Commentary on "Common Method Bias: Nature, Causes, and Procedural Remedies"

Madhubalan Viswanathan, University of Illinois at Urbana-Champaign
Ujwal Kayande, Melbourne Business School

Available at: https://works.bepress.com/ujwal_kayande/2/
Commentary on “Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies”

Madhubalan Viswanathan a,∗, 1, Ujwal Kayande b, 1

Abstract
Common method bias is a potentially serious methodological problem in research in marketing. Several statistical remedies have been proposed in the literature, and used by academic researchers. MacKenzie and Podsakoff (2012) identify the causes of common method bias, and then provide a set of procedural remedies that might prevent the occurrence of the problem. In this commentary, we expand on their contribution by articulating the different types of measurement error that could occur in survey research, how a procedural remedy might simultaneously affect more than one type of error, and how common method bias might manifest itself in the domain of stimulus-centered measures.

Keywords: Common method bias; Measurement error; Procedural remedies; Stimulus-centered measures

Introduction
In an important article, MacKenzie and Podsakoff (2012, hereafter M&P) identify the causes of common method bias, and present a comprehensive set of procedural remedies to reduce the possibility of common method bias in survey research. We applaud them for their continued, and influential, work in this area, and for guiding survey researchers across disciplines on how to improve the quality of empirical research. Our commentary will use their article as a foundation to delve into different arenas of procedural remedies for common method bias. This all-important discussion is very timely because the arena is fraught with implicit assumptions as well as implicit knowledge. A parallel stream of survey research methodology has examined the effects of a variety of factors on responses in survey research, emphasizing the estimation of accurate means (e.g., Sudman, Bradburn, and Schwarz 1996). A similar emphasis is needed in academic research in marketing and the social sciences that focuses on the accurate estimation of relationships between variables. M&P provide that emphasis, and our aim is to reinforce their views.

We focus our comments in three areas: (1) discussing a set of terms to dissect measurement error that provides the level of nuance needed in this area, (2) covering a variety of potential procedural remedies that expand on the discussion in the M&P article, and (3) applying this discussion to the domain of stimulus-centered measures, which are particularly relevant for retailing research.

Relevant types of measurement error
We draw from Viswanathan (2005, 2008, 2010) to lay out the different types of measurement error, specifically systematic error. Additive systematic error represents constant deviations (e.g., extreme means). Such error has been argued to be less problematic for academic research, which emphasizes relationships between constructs, when compared to public opinion research which is primarily concerned with estimating means of variables. Additive systematic error matters only to the extent that it may decrease item variance, thereby reducing covariation with other items (Viswanathan 2005). However, academic research is primarily concerned with correlation, which is influenced by the relative standing of values on variables (Nunnally 1978). In this sense, there is a second type of systematic error that matters more, discussed next.

Correlational systematic error is reflected in responses varying beyond true differences in the measured construct...
consistently and by different degrees. For instance, response categories (e.g., “very good” to “very bad,”) may be used by different individuals in consistent but different ways, depending on how they interpret “very bad” as more or less negative. Such error may strengthen or weaken observed relationships (Nunnally 1978), which is particularly problematic for empirical research.

Correlational systematic error can be further differentiated as being within- or across-measures. Within-measure correlational systematic error arises between different items of the same construct. For example, stronger relationships may be observed between items of a construct employing the same response format (Viswanathan 2005). Halo error – where a global impression is used to rate on distinct dimensions – is another source (Lance, LaPointe, and Stewart 1994). One or two items can create a halo, leading to consistent responses to other items of the construct. As a result, responses to later items in a measure may be more polarized and consistent (Feldman and Lynch 1988; Knowles 1988; Simmons, Bickart, and Lynch 1993).

Across-measure correlational systematic error, on the other hand leads to inaccurate but consistent observed relationships between measures, and can be reflected in increased or decreased correlations. It is referred to as method variance; however, the language we present here places it in broader perspective in terms of its nature and effects. Possible factors leading to across-measure correlational systematic error include placement of items of different measures on one page (Lennox and Dennis 1994).

