found in both the MTMM matrix and its confirmatory factor analysis. The monotrait-heteromethod diagonals produced the highest correlations within the MTMM matrix, followed by the heterotrait-monomethod triangles and the heterotrait-heteromethod blocks. Because independent methods converged in the assessment of a single trait, evidence of convergent construct validity seemed to emerge from our data. The moderate correlations of the heterotrait-monomethod triangles indicated a bias of using the same method in measuring different traits, whereas the lowest correlations of the heterotrait-heteromethod blocks seemed to show evidence of discriminant or divergent validity.

High correlations between quantity of reuse self-reports and direct observations of reuse were found in the monotrait-heteromethod diagonals. Those were higher than the ones found between subjective frequency and quantity of reuse self-reports and those between observations and the subjective frequency of reuse self-reports. This would indicate that the numerical report was closer to observations than to a subjective frequency report. It is possible that the use of numbers made quantity of reuse reporting more like the quantification of observed behavior. This is despite the greater methodological similarity between self-reports of the subjective frequency and quantity of reuse. Because the classical (subjective frequency) reporting uses nonnumerical units (*always* to *never*), this kind of self-report might be reflecting a self-perception rather than an account of behavior (Corral-Verdugo, 1997). It

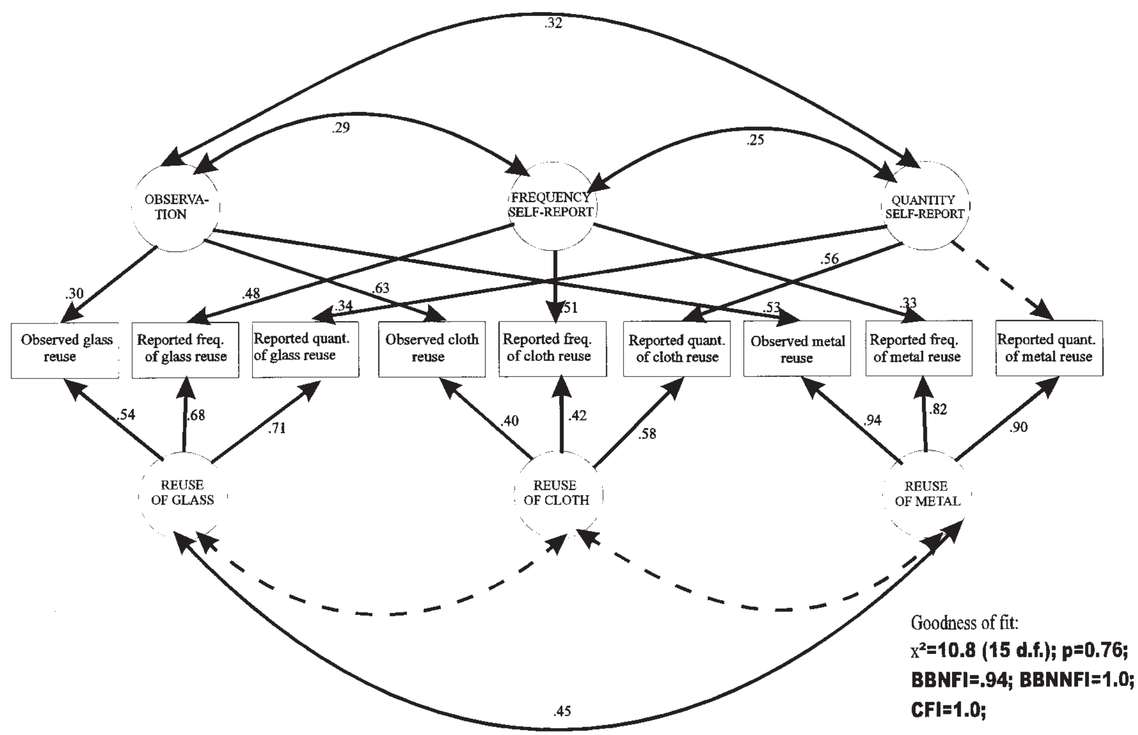




**Figure 1: Confirmatory Factor Analysis Multitrait-Multimethod Correlation Matrix**

might also reflect subjective norms regarding the amount of reuse behavior that the subject believes to be prevalent in his or her community.

