

# The Dangers of Poor Construct Conceptualization

**Scott B. MacKenzie**

*Indiana University*

For want of a nail, the shoe was lost.  
For want of the shoe, the horse was lost.  
For want of the horse, the rider was lost.  
For want of the rider, the battle was lost.  
For want of the battle, the kingdom was lost.  
And all for the want of a nail.

In some respects, this nursery rhyme summarizes the point of my brief essay on what I believe is one of the major problems in the manuscripts I review. Only instead of a missing nail, the problem is poor construct conceptualization. I chose this issue because it is one that has serious consequences for the validity of research in the discipline, but it is something that authors tend to overlook.

As noted by Cook and Campbell (1979), there are four main questions that should be considered when evaluating the validity of a study's findings. First, what are the particular cause-and-effect constructs involved? This is a question of *construct validity*. Second, is there a relationship between the presumed cause and effect? This question relates to *statistical conclusion validity*. Third, is the relationship causal? This question addresses the *internal validity* of the research and concerns whether the conditions for establishing causal priority of the causal variable over the effect variable have been established and whether the design of the study renders any rival explanations of the observed effect implausible. Fourth, how generalizable is this relationship across persons, settings, and times? Here the focus is on *external validity*.

Even though all four types of validity are important, and indeed essential, authors often give far more attention to some of these questions than to others. Most authors give a great deal of consideration to questions of internal and external validity (as they should), but far less to issues related to construct and statistical conclusion validity.

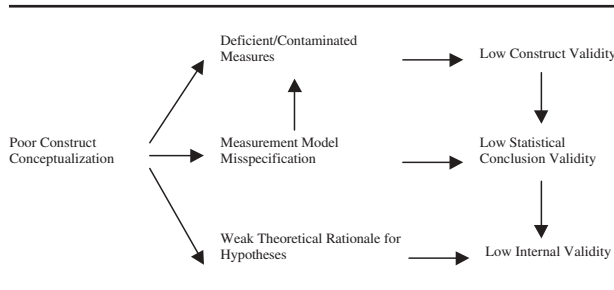
However, as a reviewer, I often find that poor construct and statistical conclusion validity can make any thoughts of internal or external validity a fantasy. In my experience, the most common cause of this disappointing state of affairs is the failure to adequately specify the conceptual meaning of the study's focal constructs. This fundamental problem triggers a sequence of events (see Figure 1) that undermines construct validity (primarily due to measure deficiency), statistical conclusion validity (due to the biasing effects of measurement model misspecification), and ultimately internal validity (due to a combination of factors).

## VICIOUS CYCLE THAT CAUSES THE PROBLEM

The downward spiral for many manuscripts begins with the failure to adequately define the focal construct(s) of the study. A considerable number of pages in the front end of most manuscripts are devoted to thoroughly reviewing the diverse, and often conflicting, conceptualizations of the focal construct(s) found in the research literature. In many instances, what emerges from this review is a rich—but not altogether clear—sense of the conceptual meaning of the focal construct(s). At this point, successful authors will wrestle with these difficult conceptual issues and attempt to synthesize the alternative conceptualizations. However, too many authors abdicate their responsibility to do this and instead move on to their discussion of the hypotheses. As a result, they never develop a clear, concise conceptual definition of the focal construct(s). This fundamental failure produces a series of subsequent problems.

First, because the focal construct was never adequately defined, it is difficult to develop measures that faithfully represent its domain. Researchers use multiple measures of constructs to (a) better capture the full domain of complex constructs (Churchill 1979), (b) unconfound the method of measurement from the construct of interest (Campbell and Fiske 1959), and (c) enhance the reliability

**FIGURE 1**  
**The Consequences of Poor Construct Conceptualization**



of measurement (Nunnally 1978). However, if you only have a vague idea of what you are trying to measure, it is easy for the measures to become contaminated by unrelated factors and/or for them to underrepresent the conceptual domain. Either way, it leads to low construct validity.