This terminology is relevant in separating sources of error from actual error, and from the consequences of each type of error. This separation as it relates to common method bias has been carefully articulated by M&P. However, the broader terminology is particularly relevant to a term like common method bias, which is not precise enough to capture the different nuances briefly described previously in terms of the nature of measurement error involved. We refer to Viswanathan (2005) who provides detailed examples linking sources of different types of error from the types of random error, additive systematic error, to actual error and outcomes. Thus, with Table 2 of M&P, the conditions that might cause method bias can in fact lead to a variety of types of measurement error, common method variance being one. For example, lack of ability may often lead to random error. Moreover, certain sources may lead to within- versus across-measure correlational systematic error, making the terminology we mention increasingly pertinent. Grouping related items together may, for instance, increase within-measure correlational systematic error, but not necessarily across-measure correlational systematic error. The sources themselves come from different arenas, ranging from item-content, to response categories, and generic individual differences.

**Procedural remedies**

M&P present some remedies that decrease the difficulty of satisficing. We build on this discussion here, covering multiple research practices in the social sciences, specifically in structuring their questionnaires. Among these practices, for example, are interspersing items from different constructs, placing items measuring a construct contiguously with or without labels, or placing items measuring different constructs on different pages. These practices may impact correlational systematic error of both types, leading to a potential effect on the estimates of relationships between constructs.

A number of survey formats and potential remedies can be generated based on two broad methodological factors: (i) item sequencing within and across constructs, and (ii) separation or grouping of items and constructs through labeling, pagination or contiguous placement.

We present a summary of the variety of formats in Table 1, along with comments about the potential impact of format on types of measurement error. In Format 1, referred to as ‘contiguous items and measures of different constructs’, items from each construct/dimension (the latter if the construct is multidimensional) are presented in proximity contiguously, a common practice and a useful baseline. No labeling or pagination is used in this format. Format 2 is similar to Format 1 with the addition of labels for unidimensional constructs and/or dimensions of multidimensional constructs. In Format 3, each construct/dimension is presented on a different page, without labels. Each of these approaches serves to provide a logical division between dimensions of a construct and the constructs themselves, labeling more explicitly and pagination more subtly. Whereas the paginated Format 3 is consistent with M&P’s recommendation to create spatial separation and therefore decrease the use of implicit theories, Format 2 might, in fact, increase the use of implicit theories because the labels might encourage respondents to do so. On the other hand, both formats may lead to greater consistency among items within constructs, or within dimensions of a construct, exploiting within-measure correlational systematic error. Stability reliability may also be enhanced, due to consistency over time among items. Because individual dimensions, rather than the overall multidimensional construct, are labeled, items are expected to have higher loadings on respective factors representing dimensions. Labeling may have stronger effects than pagination due to the explicit naming of constructs and dimensions. In this regard, Bradlow and Fitzsimons (2001) found that labeling and grouping (of items on a screen), each led to reduced variance within a subscale.

In terms of relationships across measures of different constructs, labeling and pagination serve to separate constructs (or dimensions of a construct); therefore, they may reduce observed correlations when compared to Format 1. But Format 3 (pagination) might also lead to an appearance of a lengthier questionnaire, leading to a decreased motivation for respondents to respond accurately. Such a lack of motivation might indeed lead to greater consistency within measure, and perhaps also across measures. This discussion is illustrative of how procedural remedies might at once address one type of measurement error and worsen another type of measurement error. Thus, it is useful to think about procedural remedies from the perspective of how they impact different types of measurement error, and therefore the overall ability to fit models.

In Format 4, we introduce both labeling as well as pagination of contiguous measures and constructs. Such a format
Table 1
Formats of measurement and their impact on types of error.

