Single versus Multiple Measurement of Attitudes

A Meta-Analysis of Advertising Studies Validates the Single-Item Measure Approach

LAWRENCE ANG
Macquarie University, Australia
Lawrence.ang@mq.edu.au

MARTIN EISEND
European University Viadrina, Germany
Eisend@europa-uni.de

Academic researchers tend to use multiple items to measure a construct reliably, whereas practitioners tend to use single-item measures. But when constructs such as brand attitude, attitude toward the advertisement, and attitude toward behaviors are double concrete—with a clear, singular meaning in which the object being rated also is clear and singularly identifiable—a single-item measure suffices (Rossiter, 2011). Using the results of 189 advertising studies, the authors found no difference in effect sizes when the double-concrete dependent variables were measured with single or multiple items—which means data collection is more efficient and less tedious. That is good news for advertising researchers in an era of ever-decreasing response rates and attention spans of respondents.

INTRODUCTION

Advertising research often involves assessing different characteristics of an advertisement: how creative or persuasive it is, whether it can change brand beliefs, and so forth. Depending on the objective of the research, a large number of measures may be involved.

Practitioners often use the minimum number of questions necessary, because the greater the number, the more expensive the research. Additional questions also increase respondent boredom and fatigue and decrease response rate (Bean and Roszkowski, 1995; Dillman, Sinclair, and Clark, 1993; Roszkowski and Bean, 1990). Academic advertising researchers, however, are more concerned with methodological rigor, insisting that constructs, such as brand attitude, should be measured reliably. Their concern also inevitably leads to the practice of using multiple items to demonstrate the internal reliability of the construct, on the basis of psychometric principles.

Management Slant

- Studies that use single-item measures do not differ in effect size from those that use multiple-item measures.
- The number of items used to measure attitudes does not influence the effect of the independent variables on attitudes.
- If a construct has a clear and singular meaning and the object being rated also is clear and identifiable, then a single valid item is all that is needed.
- The study's results strongly suggest that the double-concrete precept is generalizable.

Submitted May 16, 2016;
revised August 16, 2016;
accepted August 19, 2016;
published online January 13, 2017

DOI: 10.2501/JAR-2017-001

To measure attitude toward the advertisement, academic researchers often use items such as, "The advertisement is good/bad"; "The advertisement is informative/noninformative"; "The advertisement is pleasant/unpleasant"; and "I like/dislike the advertisement." By contrast, practitioners often just use one item, such as, "I like/dislike the advertisement," because they are concerned less with internal reliability but more with the managerial value of the study. Practitioners, in fact, often strive to shorten rather than lengthen their surveys, which especially is important given ever-declining response rates (Stern, Bilgen, and Dillman, 2014; Harris, 2015). Although debates about differences in mindset between practitioners and academics are not new (Bartunek and Rynes, 2014; Hunt, 2002; Nyilasy and Reid, 2007; Southgate, 2006), resolving this particular issue about the use of single or multiple items may be useful for both parties.

In recent years, a new theory has developed that argues that not all constructs need to be measured with multiple items, even for academic research (Rossiter, 2002, 2011). Under certain circumstances—that is, when the attribute of the construct is unambiguous and refers to an unambiguous object—a one-item valid measure of the construct will suffice. If this theory proves to be true, the implication is wide ranging, because if no difference can be found when outcomes are measured with either single or multiple items, then, in practical terms, using single-item measures is more efficient. Although this theory has been supported by examinations of the relationship between advertisement attitude and brand attitude (Bergkvist and Rossiter, 2007), which stimulated much debate in the marketing literature (Rigdon, Preacher, Lee, Howell, *et al.*, 2011), it has received surprisingly little attention in the advertising research literature (*cf.* Bergkvist and Rossiter, 2009). The authors of the current study aimed to address this shortcoming by

- drawing the attention of advertising researchers to this issue;
- further testing the validity of this theory using the data of eight meta-analyses, consisting of 189 advertising studies; and
- extending the test to include experiments and, hence, examine other, causal relationships.

By examining a large number of mostly experimental studies and going beyond the relationship between advertisement attitude and brand attitude, this study sought to broaden the test of the theory. It tested the generalizability of the theory, in an effort to help reconcile differences of opinion on this issue.

LITERATURE REVIEW

Theory Development

In 2002, John R. Rossiter (University of Wollongong) first suggested a new six-step approach to scale development, which he called "**c**onstruct definition, **o**bject **a**nd **r**ater identification, **s**cale formation, and **e**numeration and reporting" (C-OAR-SE). The approach, which is more reliant on the use of judgement (*i.e.*, logic or rationalism) in elucidating the validity of a construct, is diametrically opposed to the traditional psychometric approach, which relies more on the resultant statistics to prove the construct's validity (see also Boorsboom, Mellenbergh, and van Heerden, 2004). This psychometric approach, which is practiced widely in marketing (*e.g.*, Churchill, 1979), can be blamed for encouraging researchers to adopt a ritualistic measurement practice, such as being overly focused on the internal reliability scores of a scale (*e.g.*, coefficient alpha), without first thinking carefully about a construct's definition or whether the items possess a meaning similar to that of the construct (Rossiter, 2011).

