



Contents lists available at ScienceDirect

## Journal of Business Research



## Should we use single items? Better not☆

Marko Sarstedt<sup>a,d,\*</sup>, Adamantios Diamantopoulos<sup>b</sup>, Thomas Salzberger<sup>c</sup><sup>a</sup> Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany<sup>b</sup> University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria<sup>c</sup> Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria<sup>d</sup> University of Newcastle, Faculty of Business and Law, Australia

## ARTICLE INFO

## Article history:

Received 1 February 2016

Accepted 29 February 2016

Available online xxxx

## Keywords:

Single items

Multi-item scales

Measurement theory

C-OAR-SE

Expert raters

## ABSTRACT

Bergkvist (2016, this issue) and Rossiter (2016, unpublished) claim that the conclusions of Sarstedt, Diamantopoulos, Salzberger, and Baumgartner (2016) are unwarranted as the study's methodology is flawed. This paper begs to differ.

© 2016 Elsevier Inc. All rights reserved.

## 1. Thank you, but...

Keeping with Oscar Wilde's famous quote “there is only one thing in life worse than being talked about, and that is not being talked about”, we very much appreciate the comments on our paper “Selecting single items to measure doubly concrete constructs: A cautionary tale” (Sarstedt, Diamantopoulos, Salzberger, & Wilczynski, 2016) by Bergkvist (in press) and Rossiter (2016). Considering that some of the paper's results question their research findings on the appropriateness of single-item measures (Bergkvist & Rossiter, 2007, 2009), coupled with the fact that both authors enjoy lively academic debates (e.g. Bergkvist, 2015, Rossiter, 2005, 2008), their rejoinders did not exactly come as a surprise. However, while considering the comments expressed therein, two questions emerged that made it very difficult to accept their criticisms. This response will share these two questions and some potential answers, hoping that they will help readers make up their own minds regarding single-item measurement and those elements of the C-OAR-SE procedure (Rossiter, 2002, 2011a, 2011b) that build on this measurement approach.

☆ The authors would like to thank Florian Kolbe for his support with the analysis. Furthermore, the authors would like to thank Barry J. Babin and Mitch Griffin for their comments on an earlier version of this article.

\* Corresponding author at: Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany.

E-mail addresses: marko.sarstedt@ovgu.de (M. Sarstedt), adamantios.diamantopoulos@univie.ac.at (A. Diamantopoulos), thomas.salzberger@wu.ac.at (T. Salzberger).

## 2. Where have all the expert raters (and users) of C-OAR-SE gone?

Construct definition in C-OAR-SE ultimately depends on rational expert judgment; or as Bergkvist (in press) notes: “the role of expert judges should be to determine the nature of the construct (...)”. For example, to classify a construct as doubly concrete, expert raters (or a majority thereof) have to agree that the construct has a simple, clear object coupled with a single and single-meaning attribute (Bergkvist, in press; Rossiter, 2002). If the expert raters agree on the doubly concrete nature of the construct, the use of a single item is—according to the C-OAR-SE procedure—sufficient (e.g. Bergkvist & Rossiter, 2007, Rossiter, 2002). Following Rossiter (2002, p. 310), these expert raters are a “small group of judges with expertise regarding the construct”. However, these experts must not have training in psychometrics as “it seems highly unlikely that experts in psychometrics would be able to provide an unbiased judgment of whether a construct could be measured with a single item” (Bergkvist, in press).

Two points are worth noting here. First, the potential benefits regarding single-item measurement and approaches to develop/select such measures have been around for decades (see Fuchs & Diamantopoulos, 2009 for a review of relevant literature) and certainly long before Bergkvist and Rossiter (2007, 2009) articles and Rossiter's (2002) C-OAR-SE procedure. Thus to claim that experts in psychometrics would be biased against single items is not only offensive but simply incorrect. Similarly, to somehow imply that C-OAR-SE is responsible for introducing the idea of single-item measurement to the literature is also incorrect. In fact, it was not in the context of C-OAR-SE but in psychometrics that single-item measurement was first discussed. Specifically, until the 1970s, the use of single items was generally accepted. The

<http://dx.doi.org/10.1016/j.jbusres.2016.02.040>

0148-2963/© 2016 Elsevier Inc. All rights reserved.