<table>
<thead>
<tr>
<th>Survey format</th>
<th>Contiguity and separation</th>
<th>Label</th>
<th>Paginate</th>
<th>Measurement error</th>
<th>Interpretation in terms of M&amp;P’s procedural remedies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contiguous items &amp; measures of different constructs</td>
<td>No</td>
<td>No</td>
<td>Baseline</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>Contiguous items &amp; measures</td>
<td>Yes</td>
<td>No</td>
<td>Increases within-measure correlational systematic error, reduces inconsistency of responses across items. May reduce across-measure correlational systematic error because of separating of constructs compared to Format 1, but may also increase across-measure correlational systematic error because of respondents’ implicit theories of how the constructs are related.</td>
<td>Decreases difficulty of satisficing by generating consistent responses within a labeled measure. Also decreases motivation to respond accurately by invoking implicit theories.</td>
</tr>
<tr>
<td>3</td>
<td>Contiguous items &amp; measures</td>
<td>No</td>
<td>Yes</td>
<td>Similar effects as for Format 2, except that within-measure correlational systematic error might be lesser because there are no labels. Survey might appear longer, leading respondents to satisfice.</td>
<td>Does not invoke implicit theories as much as Format 2 (especially if previous responses are unavailable), but decreases motivation to respond accurately by increasing perceived survey length.</td>
</tr>
<tr>
<td>4</td>
<td>Contiguous items &amp; measures</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar and somewhat accentuated effects compared to Format 2.</td>
<td>Decreases difficulty of satisficing, and decreases motivation to respond accurately.</td>
</tr>
<tr>
<td>5</td>
<td>Interspersed items (&amp; measures)</td>
<td>NA</td>
<td>NA</td>
<td>Reduces within-measure correlational systematic error, but increases confusion (and therefore random error). May reduce across-measure correlational systematic error.</td>
<td>Increases difficulty of satisficing that comes from grouping related items together.</td>
</tr>
<tr>
<td>6</td>
<td>Contiguous items, separated measures</td>
<td>No</td>
<td>No</td>
<td>Increases correlations between items of same measure (within-measure correlational systematic error), but decreases across-measure correlational systematic error.</td>
<td>Increases consistency of responses within measure, but increases difficulty of satisficing by allowing for separation between measures of different constructs.</td>
</tr>
<tr>
<td>7</td>
<td>Contiguous items, separated measures</td>
<td>Yes</td>
<td>No</td>
<td>Increases within-measure correlational systematic error even more than Format 6 because labels invoke implicit theories, but decreases across-measure correlational systematic error when compared to Formats 1, 2, and 6.</td>
<td>Invokes implicit theories, but effect is less than when measures are contiguous (e.g., Formats 2–4).</td>
</tr>
<tr>
<td>8</td>
<td>Contiguous items, separated measures</td>
<td>No</td>
<td>Yes</td>
<td>Increases within-measure correlational systematic error even more than Format 6, but decreases across-measure correlational systematic error when compared to Formats 1, 3, and 6.</td>
<td>Lowers likelihood of invoking implicit theories, but decreases motivation to respond accurately by increasing perceived survey length.</td>
</tr>
<tr>
<td>9</td>
<td>Contiguous items, separated measures</td>
<td>Yes</td>
<td>Yes</td>
<td>Same as in Formats 7 &amp; 8, with relevant comparison being with Formats 1, 4, and 6.</td>
<td>Decreases motivation to respond accurately because labeling invokes implicit theories, but pagination serves as a separation. However, pagination increases perceived survey length.</td>
</tr>
</tbody>
</table>
combines the advantages and disadvantages of each of Formats 2 & 3. In Format 5, items from different constructs are completely interspersed. Podsakoff et al. (2003) discuss the possibility of interspersion, referred to it as intermixing of items, which increases inter-construct correlation and reduces intra-construct correlation. Similarly M&P recommend that items should be dispersed throughout the questionnaire to increase the difficulty of satisficing. Kline, Sulsky, and Rever-Moriyama (2000) include interspersion as a possible solution for reducing common method variance. Interspersion may detract from the halo effect within a construct when responses to later items are based on a general impression created by earlier items, likely reducing consistency across items within a construct or a dimension. However, interspersion can potentially cut both ways: it can serve to separate items within a construct, but can also create confusion and have the opposite effect of labeling or pagination. In this regard, the within-construct halo effect serves to increase consistency of responses to items within a construct. Subsequent items in a construct are interpreted in light of earlier items. Researchers have shown increased reliability for later items in a construct (Knowles 1988). However, interspersion also has the potential benefit of minimizing blurring across items from different dimensions of a construct when compared to the contiguous condition, suggesting higher fit for multidimensional models with CFA. Interspersion may also decrease observed relationships between measures of different constructs. This decrease when compared to Format 1 may be greater than the decrease due to labeling or pagination.

In Format 6, items are presented contiguously within dimensions of the same construct as in Format 1. However measures of different dimensions/related constructs are separated here. Thus, Format 6 is sequenced so that no two measures of dimensions of the same construct or measures of different but related constructs are contiguously placed, thereby reducing the inflated correlation between dimensions of the same constructs or between related constructs that could result from contiguous placement of dimensions. In comparison, the contiguity of measures of related but different constructs in Format 1 would lead to increased across-measure correlational systematic error, primarily because contiguous positioning might lead respondents to satisfice. Formats 7, 8, and 9 are similar to Format 6 with the addition of labeling, pagination, and both labeling & pagination respectively. Such labeling and pagination could lead to stronger relationships across related constructs when compared to Format 6 because of implicit theories being invoked, but at the same time could also lead to weaker across-construct relationships because of separation. Pagination would additionally affect the perceived burden on respondents, leading to decreased motivation to respond accurately.