Adding similarly worded items in an attempt to increase internal reliability does not necessarily add more information and may in fact decrease the validity of the measure (Drolet and Morrison, 2001). Other indices, such as interitem and item-to-total correlations, also could be considered (DeVellis, 1991; Netemeyer, Bearden, and Sharma, 2003). Although scale development texts warn against "useless redundancy" (*e.g.*, DeVellis, 1991, p. 191; Netemeyer *et al.*, 2003, p. 98), this advice seldom is heeded. Worse still, because academic journals demand the use of multiple measures, a scale's high internal reliability often is mistaken for its validity (Rossiter, 2011).

For a double-concrete construct, a single item that has valid content will suffice, and internal reliability is a nonissue (Rossiter, 2011), although it still can be estimated (Wanous and Reichers, 1996). "Double concrete" means the attribute of the construct has a clear, singular meaning (*i.e.*, is unidimensional), and the object being rated also is clear and unambiguous to the person doing the rating. The construct of purchase intentions is a good example, because its meaning is singularly clear and its reference to the object—that is, the product—also is unmistakable. Because there is nothing ambiguous or abstract about this construct, one item is all that is needed. Two other notable examples are the constructs of advertisement attitude and brand attitude.

Advertisement attitude and brand attitude especially are important in advertising research because they often are studied together and measured concurrently. When advertisement attitude increases, so does brand attitude, which then increases purchase intentions (McKenzie and Lutz, 1989). Another prominent attitude

measure in advertising and, in particular, social marketing and communication research is attitude toward behavior (*e.g.*, attitude toward smoking), which also has a singular, clear meaning and reference to an object (*e.g.*, smoking).

Further research took the double-concrete argument one step further, by demonstrating that there was no statistical difference in the simple correlations between advertisement attitude and brand attitude for four product categories—painkillers, coffee, pension plans, and jeans—whether the constructs were measured with a single or multiple items (Bergkvist and Rossiter, 2007). That study has been cited frequently because its implications are wide ranging. If a single-item construct, such as advertisement attitude, reliably can predict the outcome of another, such as brand attitude, it would save researchers considerable resources. Bergkvist and Rossiter (2007) demonstrated that a single-item assessment of constructs such as advertisement and brand attitude was as good as its multi-item equivalent. The results were replicated and extended in a later study (Bergkvist and Rossiter, 2009), which confirmed the double-concrete precept, with the exception of one product category—wine.

That line of investigation, which was about the predictive validity of a measure, was a departure from previous studies, in which the focus was more about how much the internal reliability of a scale would decrease with a reduction in the number of items (Jordan and Turner, 2008; Linden and Rosenthal, 2016; Wanous and Hudy, 2001; Wanous, Reichers and Hudy, 1997). In this regard, Bergkvist and Rossiter's study broke new ground by being the first to test empirically how well a particular construct could predict another using either a single or multiple items (Bergkvist and Rossiter, 2007).

Other scholars later challenged the researchers' findings:

- A three-part study involving five product categories—painkillers, coffee, insurance, jeans, and cars (Diamantopoulous, Sarstedt, Fuchs, Wilczynski, and Kaiser, 2012)—found limited support for their conclusion. Although the 2012 study found that some single-item measures of advertisement attitude could be just as good as the multi-item equivalent in predicting brand attitude, most were not: "While single-item measures can in specific applications produce a comparable level of predictive validity as multiple-item scales, there is no guarantee that they will" (p. 439). The researchers also argued that because the correlations in Bergkvist and Rossiter's (2007, 2009) studies came from the same sample and therefore were correlated, they had applied the wrong test of significance.

- Research published in 2015 used data from a consumer panel of milk buyers when correlating two attitudinal constructs—a person's concern about weight and preference for natural food—to milk consumption (Kamakura, 2015). This researcher found that the correlations were lower with single-item measures of the attitudinal constructs compared with a multi-item version. Because the meaning of these two attitudinal constructs was not defined and one of them possibly was not double concrete, the findings later were dismissed (Bergkvist, 2015).

- Another group of researchers (Sarstedt, Diamantopoulos, Salzberger, and Baumgartner, 2016) asked 13 experts, all well versed in psychometric principles, whether a single item was sufficient to measure advertisement attitude and brand attitude. Most concluded that it was not. Bergkvist (2016) pointed out, however, that this was not the correct way of assessing whether a construct is double concrete. The correct way is to ask whether the object, such as an advertisement or brand, is identified clearly and whether the attribute being evaluated, such as liking, possesses a singular, unambiguous meaning.

