FORMATIIVISTA MITTAAMISTA EI OLE OLEMASSA, MUTTA FORMATIIIVISET MALLIT VOIVAT OLLA HYÖDYLLISIÄ

Mikko Rönkkö
Jyväskylän yliopiston kauppakorkeakoulu

mikko.ronkko@jyu.fi
http://tiny.cc/formativemeasurement
In Defense of Causal–Formative Indicators: A Minority Report

Kenneth A. Bollen
University of North Carolina, Chapel Hill

Adamantios Diamantopoulos
University of Vienna

Causal–formative indicators directly affect their corresponding latent variable. They run counter to the predominant view that indicators depend on latent variables and are thus often controversial. If present, such indicators have serious implications for factor analysis, reliability theory, item response theory, structural equation models, and most measurement approaches that are based on reflective or effect indicators. *Psychological Methods* has published a number of influential articles on causal and formative indicators as well as launching the first major backlash against them. This article examines 7 common criticisms of these indicators distilled from the literature: (a) A construct measured with “formative” indicators does not exist independently of its indicators; (b) Such indicators are causes rather than measures; (c) They imply multiple dimensions to a construct and this is a liability; (d) They are assumed to be error-free, which is unrealistic; (e) They are inherently subject to interpretational confounding; (f) They fail proportionality constraints; and (g) Their coefficients should be set in advance and not estimated. We summarize each of these criticisms and point out the flaws in the logic and evidence marshaled in their support. The most common problems are not distinguishing between what we call causal–formative and composite–formative indicators, tautological fallacies, and highlighting issues that are common to all indicators, but presenting them as special problems of causal–formative indicators. We conclude that measurement theory needs (a) to incorporate these types of indicators, and (b) to better understand their similarities to and differences from traditional indicators.

**Keywords:** causal indicators, formative indicators, composites, measurement models, structural equation models
Formative measurement
Formative measurement?

- Measures cause the construct
- A CEO responding positively to a survey causes the firm to be innovative
- How you respond to questions about your income, education, etc. defines your socio economic status

... Formative indicators are not measures at all.
—Lee, Cadogan, and Chamberlain (2013, p. 12)

Disciplines such as IS should consider avoiding formative measurement ... editors should question and perhaps even temporarily discourage the use of formative measurement.
—Hardin and Marcoulides (2011, p. 761)

The shortcomings of formative measurement lead to the inexorable conclusion that formative measurement models should be abandoned.
—Edwards (2011, p. 383)

# Guidelines for using formative measurement

## Table 1

*Conceptual Checks and Empirical Test for Determining Whether Indicators Are Causal–Formative or Reflective (Effect)*

<table>
<thead>
<tr>
<th>Conceptual check</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mental experiment</strong>&lt;br&gt;(see Bollen, 1989; Bollen &amp; Lennox, 1991)</td>
<td>1. <em>Reflective (effect) indicators</em>: Imagine a change in latent variable net of other indicator influences. Will this lead to a change in the indicators?&lt;br&gt;2. <em>Causal–formative indicators</em>: Imagine a change in the indicator net of other latent variable influences. Will this lead to a change in the latent variable?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Empirical test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vanishing tetrads test</strong>&lt;br&gt;(see Bollen &amp; Ting, 2000)</td>
<td>1. Based on hypotheses and mental experiments, formulate one or more measurement models with causal–formative indicators, reflective (effect) indicators, or some combination.&lt;br&gt;2. Calculate chi-square test for each model and assess fit.&lt;br&gt;3. Compare any of the competing models that are nested in their vanishing tetrads.&lt;br&gt;4. Choose the hypothesized model with the best fit.</td>
</tr>
</tbody>
</table>