We note that several other formats are also possible. For example, items within dimensions or constructs could be sequenced differently when compared to their sequencing during their validation. The sequencing used at validation capitalizes on within-measure correlational systematic error due to sources such as a halo effect in responses. A different, unvalidated sequencing of items may detract from this effect, thus reducing within-measure correlational systematic error. Labeling and pagination could also be added to this format, with their associated advantages and disadvantages as previously discussed for other formats.

In summary, it is useful to think about how procedural remedies impact different types of measurement error, and these errors then manifest into problems in theory-testing and/or applied research with the measures. An important outcome from the illustrative formats shown in Table 1 is that one procedural remedy has the potential to impact a variety of measurement errors, resulting in greater overall error or in some cases, canceling out different types of error. We now move to the discussion of how the nature of measure can affect different types of measurement error.

**Stimulus-centered measures**

M&P provide an in-depth discussion of the role of respondent ability, motivation, and satisficing in leading to common method biases. Very pertinent here are the nature of the measures – specifically whether the measures are respondent-centered or stimulus-centered. Viswanathan (2005) provides a detailed discussion of what makes a measure respondent-centered or stimulus-centered. In particular, it is the purpose of measurement that determines the unit of analysis. If the purpose is to "place stimuli on a continuum," then the measure is stimulus-centered (Viswanathan 2005, p. 213). On the other hand, if the purpose is to place respondents on a continuum, irrespective of whether measures are about respondent traits or external stimuli, then the measure is respondent-centered, that is, it is focused on individual differences rather than stimuli differences. When the purpose of measurement is to place stimuli on a continuum (i.e., scale stimuli), the analysis of data must occur by obtaining a covariance matrix across stimuli, not across respondents (the typical practice in marketing).

Most, if not all, research on common method bias has focused on respondent-centered measures, indeed often on respondent traits. Further, such research has assumed that what applies to respondent-centered measures also applies to stimulus-centered measures, implicitly ignoring away how such differences in the nature of the measure might affect the bias. Yet, a number of measures in marketing and social sciences are stimulus-centered, making the unit of analysis an object other than respondents (Kayande 2010; Viswanathan and Childers 1999). For example, when a researcher measures the quality of service provided by a retail store, and then tries to relate that quality to a retail performance measure such as sales, the unit of analysis is squarely the retail store – making the respondent a mere rater of the retail store’s service quality. Such a measure is then stimulus-centered, not respondent-centered. In a series of papers, Finn and Kayande (1997, 1999, 2004) articulated the differences in not only conceptualizing such stimulus-centered measures, but also the approach used to assess psychometric properties of such measures. The question then becomes, “how does common method bias manifest itself when the measure is stimulus-centered?” We provide some thoughts on how to answer this question.
At a qualitative level, respondents may be more certain and knowledgeable about traits and characteristics pertaining to themselves, rather than about external stimuli. Literature from a number of areas of research including self-concepts, self-referencing, and autobiographical memory (e.g., Krishnamurthy and Sujan, 1999) supports this line of reasoning, arising from the highly organized memory structure of the self (Greenwald and Banaji, 1989) and leading to advantages in elaboration of incoming information and memory. When using validated scales with items relating to the self, student respondents are perhaps able to complete questionnaires based on item content. It is not entirely clear how able students are in completing questionnaires that relate to multiple stimuli. In particular, students, and/or other respondents, might not have had experience with the stimuli, or experience in thinking about the stimuli (M&P), leading to a decreased ability to respond accurately to stimulus-centered measures. Ability, therefore, is likely to be a greater concern with stimulus-centered measures than with respondent-centered measures. The decreased ability to respond accurately to questions might lead to increased random error, but also increased within- and across-measure correlational systematic error, as a result of using consistent responses within and across measures of different constructs/dimensions. It should be noted that additive systematic error is also a possibility with stimulus-centered measures, making for an inaccurate scaling of stimuli. Beyond just lack of ability in responding to stimulus-centered measures, there might be a large variance across individuals in terms of their experience with the stimuli, leading to variation across individuals on their ability to respond to questions about the stimuli. M&P suggest a careful selection of respondents who have the ability to think about the stimuli. One way to do so is to screen respondents for level of experience or familiarity with the stimulus, but researchers should be wary of sample self-selection problems in drawing conclusions from their data.