Second, the failure to clearly define the focal construct makes it difficult to correctly specify how the construct should relate to its measures. Two different types of measurement relations have been identified in the literature (cf. Bagozzi and Fornell 1982; Bollen and Lennox 1991). *Reflective* indicator measurement models posit that the direction of causality flows from the construct to the measures, whereas *formative* indicator measurement models posit that causality flows from the measures to the construct. Obviously, when a construct is not clearly defined, it is more difficult to determine which of these two fundamentally different types of measurement relations is the most appropriate. This is important because measurement model misspecification can undermine both construct validity and statistical conclusion validity.

Measurement model misspecification undermines statistical conclusion validity by biasing estimates of the structural relationships between constructs. Indeed, a recent research study that I was part of (Jarvis, MacKenzie, and Podsakoff forthcoming) provides a vivid illustration of the biasing effects of measurement model misspecification and its pervasiveness. We reviewed 178 articles published in the *Journal of Marketing Research*, the *Journal of Marketing*, *Marketing Science*, and the *Journal of Consumer Research* over a 24-year period and found that approximately 29 percent of the 1,192 latent constructs found in these articles had incorrectly specified measurement models. In nearly all of these cases, treating formative indicators as if they were reflective caused the misspecification. The effect of this type of measurement model misspecification was examined with a Monte Carlo simulation, and the results showed that it could bias estimates of the relationships between constructs by 300 to 500 percent, depending on the magnitudes of the

correlations between the measures. Obviously, a bias this large could have a huge impact on the validity of any statistical inferences made about whether the constructs are related.

Measurement model misspecification can also undermine construct validity in two main ways. First, it misrepresents the fundamental nature of the relations between the measures and the focal construct. Second, when the misspecification arises from treating formative indicators as if they were reflective, the use of standard scale refinement procedures can result in the dropping of measures that are essential to the domain of the construct. For example, it is standard practice (cf. Churchill 1979) when purifying a new scale to drop the items with the lowest item-to-total correlations or the lowest factor loadings. This is good advice when the measures are reflective of the underlying construct. However, when the measures are formative in nature, dropping items with the lowest item-to-total correlation will result in the removal of precisely those items that would most alter the empirical meaning of the composite latent construct. This can increase the likelihood that a unique part of the conceptual domain will be omitted, make the measures deficient, and undermine construct validity.

The third problem caused by inadequate construct definitions is that they undermine the credibility of a study's hypotheses. Without well-developed construct definitions, it is impossible to develop a coherent theory because constructs are the building blocks of theory. One simply cannot develop a meaningful theoretical rationale for why Construct A should be related to Construct B if the exact meaning of each of these two constructs has not been established. Indeed, when a construct is poorly defined (or not defined at all), there is a tendency for its meaning to be stretched in a slightly different conceptual direction each time a theoretical rationale is provided in support of one of the hypotheses about it. For example, when explaining why "perceived risk" influences an entrepreneur's investment decisions, the theoretical rationale might implicitly focus on "financial" risk, but when explaining how "perceived risk" influences decisions about becoming an entrepreneur in the first place, the focus may subtly shift to "psychological/social" risk. This leaves reviewers with the impression that the hypotheses are inconsistent and that their theoretical foundation is built on shifting sand, thus undermining perceptions of internal validity.

Thus, poor construct conceptualization makes it difficult to avoid measure deficiency and contamination, often leads to measurement model misspecification, and raises doubts about the credibility of the hypotheses. This undermines construct validity, statistical conclusion validity, and ultimately internal validity and usually results in the rejection of the manuscript. So it really *is* true that "for want of a definition the construct was lost" and so on.

## RECOMMENDATIONS FOR BREAKING THIS DESTRUCTIVE CYCLE

### Carefully Define the Constructs of Interest

Good definitions should (a) specify the construct's conceptual theme, (b) in unambiguous terms, (c) in a manner that is consistent with prior research, and that (d) clearly distinguishes it from related constructs. A good definition should also specify the extent to which values of the construct are expected to differ across cases, conditions, and time. Also, when the construct is conceptualized as being multidimensional, the relations between the subdimensions and the superordinate construct should be specified (cf. Jarvis, MacKenzie, and Podsakoff forthcoming).