Please cite this article as: Sarstedt, M., et al., Should we use single items? Better not, *Journal of Business Research* (2016), <http://dx.doi.org/10.1016/j.jbusres.2016.02.040>

seminal works by Churchill (1979), Jacoby (1978), and Peter (1979) triggered a rethinking process that marked the start of multi-item use in marketing and business research in general. The widespread use of Confirmatory Factor Analysis further contributed to the advancement of psychometric adoption. Apart from that, the benefits and limitations of single-item use have long been discussed in a variety of disciplines outside marketing such as management (e.g., Kwon & Trail, 2005, Loo, 2002), psychology (e.g. Sackett & Larson, 1990, Wanous & Reichers, 1996), and various fields of medical research (e.g. Pomeroy, Clark, & Philp, 2001, Robins, Fein, Barton, & Green, 2001).

Second, in Sarstedt et al. (in press), the experts used were not only methodological experts but also had substantive knowledge/expertise in advertising and branding (a point ignored by Bergkvist, in press), which is exactly in keeping with the C-OAR-SE procedure. Moreover, there were thirteen experts that participated in the study. In comparison, all of Bergkvist and Rossiter's research (e.g. Bergkvist & Rossiter, 2007, 2009, Rossiter, 2002, 2008, 2009, 2011a, 2011b, 2012) relies on a single set of two expert raters: Lars Bergkvist and John R. Rossiter. This absence of variability in expert raters not only decreases confidence in the validity of the study findings but also may set the wrong standard if Rossiter's (2016, p. 6) prediction that “psychometrics in about 10 years' time will be found only in a museum” and will be substituted by C-OAR-SE becomes reality.

To date, however, the C-OAR-SE procedure has not been nearly as well received in applied research as the original paper's citation count and the winning of the 2012 Jan-Benedict E.M. Steenkamp Award for Long Term Impact, would suggest. To explore the dissemination of C-OAR-SE, we analyzed all articles in the Scopus database that have cited Rossiter's (2002) original C-OAR-SE article. The search in February 2016 yielded 645 citations from a variety of fields such as management, marketing, accounting, finance, and psychology. To narrow down the list of articles, the analysis focused on the field of marketing in which the C-OAR-SE procedure has been originally proposed and widely discussed (e.g. Diamantopoulos, 2005, Rigdon, Preacher, Lee, Howell, Franke, & Borsboom, 2011, Rossiter, 2005). More precisely, the review of articles in marketing consists of studies published in the top 20 marketing journals identified in Hult, Reimann, and Schilke's (2009) journal ranking, which related research also used (e.g. Hair, Sarstedt, Pieper, & Ringle, 2012, Hair, Sarstedt, Ringle, & Mena, 2012). The search yielded a total of 108 articles. Next, each article was read in detail and subsequently evaluated in terms of the context in which the Rossiter (2002) paper was cited. A senior graduate student and a professor proficient in the C-OAR-SE procedure independently developed a classification scheme based on an initial coding of 20 articles. The two classification schemes were then discussed and merged into a common one (see Table 1), which served as basis for the analysis of the remaining articles.

The results in Table 1 show that 38 of 108 articles (35.19%) cite Rossiter (2002) to justify the use of single items. 15 of these articles explicitly refer to the doubly concrete nature of the constructs under consideration (e.g., Fortenberry & McGoldrick, 2011), while the others generically indicate that the use of single items is appropriate in the