Meta-Analysis and Effect Size

In consideration of the controversy and inconsistent findings of previous investigations, the current research analyzed the findings of 189 advertising studies. It contributes to the debate by asking whether the double-concrete precept holds across larger samples and more product categories. The objective was to examine whether the effect size that describes the relationship between two variables, such as between humor in advertising and brand attitudes, changes when attitudes are measured with single or multiple items. If there were no change, one could conclude that the number of items used to measure attitudes or, more generally, the dependent variables, is immaterial.

Meta-analysis is a statistical technique in which the findings from several studies referring to the same hypothesis or research question are combined. These findings are described by effect size estimates. An effect size is a statistical measure, such as a correlation coefficient, that describes how strongly two variables are related or depend on each other. Meta-analysis combines similar effect sizes across studies. The result, called the meta-analytic effect size, provides a generalized outcome of the strength of the relationship between the two variables.

Furthermore, the variation in effect sizes can be explained by third variables. If, for example, the effect size that describes the relationship between humor in advertising and brand attitudes shows large variation, one could divide the effect sizes into those that were retrieved in experimental

studies and those that were retrieved in surveys and test whether the study design influenced the effect sizes. Effects found in experimental studies usually are stronger than those found in surveys, and therefore the authors would expect the effect size to differ. If they do not differ across a large number of studies, however, the implication is that the study design does not influence the findings, and it does not matter whether the data were collected through experiments or surveys.

In the context of the current study, the question was whether the effect size was different between studies that used a single-item measure of attitudes and those that used multi-item measures. If, on the one hand, single-item measures of attitudes were less valid than the multi-item measures, one would expect to see a difference in the effect size. If, on the other hand, no difference were found, then the number of items used to measure the attitudes would be immaterial—and single-item measures would be as good as multi-item ones.

Because meta-analyses are based on multiple studies and large samples, they provide a high degree of generalization, thus addressing the problems of prior studies, which typically are based on smaller samples. The previous debate on this issue involved only six studies (Bergkvist and Rossiter, 2007, 2009; Diamantopoulous *et al.*, 2012; Sarstedt *et al.*, 2016) covering eight product categories: painkillers, coffee, insurance, jeans, beer, breakfast cereals, wine, and cars. By contrast, the current authors examined the findings provided by eight meta-analyses, which represented 189 studies with nearly 40,000 participants and covered a wide range of product categories.

Predictive Validity and Experiments

One of the hallmarks of modern science is the ability to make accurate predictions. The ultimate achievement of advertising pretesting is to be able to use pretest scores in time 1 to predict subsequent sales level in time 2. If this reliably can be achieved, the pretest methodology is said to have predictive validity. Another way of inferring causality is to use experiments (Shadish, Cook, and Campbell, 2002). In experiments, certain variables of interest, such as level of message credibility, are manipulated. If differences in the outcome measures—advertisement attitude, brand attitude—are observed, one can conclude that the manipulated variables are responsible for this effect and therefore have predictive validity. Most of the studies in the meta-analyses were based on experiments. The result of the current study, therefore, was intended to provide strong evidence about the predictive validity of single- versus multi-item measures.

The current approach is different from previous studies, which largely were about the relationship, or correlation, between two outcome measures, such as advertisement attitude and brand attitude. In those studies, outcome measures were related and often measured concurrently, and variables were not manipulated. As such, one even could argue that what was demonstrated in past studies was more akin to concurrent validity (Cronbach and Meehl, 1955), rather than strictly predictive validity.

The current study, however, did not examine the relationship between the outcome measures—that is, between advertisement attitude and brand attitude. Instead, it examined the relationship between the manipulated variables and the outcome measures—that is, whether the outcomes were affected individually by the variables manipulated in the experiments, such as humor, credibility, and two-sided advertising. Such an approach allows one to assess the predictive validity of single- versus multi-item measures. The focus of this study, therefore, was on the relationship between the independent and dependent variables, as expressed by a correlation-based effect size, rather than the relationship between the two dependent measures. Also included in the current study was another dependent variable, called "attitude toward a behavior," such as disease-prevention behavior. This dependent variable, which is used more commonly in health-communications studies, is also a double-concrete variable and has not been studied before in this context.

Finally, the current study also directly tested whether the number of items used to measure the outcomes would influence the effect size, through meta-regression. Previous research, which often was limited to a few studies, could not adopt this analytical approach. It is feasible with meta-analysis because a large number of studies are aggregated.