A second, perhaps more serious, problem that arises with stimulus-centered measures is that respondents might be required to respond to a number of items about a large number of stimuli. The demands on the respondent can be at such a high level as to make satisficing a natural response, lacking the motivation to respond in a truthful manner. For example, consider a study that relates service quality of a sample of retail stores to their sales performance. Using a standard survey design. We can now use Eq. (1) to work out what happens when only one method is used to obtain the service quality scores. With a single method, the variance-partitioning model in Eq. (1) reduces to:

\[ \sigma_y^2 = \sigma_r^2 + \sigma_i^2 + \sigma_m^2 + \sigma_{r \times i}^2 + \sigma_{r \times m}^2 \]

(1)

where \( \sigma_y^2 \) is the variance in observed scores, \( \sigma_r^2 \) is the variance attributable to respondents, \( \sigma_m^2 \) is the variance attributable to items, \( \sigma_i^2 \) is the variance attributable to stores, and \( \sigma_{r \times m}^2 \) is the variance attributable to the two-way and three-way interactions between respondents, stores, items, and method, and random error. The framework in Eq. (1) assumes that the scores have been obtained using multiple methods, and that all stores have been rated by all respondents on all items across all methods. In other words, Eq. (1) assumes a fully crossed survey design. We now use Eq. (1) to work out what happens when only one method is used to obtain the service quality scores. With a single method, the variance-partitioning model in Eq. (1) reduces to:

\[ \sigma_y^2 = \sigma_{r(m)}^2 + \sigma_{i(m)}^2 + \sigma_{r \times i(m)}^2 + \sigma_{r \times x(m)}^2 \]

(2)

where all variances are now nested within method. M&P note that the systematic trait variance (i.e., \( \sigma_r^2 \)) is perfectly confounded with systematic method variance (\( \sigma_m^2 \)), which is exactly what nesting does. Eq. (2) provides clarity on why such confounding occurs. Further, we can also show how nesting can affect reliability of measures. Classical reliability, in its most fundamental form, is:

\[ r_{multi-method} = \frac{\sigma_r^2}{\sigma_r^2 + \sigma_i^2}, \]

(3)

where \( \sigma_i^2 \) captures all sources of variance in Eq. (2) that involve respondents.

When only one method is used, that is, all measures are nested within one method, reliability is:

\[ r_{single-method} = \frac{\sigma_{r(m)}^2}{\sigma_{r(m)}^2 + \sigma_i^2} = \frac{(\sigma_r^2 + \sigma_{r \times m}^2)}{(\sigma_r^2 + \sigma_{r \times m}^2) + \sigma_i^2} \]

(4)
Table 2  
Stimulus-centered measurement designs, risks, and impact on measurement error.

<table>
<thead>
<tr>
<th>Stimuli crossed or nested with respondents</th>
<th>Risks</th>
<th>Impact on measurement error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully crossed</td>
<td>Halo; unfamiliarity with stimulus; extreme fatigue</td>
<td>Increases within-measure and across-measure correlational systematic error; may increase random error because of unfamiliarity with stimulus and fatigue.</td>
</tr>
<tr>
<td>Fully nested</td>
<td>Confounding of stimuli and respondents; halo</td>
<td>Not possible to separate stimuli and respondent effects; increases within- and across-measure correlational systematic error but not as much as when fully crossed.</td>
</tr>
<tr>
<td>Partially nested</td>
<td>Fatigue; halo; respondent group heterogeneity across stimuli</td>
<td>Increases within- and across-measure correlational systematic error compared with fully nested, but not as much as when fully crossed. Allows separation of stimuli and respondent effects, but only if respondent groups are homogenous across stimuli.</td>
</tr>
</tbody>
</table>

If the variance attributable to the interaction of respondents and method (\(\sigma^2_{r \times m}\)) is non-zero, then \(r_{\text{single-method}}\) will be greater than \(r_{\text{multi-method}}\) because all sources of error variance in Eq. (3) are the same as in Eq. (4), except for \(\sigma^2_{r \times m}\) which is now confounded with \(\sigma^2_{r}\). This result indicates that using a single method produces biased estimates of reliability, which in turn may produce biased estimates of correlations between variables.