---

Items as causes of constructs

Third, I hypothesize that the conflation of item scores with the attributes that one interprets the items as assessing poses a third conceptual impediment to progress. Consider the following characterization: “That is, a change in SES is not expected to lead to a change in each of these indicators. Rather, a change in one or more of these indicators leads to a change in SES” (Bollen, 1984, p. 380). Presumably, no one means to suggest that one can improve one’s SES by changing one’s answers to three survey items. Clearly, then, one must take Bollen to refer here to the attributes that the items assess. However, scales do not comprise attributes, they comprise items. Likewise, item scores determine scale scores. Thus, formative measurement theory needs to better maintain this distinction in order to provide a clear-headed account of formative scales.
Measurement and modeling

Formative models might be useful but are not measurement (Markus & Borsboom, 2013, Chapter 6, Rhemtulla et al 2015)


Alternatives to formative measurement models

Figure 5. Model that replaces formative measures with facet constructs and single reflective measures

Alternatives to formative measurement models

Figure 5. Model that replaces formative measures with facet constructs and single reflective measures


Alternatives to formative measurement models

Figure 6. Model that replaces formative measures with facet constructs and multiple reflective measures


Summary of formative measurement

• Formative measurement does not exist
  • Items do not cause the construct
  • If they did, this would be easy to demonstrate by manipulating the measures

• Formative models (indices) can be useful
  • Items that go into the index must be validated outside the index
  • An index must be justified
Statistical issues in formative models
Identification of formative models

Formative LV receives its identify and meaning from the "reflective" measures


Composite model and weight calculation

Weights are defined to maximize $\beta$
- Possible positive bias
- Full mediation assumed
- Interpretational confounding

Solution: defined the weights based on theory

Indices

Multiple item indices
Scale and non-scale variables

1. **How well do the following statements describe your firm?**

   Our firm is very innovative ................................................................. 1 2 3 4 5
   We are the technological leader in our industry ................................... 1 2 3 4 5
   We are often the first to bring new product concepts to markets .......... 1 2 3 4 5

**Unidimensional scale**
- Items measure the same quantity
- Items are highly correlated
- If it were not for random error (unreliability), the items would be perfectly correlated

**Non-scale variables**
- Items measure distinct quantities
- Items may not be highly correlated
- Example: beer, wine, and hard alcohol as sources of alcohol
- Example: various alternative ways that a company can do supply chain redesign
Perhaps an example at this point will help clarify some of the issues raised concerning multicollinearity. Suppose we are interested in estimating the effect of various school expenditure categories on student performance. It is likely that expenditures on teacher salaries, instructional materials, athletics, and so on are highly correlated: Wealthier schools tend to spend more on everything, and poorer schools spend less on everything. Not surprisingly, it can be difficult to estimate the effect of any particular expenditure category on student performance when there is little variation in one category that cannot largely be explained by variations in the other expenditure categories (this leads to high \( R^2 \) for each of the expenditure variables). Such multicollinearity problems can be mitigated by collecting more data, but in a sense we have imposed the problem on ourselves: we are asking questions that may be too subtle for the available data to answer with any precision. We can probably do much better by changing the scope of the analysis and lumping all expenditure categories together, since we would no longer be trying to estimate the partial effect of each separate category.
Index as dependent and independent variable

Define an index:
\[ C = x_1 + x_2 + x_3 \]

Independent variable:
\[ y = \beta_0 + \beta_1 C + u \]
\[ y = \beta_0 + \beta_1 x_1 + \beta_1 x_2 + \beta_1 x_3 + u \]

Dependent variable:
\[ C = \beta_0 + \beta_1 z_1 + u \]
\[ x_1 = \beta_{01} + \beta_1 z_1 + u \]
\[ x_2 = \beta_{02} + \beta_1 z_1 + u \]
\[ x_3 = \beta_{03} + \beta_1 z_1 + u \]

Summary of indices

1. Items that form the index must be validated outside index construction

2. The index needs to be justified
   A. Different sources of a quantity
   B. Alternative ways that the behavior can manifest

3. Justify the use of index over separate items

4. Set the weights on conceptual grounds
   A. When in doubt, use equal weights

Probably a good idea to avoid the term “formative”
KYSYMYSIÄ JA KESKUSTELUA

mikko.ronkko@jyu.fi
http://tiny.cc/formativemeasurement