METHODOLOGY

In a recent meta-meta-analysis, the current researchers collated all previous meta-analyses in advertising published by early 2015, for the purpose of understanding the advertising's effectiveness (Eisend and Tarrahi, 2016). For the current study, the authors selected meta-analyses published since 2000 with attitudes as the dependent variables (See Technical Appendix). The authors of the selected meta-analyses then were contacted, and information was sought on effect sizes related to attitudes, sample size, and scale items of the attitude variables. If the authors had not coded the scales but could provide all other information, the current researchers coded the scale items for this study by referring to the individual studies that were included in the meta-analysis. Eventually, the following eight meta-analyses were used in the current study:

- a meta-analysis that dealt with the effects of advertising exposure on children's attitudes (Desmond and Carveth, 2007);
- a meta-analysis that dealt with the effects of credibility in marketing communication on attitudes (Eisend, 2006a);
- a meta-analysis that dealt with the effects of two-sided advertising on advertisement and brand attitudes (Eisend, 2006b);
- a meta-analysis that dealt with the effects of humor in advertising on advertisement and brand attitudes (Eisend, 2009);
- a meta-analysis that dealt with the effects of public relations versus advertising on advertisement and brand attitudes (Eisend and Küster, 2011);
- a meta-analysis that dealt with the effects of cultural adaptation of advertising on advertisement liking (Hornikx and O'Keefe, 2009);
- a meta-analysis that dealt with the effects of gain-framed versus loss-framed messages on attitudes toward disease-prevention behaviors (O'Keefe and Jensen, 2007);
- a meta-analysis that dealt with the effects of gain-framed versus loss-framed messages on attitudes toward disease-detection behaviors (O'Keefe and Jensen, 2009).

For each study in the meta-analysis, the attitudes were coded by whether they were measured with a single- or multi-item scale. To analyze whether the effect sizes differed depending on whether a single- or multi-item scale was used in the study, the researchers applied an inverse-variance weighted one-way analysis of variance mixed-effects model to each meta-analysis. This technique commonly is used in meta-analytical methodology to test the influence of categorical variables on the effect sizes. The procedure partitions the total variability of the effect sizes into the portion explained by the categorical variable ($Q_{explained}$) and the residual ($Q_{unexplained}$). Q follows a chi-square distribution and can be used to test whether the explained variance due to the categorical variable is significant—that is, whether the effect sizes differ between single- and multi-item scales.

Because some studies measuring attitudes used single-item scales, whereas others used multi-item scales, they were coded as the independent variable in this study's analysis (0 = a study that used a single-item scale; 1 = a study that used a multi-item scale). The dependent variable was the effect size—that is, the correlation coefficient that measures the relationship between the manipulated variables (*i.e.*, humor, credibility, public relations versus advertising, two-sided advertising, children's advertisement exposure, cultural adaptation, gain- or loss-framed message) and the outcome variables (*i.e.*, advertisement attitude, brand attitude, and attitude toward behavior). The meta-analyses used different weights for integration, such as sample size, inverse variance; and artifact corrections, like reliability correction, and therefore the correlation-based effect sizes were not directly comparable between meta-analyses. They can be compared within a meta-analysis, because each meta-analysis used consistent weighting and correction procedures.

The actual number of items used to measure attitudes (*i.e.*, scale length) also was included in this analysis. To investigate whether the number of items, which is a continuous variable, influenced the effect size, the researchers applied an inverse-variance weighted meta-regression, mixed-effects model, a technique that commonly is used to test the influence of continuous variables on the effect sizes. The variance explained by the independent variable again is indicated by $Q_{explained}$ and can be used to test whether the number of items influences the effect sizes.

RESULTS

The objective of this study was to investigate whether the use of either single- or multi-item scales for the dependent variable (*i.e.*, attitude) influenced the effect size. Because the choice of single-item versus multi-item scales was not related to other characteristics of the studies, a significant difference in the effect size would have indicated that the choice of scale (*i.e.*, single-item versus multi-item) matters. Conversely, if no difference in effect size were found, one would conclude that the use of a single- or multi-item scale in measuring attitude is immaterial. In addition, the influence of the scale length (*i.e.*, the number of items) on the effect size also was tested directly through meta-regression.

The Q-test results showed that, except for one test (the influence of two-sided advertising on advertisement attitude; Eisend 2006b, See Table 1), all others were not significant when the researchers assumed a significance threshold of 5 percent. That is, studies that used a single-item scale did not differ from those that used multi-item scales in their effect size (See Table 1). Because the use of single-item or multi-item scales did not influence the effect on attitudes of various advertising tactics (*i.e.*, humor, credibility, public relations versus advertising, two-sided advertising, children's advertisement exposure, cultural adaptation, or gain- or loss-framed messages) the results suggest that single-item measures are as good as multi-item measures.

Null effects might have been due to weak statistical tests based on low sample sizes. The magnitude and the signs of the effect sizes for single- and multi-item scales, however, showed considerable variation and no systematic difference: Single-item measures led to greater effect sizes in seven cases, smaller effect sizes in four cases, and different signs in two cases.