One common mistake is to define a construct as the result of, and/or the cause of, some other construct. For example, to define a construct as "*beliefs formed by early childhood experiences that guide consumption behaviors*" is not helpful because this definition does not specify the nature of the "beliefs" included in the construct. Defining a construct in this manner also makes it impossible to empirically test the proposed theoretical linkages between the construct and the specific antecedents and consequences mentioned, because these relationships are assumed to be true by definition. Therefore, one should not define a construct solely in terms of its antecedents or consequences.

Another common type of pseudo definition noted by Summers (2001) involves merely giving examples of what is included in a construct. For example, rather than defining the construct, an author will sometimes report that "Construct A includes things such as x, y and z" (usually this list of exemplars bears a strong resemblance to the list of measures in the Method section). The problem with defining a construct solely by exemplars is that there is no way to know whether the exemplars provide a complete listing of the construct's domain and/or whether new exemplars should be excluded from the construct's domain. Therefore, a good conceptual definition should specify the underlying theme that ties the exemplars together. Otherwise, it is impossible to evaluate whether the measures adequately tap the domain of the construct and/or are contaminated by extraneous factors. It will also be difficult for others to develop new "maximally dissimilar" measures if the conceptual principle linking the measures has not been articulated.

Although these recommendations might seem obvious, following them often proves to be difficult for many authors. If my own experience is any indication, when you are having difficulty defining a construct, it is usually because you have not really decided what you want the construct to represent. People often mistakenly think they are having a writing problem (i.e., finding the proper

words to represent their ideas) when the real problem is that their ideas are not well developed.

### Develop Measures That Adequately Represent the Construct(s) of Interest

When developing measures of a construct, the goal is to make sure that (a) all key aspects of the conceptual definition are reflected in the measures, (b) the items are not contaminated by the inclusion of things that are not part of the conceptual domain, and (c) the items are properly worded (e.g., unambiguous, specific, no leading questions, no double-barreled questions, etc.).

Once an initial set of items tapping the key aspects of the conceptual domain has been generated, you should try to critically evaluate them—always remembering that if you don't, the reviewers will! A reasonable procedure for evaluating the measures might begin with the creation of a table that lists each essential aspect of the conceptual domain in a separate column and each measure in a separate row. Next, you would carefully evaluate which essential aspects of the conceptual domain are tapped by each of the measures listed. Following this, you might add a final column and list any confounding constructs that may be tapped by the measure that are not part of the focal construct's conceptual domain. Finally, you could add rows for any new items needed to tap essential aspects of the conceptual domain that have been neglected and/or for items that need to be reworded in order to avoid conceptual confounds.

### Think Carefully About the Relations Between the Constructs and Measures

This is one of the most frequently neglected steps. Far too many authors (and reviewers) assume that anytime a construct has multiple measures, it should be modeled as having reflective indicators. However, this is not always conceptually appropriate (cf. Bagozzi and Fornell 1982; Bollen and Lennox 1991; Diamantopoulos and Winklhofer 2001). Researchers ought to give as much thought to their hypotheses about measurement relations as they do to their hypotheses about structural relations. In a recent article (Jarvis et al. forthcoming), my coauthors and I suggest that constructs should be modeled as having formative indicators if (a) the indicators are viewed as defining characteristics of the construct rather than manifestations of it, (b) changes in the indicators are expected to cause changes in the construct, (c) changes in the construct are not expected to cause changes in the indicators, (d) the indicators do not necessarily share a common theme, (e) eliminating an indicator may alter the conceptual domain of the construct, (f) a change in the value of one of the indicators is not necessarily expected to be

associated with a change in all of the other indicators, and (g) the indicators are not expected to have the same antecedents and consequences. If the opposite tends to be true, the construct should be modeled as having reflective indicators.