The confirmatory factor analysis confirmed these results, adding significance testing to the assessment of validity and permitting the determination of construct validity by means of the specification of latent factors. As expected, all (but one) factor loadings connecting constructs and their corresponding indicators were statistically significant ($p < .05$), which reinforces the assumption of convergent construct validity. Because the covariances between methods and between traits were, in general, lower than those factor loadings, there are reasons for also concluding divergent, or discriminant, construct validity. In summary, these results imply that the measures of any kind of reuse were more strongly related regardless of which method was used to collect the data. Despite these results, the between-methods covariances were significant, indicating that, at least partially, the used methods were similar in their assessing reuse behavior. However, because those correlations ranked from .25 to .32, it seems that their difference was higher than their covariation. According to Ferketich et al. (1991), to produce a meaningful contrast in an MTMM matrix the used methods must really be different, which seemed to be the case in this study.

In the confirmatory factor analysis, the value of covariances between traits was only significant for the association between reuse of glass and reuse of metal: Those who reused metal did not necessarily reuse clothing, and those who reused clothing did not necessarily reuse glass. These results replicate previous reports indicating that particular forms of conservation behavior are not necessarily interrelated (Corral-Verdugo, 1996; Daneshvary et al., 1998; Lee et al., 1995).

Some of our findings are in apparent conflict with previous results that show a discrepancy between self-reports and observations of conservation behavior (Corral-Verdugo, 1997; Corral-Verdugo et al., 1995; Cote, 1984). Although in this study the correlations between observations and the subjective frequency and the quantity versions of self-report were from low to moderate, they were significant and higher than those reported in the previous reports. It is possible that such higher correlation might be due to particular conditions of this study: First, people were notified (and asked) that observations of their behavioral traces would be conducted and, second, numerical reports were collected of just frequent (salient) reuse activities. The first situation could have helped in decreasing the self-report reactivity and thus made people to produce a description of their own behavior closer to observed reuse. It is likely that people, being alerted that the products of their actions were being observed, made descriptions of their behavior more genuine than otherwise. The second situation could have had an impact by increasing memory retrieval or reuse actions. Individuals remembered better maybe because what they reused was relevant or salient for them. Salience of information is an important feature of a good retrieval of autobiographical memory (Wagenaar, 1986). If memory plays a role in explaining the inaccuracy of self-reports (Warriner et al., 1984), then making salient the information to self-reporting would increase its accuracy. A limitation of this study is that it fails in responding to the question: How good are the findings of this article in assessing validity of unattended and unobserved behavior? Because we did not assess responses of unadverted subjects (i.e., subjects that were not told they would be observed), it is difficult to reveal the extent of the real measurement error of self-reports. This limitation should be addressed in future studies.

This study's results seem to demonstrate that it is possible to use multiple facets of conservation behavior, together with multiple methods, to assess the validity of PEB measures. As the description of our design shows, most—if not all—of the issues considered as critical (Ferketich et al., 1991) in the use of the MTMM approach were accomplished in this study: The methods were independent, traits significantly discriminated the intended conservation practice, hypothesis testing of correlations was obtained, and the results of the MTMM matrix were found as expected.

In addition, we suggest an additional condition for using and interpreting an MTMM matrix, in the context of conservation behavior: Biases of methods should be minimized prior to starting the research so as not to expect that the combined use of multiple measures will cancel each other's method bias after the research is conducted; this will produce an optimal convergence. In this study, we attempted to reduce reactivity of self-reports by making people aware of their being observed and tried to increase the accuracy of quantitative self-reporting by asking individuals about salient reuse behaviors. These manipulations apparently made possible a higher covariance between self-reports and observations.