TABLE 1

Influence of Single-Item versus Multi-Item Scale on Effect Sizes (Weighted Analysis of Variance, Mixed-Effects Model)

Authors and Year	Relationship Investigated	No. of Studies	No. of Participants	No. of Effect Sizes	Effect Size Single-Item	Effect Size Multi-Item	*Q* Explained	Significance of *Q*
Desmond and Carveth (2007)	Advertisement exposure of children—attitudes	12	4,091	12	0.138	0.160	0.126	0.723
Eisend (2006a)	Credibility—attitudes	17	4,425	28	0.317	0.385	0.221	0.639
Eisend (2006b)	Two-sided advertising—advertisement attitude	14	4,853	56	–0.097	0.168	6.985	0.008
Eisend (2006b)	Two-sided advertising—brand attitude	16	4,586	64	0.174	0.064	3.578	0.059
Eisend (2009)	Humor—advertisement attitude	25	3,391	84	0.232	0.174	0.445	0.505
Eisend (2009)	Humor—brand attitude	17	2,227	48	0.053	0.158	3.024	0.082
Eisend and Küster (2011)	PR vs. advertising—advertisement attitude	17	3,371	63	0.200	0.108	1.325	0.250
Eisend and Küster (2011)	PR vs. advertising—brand attitude	12	2,766	26	–0.215	0.202	2.364	0.124
Hornikx and O'Keefe (2009)	Cultural adaptation of ads—advertisement liking	28	5,423	54	0.237	0.057	2.676	0.102
O'Keefe and Jensen (2007)	Gain- vs. loss-framed messages—attitude toward disease prevention behaviors	15	2,699	20	0.194	0.099	0.151	0.698
O'Keefe and Jensen (2009)	Gain- vs. loss-framed messages—attitude toward disease-detection behaviors	16	2,133	22	–0.143	–0.010	1.838	0.175
Total		189	39,965	477	–	–	–	–

Note: The number of effect sizes does not always correspond to the total number of effect sizes that is reported in the meta-analysis, because not all studies included in the meta-analysis provided information on scales. "Studies" refers to independent samples. That is, a single article can include more than one study.
PR = public relations.

These results cannot easily be explained by lack of statistical power.

The analysis also examines how well the number of items can influence effect size using meta-regression (See Table 2). The *Q* test of the explained variance of each model indicated that none of the effects was significant when the researchers assumed a significance threshold of 5 percent That is, the number of items used to measure attitudes did not influence the effect of the independent variables (*i.e.*, humor, credibility, public relations versus advertising, two-sided advertising, children's advertisement exposure, cultural adaptation, gain- or loss-framed message) on attitudes. This, again, suggests that the use of a single-item measure of a double-concrete construct is as good as the use of a multi-item measure. The signs of the beta coefficients showed considerable variation, and there was no systematic relationship between effect size and number of items. These results cannot be explained easily by lack of statistical power and indicate low validity threat.

DISCUSSION

In this study, 189 advertising studies from eight previous meta-analyses were examined. In all these studies, attitudes were the outcome variables—advertisement attitude, brand attitude, and attitude toward

TABLE 2

Influence of Number of Items on Effect Sizes (Weighted Meta-Regression, Mixed-Effects Model)

Authors and Year	Relationship Investigated	No. of Studies	No. Participants	No. Effect Sizes	β	*Q* Explained	Significance of *Q*
Desmond and Carveth (2007)	Advertisement exposure of children—attitudes	11	3,665	11	0.702	2.698	0.101
Eisend (2006a)	Credibility—attitudes	17	4,425	28	–0.389	2.964	0.085
Eisend (2006b)	Two-sided advertising—advertisement attitude	14	4,853	56	0.277	2.808	0.094
Eisend (2006b)	Two-sided advertising—brand attitude	16	4,586	64	–0.107	0.620	0.431
Eisend (2009)	Humor—advertisement attitude	25	3,391	84	–0.107	0.421	0.517
Eisend (2009)	Humor—brand attitude	17	2,227	48	0.244	2.837	0.092
Eisend and Küster (2011)	PR vs. advertising—advertisement attitude	17	3,371	63	0.045	0.080	0.777
Eisend and Küster (2011)	PR vs. advertising—brand attitude	12	2,766	26	0.041	0.019	0.891
Hornikx and O'Keefe (2009)	Cultural adaptation of ads—advertisement liking	28	5,423	54	–0.053	0.153	0.696
O'Keefe and Jensen (2007)	Gain- vs. loss-framed messages—attitude toward disease prevention behaviors	15	2,699	20	0.198	0.493	0.483
O'Keefe and Jensen (2009)	Gain- vs. loss-framed messages—attitude toward disease detection behaviors	16	2,133	22	–0.041	0.036	0.849
Total		188	39,539	476	–	–	–