We now turn to the problem of stimulus-centered measures, that is, when the units of analysis implied by the research are the stimuli. If this is the case, reliability is not given by Eq. (3) or (4); rather the focus of interest is the variance that is attributable to the stimuli, that is, stores (i.e., \(\sigma^2_{s}\)). Variance that is attributable to the interaction between respondents and stores (i.e., the stimuli) is now a component of the error term which reduces reliability. In this regard, Viswanathan (2005) states that with stimulus-centered measures, what is required is consistency across respondents. Similarly, the variance attributable to the interaction of stores and method is also error, but if a single method is used, this component is confounded with the variance of interest – that is the variance attributable to stores. Thus, the question we now ask is if we were to use multiple methods, would there be a significant variation attributable to the interaction of stores and method? In other words, would the scaling of stores depend on which method is used? If this is the case, then using a single method certainly biases the reliability of stimulus-centered measures. The extent to which it does so is an empirical question, although we see many reasons for the scaling of stimuli to depend on the method used. We note that our framework here provides an approach to understand this problem.

The framework also makes clear why stimulus-centered measures create additional problems in survey research. Asking a respondent to respond to items about only one stimulus reduces the burden on the respondent, but at the same time completely confounds the variance due to respondents and stimuli. When the interest is in stimuli, not respondents, the researcher unfortunately will not know if the differences across scores for stimuli are because stimuli are different or because respondents are different. Thus, our earlier comment about stimulus-centered measures increasing the burden on respondents is only reinforced. Such an increase in burden has the potential to lead to greater levels of common method bias than with respondent-centered measures.

In the case of stimulus-centered measures, there are a number of ways in which the questionnaire could be designed, but each approach has its associated measurement error (and bias) problems. In Table 2, we identify three different types of designs for stimulus-centered measures (i.e., measures that are designed to place multiple stimuli on a continuum). With stimulus-centered measures, halo is likely to be a major driver of responses because each stimulus provides the halo which may then be used by the respondent to respond to all items. Such a halo effect may lead to increased within- and across-measure correlational systematic error, the latter for items of different dimensions being measured. Respondents are also unlikely to know about all stimuli, leading to increased random error; moreover, respondents are likely to be consistent within measures, leading to increased within-measure correlational systematic error. With multiple stimuli, fatigue is likely to be a serious problem because of the number of items for each of the multiple stimuli (the first case in Table 2). Fatigue may even lead to across-measure correlational systematic error due to consistency in responses on items across measures of different dimensions/constructs.

A fully nested design with respondents only responding to one stimulus is likely to ameliorate the fatigue issues, but the design fully confounds respondents and stimuli. Such a confound in the domain of stimulus-centered theories might invalidate the test of the theory. The across-stimuli covariance matrix is simultaneously across-respondents. One possibility to manage this problem of confounding is partially nesting respondents within stimuli, that is, respondents respond to a few stimuli. Such partial nesting might ameliorate the confounding and fatigue problems, but the number of items used in the survey must be severely restricted so as to not dramatically decrease motivation to respond accurately and/or increase satisficing. Furthermore, confounding can only be ameliorated to the extent that there is homogeneity across respondent groups that respond to specific stimuli. We note that confounding is likely to continue to be a problem when stimuli vary widely in terms of their attractiveness to different customer segments because respondents from within heterogeneous customer segments might simply respond to those stimuli with which they are familiar. We note that a complete interspersion of items, constructs, and stimuli is unlikely to yield data of good quality because of the large amount of random error such a confusing questionnaire might generate.
In summary, consistent with M&P’s assessment of the causes of bias, responding to multiple stimuli may lead to a decreased ability to respond accurately (due to lack of ability and/or experience), a decreased motivation to respond accurately (due to length of questionnaires), and decreased difficulty of satisficing (due to items related to a stimulus being grouped together). Systematic research is thus needed to study common method bias, and more broadly the impact of procedural remedies on different types of measurement error in the domain of stimulus-centered measures.

Concluding comments

M&P provide a set of procedural remedies to ameliorate the problem of common method bias. In this commentary, we have focused on (i) providing a general typology of types of measurement errors, (ii) identifying how M&P’S procedural remedies for common method bias affect those errors, (iii) identifying how a remedy can affect many types of errors, sometimes creating additional problems, and (iv) discussing how such errors and biases can be even more problematic in stimulus-centered measures. We hope that researchers consider these issues in the design of empirical studies.

References