## REFERENCES

Bechtel, R. B. (1990). The ubiquitous world of paper and pencil tests. In R. B. Bechtel, R. W. Marans, & W. Michelson (Eds.), *Methods in environmental and behavioral research* (pp. 82-119). Malabar, FL: Krieger.

Bentler, P. (1993). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.

Berger, I. E. (1997). The demographics of recycling and the structure of environmental behavior. *Environment and Behavior, 29*, 515-531.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows*. Thousand Oaks, CA: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Corral-Verdugo, V. (1996). A structural model of reuse and recycling in Mexico. *Environment and Behavior*, *28*, 665-696.

Corral-Verdugo, V. (1997). Dual "realities" of conservation behavior: Self-reports vs observations of re-use and recycling behavior. *Journal of Environmental Psychology*, *17*, 135-145.

Corral-Verdugo, V., Bernache, G., Encinas, L., & Garibaldi, L. C. (1995). A comparison of two measures of reuse and recycling behavior: Self-reports and material culture. *Journal of Environmental Systems*, *23*, 313-327.

Corral-Verdugo, V., Zaragoza, F., & Guillén, A. (in press). The effect of quantification on the accuracy of proenvironmental behavior self-reports. *Journal of Environmental Systems*.

Cote, J. A. (1984). Use of household refuse analysis to measure usual and period-specific food consumption. *American Behavioral Scientist*, *28*, 129-138.

Daneshvary, N., Daneshvary, R., & Schwer, R. K. (1998). Solid-waste recycling behavior and support for curbside textile recycling. *Environment and Behavior, 30*, 144-161.

De Young, R. (1990-1991). Some psychological aspects of living lightly: Desired lifestyle patterns and conservation behavior. *Journal of Environmental Systems*, *20*, 215-227.

Ferketich, S. L., Figueredo, A. J., & Knapp, T. R. (1991). Focus on psychometrics. The multitrait-multimethod approach to construct validity. *Research in Nursing & Health*, *14*, 315-320.

Geller, E. S. (1981). Evaluating energy conservation programs: Is verbal report enough? *Journal of Consumer Research*, *8*, 331-335.

Hartley, S. (1977). Subjective time estimates of work tasks by office workers. *Journal of Occupational Psychology*, *50*, 23-36.

Lee, Y. J., De Young, R., & Marans, R. W. (1995). Factors influencing individual recycling behavior in office settings. *Environment and Behavior, 27*, 380-403.

Linn, N., Vining, J., & Feeley, P. A. (1994). Toward a sustainable society: Waste minimization through environmentally conscious consuming. *Journal of Applied Social Psychology, 24*, 1550-1572.

Margai, F. L. (1997). Analyzing changes in waste reduction behavior in a low-income urban community following a public outreach program. *Environment and Behavior, 29*, 769-792.

Marutza, V. R. (1977). *Applying norm-referenced and criterion-referenced measurement in education*. New York: Allyn & Bacon.

Oskamp, S., Burkhardt, R. L., Schultz, P. W., Hurin, S., & Zelezny, L. (1998). Predicting three dimensions of residential curbside recycling: An observational study. *Journal of Environmental Education*, *29*, 37-42.

Rathje, W. (1989). The three faces of garbage—Measurements, perceptions, behaviors. *Journal of Management and Technology*, *17*, 61-65.

U. S. Environmental Protection Agency. (1989). *The solid waste dilemma. An agenda for action*. Washington, DC: Government Printing Office.

Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, *18*, 225-252.

Warriner, G. K., McDougal, G. H., & Claxton, J. D. (1984). Any data or none at all? Living with inaccuracies in self-reports of residential energy consumption. *Environment and Behavior*, *16*, 503-526.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, *9*, 1-26.

Zeisel, J. (1993). *Inquiry by design: Tools for environment-behavior research*. Cambridge, MA: Cambridge University Press.