Note: The number of effect sizes does not always correspond to the total number of effect sizes reported in the meta-analysis, because not all studies included in the meta-analysis provided information on scale length. "Studies" refers to independent samples. That is, a single article can include more than one study.
PR = public relations.

behavior—measured with either a single or multiple items. This allowed the researchers to test whether the effect size changed when these outcome variables were measured with either a single or multiple items. If a single-item measure is as valid as a multi-item measure, then the effect size should be the same across the studies. This is exactly what the researchers found. In the 22 tests, no significant difference was found in the effect size, except for one (See Tables 1 and 2). That single significant finding can be considered a type 1 error (*i.e.*, one out of 22 tests). The study therefore supports the double-concrete precept (Bergkvist and Rossiter, 2007). That is, if a construct has a clear and singular meaning and the object being rated also is clear and identifiable, then a single valid item is all that is needed.

This finding also is notable for three reasons:

- A much larger number of product categories, respondents, and studies were examined, as compared with previous studies.

- The effect size was based on the relationship between a large number of independent variables (humor, credibility, public relations versus advertising, two-sided advertising, children's advertisement exposure, cultural adaptation, gain- or loss-framed message) and the outcome measures (advertisement attitude, brand attitude, and attitude toward behavior). Because these independent variables and the outcome measures were not directly related, this study is a more stringent test of the double-concrete precept. Past studies tended to examine the relationship between the two related outcome measures of advertisement attitude and brand attitude, rather than their relationships with advertising input variables. This result strongly suggests that

the double-concrete precept is, indeed, generalizable.

- Because most of the studies were based on randomized controlled experiments in which different kinds of independent variables were manipulated, one also can make a stronger inference about the predictive value of using a single-item measure. The fact that there was no difference in the effect size implies that the predictive powers of a single-item measure are as good as those of a multi-item measure. This was confirmed further in a meta-regression, which found that the number of items used in measuring the attitudinal constructs had no influence on effect size. This result, again, supports the double-concrete precept.

In summary, although the double-concrete precept has not been accepted unanimously (*e.g.*, Diamantopoulous *et al.*, 2012), the current large-scale study provides highly generalized support for this precept. To this end, this study has extended the investigation into attitude toward behavior, as well as examining the relationship between the independent variables and attitudes in experiments. The consistency of these findings across a large number of studies provides further evidence that the double-concrete precept is correct.

Future research further can validate this theory by investigating other constructs measured in different contexts. It also can examine in greater depth why multi-item measures are not necessarily better than single-item measures. There may be many reasons:

- the wording of some items may not adequately reflect the focal meaning of the construct;
- some items may suffer from inconsistent object reference; and
- other items may be worded in a complex manner.

All these factors may result in confusion, irritation, fatigue, and misinterpretation, leading to errors in the rating. Unless care is taken, the more items there are, the greater is the likelihood of such errors occurring.

PRACTICAL IMPLICATIONS

Attitudes can, indeed, be measured with a single item. Although academic advertising research tends to favor the use of multi-item measures, this study shows that such measures are not always necessary. When an attitudinal construct is double concrete, all that is needed is one good item. The item fundamentally must be valid; it should measure what it is supposed to measure.

This means that all the hard thinking about the construct's definition and how it can be operationalized meaningfully still needs to be completed upfront. The support for the use of single-item measures found in the current study does not obviate this requirement. In fact, it may actually raise its importance, given that everything now hinges on that single item—it has to be the best measure yet for that construct. This is good news in an era of ever-declining response rate and attention span.

Finally, this study suggests that the difference in mind-set between academics and practitioners can be reconciled. In this case, the current findings favor the practitioners' perspective. This implies that it is better to have one valid measurement item that fully captures the semantic meaning of the construct rather than having multiple bad ones, no matter how internally consistent the measurement scale may be. With valid measures, theories can be developed and tested better. The reward may be better data and better theories, but at lower cost. JAR

ABOUT THE AUTHORS

Lawrence Ang is an associate professor of marketing at Macquarie University, Australia. Prior to entering academia, he was a senior marketing research consultant. He has a keen interest in customer relationship management, advertising effectiveness, media studies, and consumer behavior. Ang's work has been published in the *European Journal of Marketing*, *Journal of Economic Psychology*, *International Journal of Management Reviews*, *Journal of Advertising*, *Journal of Advertising Research*, *Journal of Marketing Management*, and *Celebrity Studies*. His latest book is *Principles of Integrated Marketing Communications* (Port Melbourne: Cambridge University Press, 2014).

Martin Eisend is a professor of marketing at the European University Viadrina in Frankfurt (Oder), Germany. His research focuses on consumer behavior, marketing communication, and empirical generalizations. Eisend's work has been published in the *Journal of Marketing*, *Journal of Advertising*, *International Journal of Advertising*, *International Journal of Research in Marketing*, *Journal of Product Innovation Management*, *Journal of the Academy of Marketing Science*, *Journal of Retailing*, and *Research Policy*, among other publications.