In addition, if the construct is conceptualized at a fairly abstract level and is expected to have several distinct subdimensions, it is important to think about not only the relationships between the measures and these first-order subdimensions but also about the relationships between the first-order subdimensions and the second-order construct they measure. My colleagues and I (Jarvis et al. forthcoming) provide several examples of marketing constructs conceptualized at this level of abstraction (e.g., market orientation, trust, helping behavior, perceived risk, etc.) and call attention to the fact that the measurement model relating the first-order subdimensions to the measures need not be the same as the measurement model relating the second-order construct to its first-order subdimensions. Although one could reasonably argue that all constructs should be unidimensional, our review suggests that such a view is often inconsistent with the way constructs are defined in the marketing literature. So as a practical matter, this is something authors should think about carefully.

### **Defend the Construct Domain and Insist on the Conceptually Appropriate Measurement Model**

Do not sacrifice construct validity at the altar of internal consistency reliability. Although this is good advice regardless of whether the measures are reflective or formative, it is particularly important to remember when your construct has formative measures, because formative indicator measurement models do *not* imply high levels of internal consistency reliability (Bollen and Lennox 1991). Therefore, when your measures are formative, it is important to resist the temptation to delete items as a means of improving Cronbach's alpha (internal consistency reliability).

Following this advice may be difficult if the temptation comes from a reviewer in the form of a recommendation. However, you must be vigilant, because the likelihood of inappropriately restricting the domain of the construct and threatening construct validity tends to be greater when formative indicators are eliminated than when reflective indicators (of a unidimensional construct) are eliminated (cf.

Bollen and Lennox 1991). The best way to avoid this unpleasant situation is to head it off by carefully discussing the construct domain and the hypothesized relations between the construct and its measures, and explicitly noting the implications of your measurement model for how it should be evaluated. If internal consistency reliability is irrelevant, consider providing other evidence of reliability (e.g., item test-retest reliability) and/or empirically examining the sensitivity of your findings to different assumptions of reliability. In other words, try to gently educate the reviewers of your manuscript as a means of avoiding problems in the review process, always remembering that your arguments will be more convincing if they come in advance rather than in response to a reviewer's criticisms.

In conclusion, the problems of poor construct validity and statistical conclusion validity that plague many manuscripts can be minimized if you carefully define the focal constructs in your research, make sure that your measures fully represent them, correctly specify the relations between the measures and constructs in your measurement model, and stick to it. I believe that following this advice will greatly improve your chances of publication success—perhaps more than anything else I might recommend.

### **REFERENCES**

- Bagozzi, Richard P. and Claes Fornell. 1982. "Theoretical Concepts, Measurement, and Meaning." In *A Second Generation of Multivariate Analysis*, Vol. 2. Ed. Claes Fornell. New York: Praeger, 24-38.
- Bollen, Kenneth A. and Richard Lennox. 1991. "Conventional Wisdom on Measurement: A Structural Equation Perspective." *Psychological Bulletin* 110:305-314.
- Campbell, Donald T. and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81-105.
- Churchill, Gilbert A., Jr. 1979. "A Paradigm for Developing Better Measures of Marketing Constructs." *Journal of Marketing Research* 16 (February): 64-73.
- Cook, Thomas D. and Donald T. Campbell. 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Diamantopoulos, Adamantios and Heidi M. Winklhofer. 2001. "Index Construction With Formative Indicators: An Alternative to Scale Development." *Journal of Marketing Research* 38 (May): 269-277.
- Jarvis, Cheryl Burke, Scott B. MacKenzie, and Philip M. Podsakoff. Forthcoming. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research." *Journal of Consumer Research* 30 (2).
- Nunnally, Jum C. 1978. *Psychometric Theory*. New York: McGraw-Hill.
- Summers, John O. 2001. "Guidelines for Conducting Research and Publishing in Marketing: From Conceptualization through the Review Process." *Journal of the Academy of Marketing Science* 29 (4): 405-415.