context of their study (e.g., Patterson, Yu, & Kimpakorn, 2014). 19 of 108 articles (17.59%) referred to C-OAR-SE as an alternative approach to scale development (e.g., Ford, Mueller, & Taylor, 2011), sometimes indicating that they had considered Rossiter's (2002) criticism of standard scale development and testing procedures in their studies (e.g., Okazaki, Mueller, & Taylor, 2010). 14 articles (12.96%) cite Rossiter (2002) to highlight differences between reflective and formative measures and the need to carefully specify measurement models (e.g., Herington, Johnson, & Scott, 2009). 12 of the 108 articles (11.11%) cite the study to stress the need for careful construct definition (e.g., Nasco, Kulviwat, Kumar, & Bruner, 2008) or use (parts of) the C-OAR-SE procedure in this context (e.g., Brocato, Brocato, Voorhees, & Baker, 2012). Further contexts in which Rossiter (2002) has been cited include the need to consider content validity in measurement (9.26%), comments on the original and extensions of the C-OAR-SE procedure (7.41%), and the use of specific answer scale formats (2.78%). Most importantly, however, only 4 of 108 (3.70%) studies citing the original C-OAR-SE article actually used the procedure to develop construct measures. Cadeyux and Ng (2012) as well as Dickinger and Stangl (2013) use the C-OAR elements for construct measure validation and development, but do not use the procedure in full. In contrast, Sabri and Obermiller (2012) use C-OAR-SE to develop a measure of consumer perception of taboo in advertising but do not fully adhere to Rossiter's (2011b, p. 1585) suggestions as they rely on “faulty Likert answer scales”. Finally, Rossiter (2012) introduces a contrastive measure that distinguishes brand love from brand liking, which focuses on content validity of the answer scale but does not make the application of the other aspects of the C-OAR-SE procedure transparent.

In short, the review shows that while many researchers cite C-OAR-SE, only very few actually apply the procedure (or even parts of it). Needless to say, this is not John R. Rossiter's fault and he frequently laments the fact that so few researchers have actually implemented his procedure (e.g., Rossiter, 2015). Nevertheless, we subscribe to Boshoff and Theron (2015, p. 264) who commented on the absence of C-OAR-SEicans by concluding that “as marketers this response should tell us something. A lot of people walked into the showroom, kicked the tires, but left without buying. Force feeding a market will not work. Fatigue is setting in. It's time to move on”.

### 3. To predict or not predict: And if not, how?

The notion that single items of doubly concrete constructs exhibit at least the same predictive validity as their multi-item counterparts is one of the cornerstones of Rossiter's (2002, 2011a, 2011b) C-OAR-SE procedure and has been empirically tested by Bergkvist and Rossiter (2007, 2009). However, the relevance of the latter two studies for C-OAR-SE remains highly unclear. On the one hand, Bergkvist and Rossiter (2009, pp. 607–608) note that “Rossiter (2002) argues that single-item measures provide valid measurement of ‘doubly concrete’ constructs” and that “Bergkvist and Rossiter (2007) show that this argument holds empirically by demonstrating that the predictive validity of single-item measures of the doubly concrete constructs attitude towards the ad ( $A_{Ad}$ ) and brand attitude ( $A_{Brand}$ ) is equal to the predictive validity of multiple-item measures of the same constructs”. On the other hand, Rossiter (2016, p. 5) emphasizes that “C-OAR-SE cannot be attacked by empirical, statistical arguments”, which necessarily implies that empirical, statistical arguments cannot provide support for C-OAR-SE either. Bergkvist (2016, p. 2) notes that “the purpose of the Bergkvist and Rossiter studies was to empirically test claims made by psychometricians using tests favored by psychometricians (and the studies showed that those claims did not withstand empirical testing)”. On the contrary, Rossiter (2016, p. 4) states that the only reason they considered predictive validity in their study titled “The predictive validity of multiple-item versus single-item measures of the same constructs”, was “because the then-Editor of *JMR* insisted that we provide ‘empirical proof’”. Indeed, both authors seem to lack “unanimous agreement”

**Table 1**  
Citation context of Rossiter (2002).

Context of citation	Number of articles (%)
Justify use of a single item	38 (35.19%)
Mention of C-OAR-SE as an approach to scale development	19 (17.59%)
Differentiate between formative and reflective measures	14 (12.96%)
Construct definition	12 (11.11%)
Emphasis on measures' content validity	10 (9.26%)
Comments and extensions of C-OAR-SE	8 (7.41%)
Application of the C-OAR-SE procedure	4 (3.70%)
Use of specific answer scale formats	3 (2.78%)
Total	108 (100%)

Download English Version:

<https://daneshyari.com/en/article/10492635>

Download Persian Version:

<https://daneshyari.com/article/10492635>

[Daneshyari.com](https://daneshyari.com)