REFERENCES

Bartunek, J. M., and S. L. Rynes. "Academics and Practitioners Are Alike and Unlike: The Paradoxes of Academic–Practitioner Relationships." *Journal of Management* 40, 5 (2014): 1181–1201.

Bean, A. G., and M. J. Roszkowski. "The Long and Short of It: When Does Questionnaire Length Affect Response Rate?" *Marketing Research* 7, 1 (1995): 21–26.

Bergkvist, L. "Appropriate Use of Single-Item Measures Is Here to Stay." *Marketing Letters* 26 (2015): 245–255.

Bergkvist, L. "The Nature of Doubly Concrete Constructs and How to Identify Them." *Journal of Business Research* 69 (2016): 3427–3429.

Bergkvist, L., and J. R. Rossiter. "The Predictive Validity of Multiple-Item versus Single-Item Measures of the Same Constructs." *Journal of Marketing Research* 44 (2007): 175–184.

Bergkvist, L., and J. R. Rossiter. "Tailor-Made Single-Item Measures of Doubly Concrete Constructs." *International Journal of Advertising* 28, 4 (2009): 175–184.

Boorsboom, D., G. J. Mellenbergh, and J. van Heerden. "The Concept of Validity." *Psychological Review* 111, 4 (2004): 1061–1071.

Churchill, G. A., Jr. "A Paradigm for Developing Better Measures of Marketing Constructs." *Journal of Marketing Research* 16, 1 (1979): 64–73.

Cronbach, L. J., and P. E. Meehl. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52 (1955): 281–302.

Desmond, R., and R. Carveth. "The Effects of Advertising on Children and Adolescents: A Meta-Analysis." In *Mass Media Effects Research: Advances Through Meta-Analysis*, R. W. Preiss, B. M. Gayle, N. Burrell, M. Allen, and J. Bryant, eds. New York: Routledge, 2007.

DeVellis, R. F. *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage, 1991.

Diamantopoulos, A., M. Sarstedt, C. Fuchs, P. Wilczynski, and S. Kaiser. "Guidelines for Choosing between Multi-Item and Single-Item Scales for Construct Measurement: A Predictive Validity Perspective." *Journal of the Academy of Marketing Science* 40, 3 (2012): 434–449.

Dillman, D. A., M. D. Sinclair, and J. R. Clark. "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys." *Public Opinion Quarterly* 57 (1993): 289–304.

Drolet, A. L. and D. G. Morrison. "Do We Really Need Multiple-Item Measures in Service Research?" *Journal of Service Research* 3, February (2001): 196–204.

Eisend, M. "Source Credibility in Marketing Communication: A Meta-Analysis." *Marketing – Journal for Research and Management* 2, 1 (2006a): 43-60.

Eisend, M. "Two-Sided Advertising: A Meta-Analysis." *International Journal of Research in Marketing* 23, 2 (2006b): 187–198.

Eisend, M. "A Meta-Analysis of Humor in Advertising." *Journal of the Academy of Marketing Science* 37, 2 (2009): 191–203.

Eisend, M., and F. Küster. "The Effectiveness of Publicity versus Advertising: A Meta-Analytic Investigation of Its Moderators." *Journal of the Academy of Marketing Science* 39, 6 (2011): 906–921.

Eisend, M., and F. Tarrahi. "The Effectiveness of Advertising: A Meta-Meta-Analysis of Advertising Inputs and Outcomes." *Journal of Advertising* 45, 4 (2016): 519–531.

Harris, D. F. (2015, September 3). "How to write shorter surveys: Treat questionnaire design as a professional discipline." Retrieved May 12, 2016, from Marketing Research website: http://www.marketingresearch.org/article/how-write-shorter-surveys-treat-questionnaire-design-professional-discipline

Hornikx, J., and D. J. O'Keefe. "Adapting Consumer Advertising Appeals to Cultural Values: A Meta-Analytic Review of Effects on Persuasiveness and Ad Liking." *Communication Yearbook* 33 (2009): 39–71.

Hunt, S. D. "Marketing as a Profession: On Closing Stakeholder Gaps." *European Journal of Marketing* 36, March (2002): 305–312.

Jordan, J. S., and B. A.Turner. "The Feasibility of Single-Item Measures for Organizational Justice." *Measurement in Physical Education and Exercise Science* 12, (2008): 237–257.

Kamakura, W. A. "Measure Twice and Cut Once: The Carpenter's Rule Still Applies." *Marketing Letters* 26 (2015): 237–243.

Linden, S., and A. A. Rosenthal. "Measuring Narcissism With a Single Question? A Replication and Extension of the Single-Item Narcissism Scale (SINS)." *Personality and Individual Differences* 90, February (2016): 238–241.

McKenzie, S. B., and R. J. Lutz. "An Empirical Examination of the Structural Antecedents of Attitude toward the Ad in an Advertising Pretesting Context." *Journal of Marketing* 53, 2 (1989): 48–65.

Netemeyer, R. G., W. O. Bearden, and S. Sharma. *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: Sage, 2003.

Nyilasy, G., and L. N. Reid. "The Academician–Practitioner Gap in Advertising." *International Journal of Advertising* 26, 4 (2007): 425–445.

O'Keefe, D. J., and D. J. Jensen. "The Relative Persuasiveness of Gain-Framed and Loss-Framed Messages for Encouraging Disease Prevention Behaviors: A Meta-Analytic Review." *Journal of Health Communication*, 12 (2007): 623–644.

O'Keefe, D. J., and D. J. Jensen. "The Relative Persuasiveness of Gain-Framed and

Loss-Framed Messages for Encouraging Disease Detection Behaviors: A Meta-Analytic Review," *Journal of Communication* 59 (2009): 296–316.

Rigdon, E. E., K. J. Preacher, N. Lee, R. D. Howell, *et al.* "Avoiding Measurement Dogma: A Response to Rossiter." *European Journal of Marketing* 45, 11/12 (2011): 1589–1600.

Rossiter, J. R. "The C-OAR-SE Procedure for Scale Development in Marketing." *International Journal of Research in Marketing*, 19 (2002): 305–335.

Rossiter, J. R. *Measurement for the Social Sciences: The C-OAR-SE Method and Why It Must Replace Psychometrics.* Berlin: Springer, 2011.

Roszkowski, M. J., and A. G. Bean. "Believe It or Not! Longer Questionnaires Have Lower Response Rates." *Journal of Business and Psychology* 4, 4 (1990): 495–509.

Sarstedt, M., A. Diamantopoulos, T. Salzberger, and P. Baumgartner. "Selecting Single Items to Measure Doubly Concrete Constructs: A Cautionary Tale." *Journal of Business Research* 69 (2016): 3159–3167.

Shadish, W. R., T. D. Cook, and D. T. Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin, 2002.

Southgate, N. "The Academic–Practitioner Divide: Finding Time to Make a Difference." *Marketing Intelligence and Planning* 24, 6 (2006): 547–551.

Stern, M.J., I. Bilgen, and D. A. Dillman. "The State of Survey Methodology: Challenges, Dilemmas, and New Frontiers in the Era of the Tailored Design." *Field Methods* 26, 3 (2014): 284–301.

Wanous, J. P., and M. J. Hudy. "Single-Item Reliability: A Replication and Extension." *Organizational Research Methods* 4 (2001): 361–375.

Wanous, J. P., and Reichers, A. E. "Estimating the Reliability of a Single-Item Measure." *Psychological Reports* 78 (1996): 631–634.

Wanous, J. P., A. E. Reichers, and M. J. Hudy. "Overall Job Satisfaction: How Good Are Single-Item Measures?" *Journal of Applied Psychology*, 82 (1997): 247–252.

Technical Appendix

The meta-analyses of the study were selected from a meta-meta-analysis (Eisend and Tarrahi, 2016). That meta-meta-analysis covered all meta-analyses on advertising topics available as of May 2015. To be considered as a meta-analysis, a study had to report a numerical measure of a relationship between two variables and systematically summarize the evidence of this effect collected in at least two primary scholarly studies by at least two researchers or research teams. The authors considered all meta-analyses published in traditional advertising, communication, and marketing outlets as well as unpublished meta-analyses. They were open to including meta-analyses that did not appear in traditional publication outlets as long as they addressed an advertising topic. The meta-analyses were retrieved from electronic databases, Internet searches, and a systematic search of journals and proceedings.

The authors of the current study selected those meta-analyses that dealt with attitudes as dependent variables, because attitudes typically are measured as either single- or multi-item measures and because they have been applied and debated in the literature discussing single- versus multi-item scales. Attitudes are prominent measures in advertising research and more often are studied than, for instance, recall, recognition, emotions, or beliefs.

It was necessary to contact several authors to retrieve the meta-analytic data of their studies. The likelihood of retrieving such data from authors who collected their data more than 16 years ago was close to zero. For pragmatic reasons, the study therefore focused on recent meta-analyses—that is, meta-analyses published in 2000 and later. Note that more than two-thirds of the meta-analyses in the study by Eisend and Tarrahi (2016) were published in 2000 and later. Furthermore, meta-analyses consisting of studies that measured attitudes using *only* multi-item scales were excluded from the dataset. The authors included all effects sizes from the selected meta-analyses, except for effect sizes from studies that did not indicate whether attitudes were measured by single- or multi-item scales or did not provide the number of scale items.