# Combating scale proliferation

## Gordon C. Bruner II

is an associate professor, Department of Marketing, Southern Illinois University, Illinois, USA. He is Director of the Office of Scale Research and the lead author of the Marketing Scales handbook series (AMA). His research has been published in *Journal of Marketing*, *Journal of the Academy of Marketing Science*, *Journal of Retailing*, *Journal of Marketing Education* and many other journals and conference proceedings.

**Abstract** It is not unusual for new scales to be created in scholarly marketing research with little concern for their relationship with previous measures of the same construct and little justification provided for their development. This practice is challenged and it is argued that this is one of the things that is preventing marketing from progressing as a science. An empirical example is provided to illustrate how this lack of concern can have negative consequences on a stream of research. Increased emphasis on scale standardisation, equivalency and justification are suggested as major components of the solution to the problem of scale proliferation. A scale selection and usage paradigm is offered to guide researchers in reaching these triple goals.

**Dr Gordon C. Bruner II**
Office of Scale Research,
Department of Marketing,
Southern Illinois University,
Carbondale, IL 62901-4629,
USA.

Tel: +1 618 453 4341;
Fax: +1 618 453 7747;
e-mail: gcbii@siu.edu

Usage of multi-item scales has become common over the last couple of decades in conducting scholarly market research. What also is clear, however, is that for most constructs there is little agreement about which pre-existing scale to use or whether it is just as good to create a new scale for the purpose. A study of attitude-toward-the-ad, for example, concluded that of 75 instances of its measurement, almost half of the scales had been used as a set of items just once.[1] In other words, it has not been unusual for new or substantially modified measures to be used despite the existence of other scales. The result is that while authors may be calling what they are measuring the same thing eg attitude-toward-the-ad), the lack of information about measure equivalency leaves open the possibility that variance in findings across studies is occurring because different scales are being used.

This paper proposes that scholarly marketing research has matured to the point where the typically provided evidence of a scale's reliability and validity is important but insufficient reasoning for use of a new scale in a study. The potential negative effects that result from the practice of scale proliferation will be described and illustrated. As a response to this problem it is argued that standardisation, equivalency and justification should take on greater importance in scale selection and usage paradigm. Finally, the paper concludes by offering a guide that typical users of scales may follow as they select, use and/or create scales.

## BACKGROUND

### Standardisation

With regard to psychometric scales, standardisation can be defined as the extent to which the same rules are used to assign numbers to a construct. The use of agreed-upon measures is common

in most fields that attempt to be scientific. In fact, a fundamental principle of science is that any observation made by one researcher should be independently verifiable by other researchers and this 'principle is violated if scientists can disagree about the measure.'[2] The implication is that a field's evolution from art to science is impeded as long as researchers can measure constructs any way they see fit because there is no way to say one scale is better than another.

There are several reasons for standardising usage. First, some degree of standardisation is necessary in order to validate a scale because validation is best viewed as a process; it is not likely to occur in one study but requires the methodical testing of a measure in multiple studies to produce a knowledge base of its psychometric properties.[3,4] Similarly, no single use of a scale is likely to examine all of the facets of generalisation that should be addressed by any measure for which wide application is desired.[5,6] Thirdly, a proper 'final' step in scale construction is development of norms.[7] This is helpful because scale scores are best interpreted in light of normative data regarding a measure's use with different groups and in a variety of situations.[8] Thus, multiple studies across conditions and populations should be conducted over time utilising the same scale in order to achieve the important psychometric goals of measure validation, generalisability and norm development.

In contrast, the continuity and value of the research process in a discipline is undermined when scales are developed, used once and then rarely if ever used again.[9] This was identified as a problem decades ago in the parent disciplines of marketing. In sociology journals for the period of 1954 to 1965, 72 per cent of scales were found to have been used just once.[10] Similarly, for the period of

1960–1969 in psychology, 63 per cent of scales were used just once.[11] Although there has been no known broad examination of marketing scale usage, a study of consumer scales estimated that 79 per cent of those published during the 1980s were used just once.[12]

Beyond these general reasons for the importance of standardisation, another concern is that measurement diversity has the potential to affect research results. Alternative measures of the same construct can vary in the amount of a construct's variance that is explained. It is quite possible that a construct has multiple facets and one apparently 'good' scale taps more heavily into one facet of interest than another apparently 'good' scale with the potential to lead to different findings. For example, in several studies where multiple measures of attitude-toward-the-ad have been used, the conclusions drawn about the significant relationships of one version were different from those based on the other version(s).[13–16] This diversity of operationalisation is a possible reason why significant variation in findings was found across studies in a major meta-analysis of advert attitudes.[17]

## Equivalency

While standardisation has to do with use of the same scale, equivalency suggests that it is possible to use alternative measures of the same construct and yet arrive at the same conclusion. Analyses can be conducted to help determine the degree of equivalency between scales that are supposed to measure the same construct. Measures can range from being parallel to congeneric to arbitrary.[18] If it is established that a set of scales are parallel then it can be safely assumed that use of any of them will lead to the same conclusion. This is also possible with scales that are congeneric if

**Table 1:** Examples of scale proliferation in scholarly marketing research

| Construct | Sample of different scales | | |
|---|---|---|---|
| Attitude-toward-the-ad | Homer and Kahle[49] | Neese and Taylor[50] | Rosbergen, Pieters and Wedel[51] |
| Attitude-toward-the-brand | Mitchell and Olson[52] | Peterson, Wilson and Brown[53] | Bhat, Leigh and Wardlow[54] |
| Innovativeness | Darden and Perreault[55] | Goldsmith and Hofacker[56] | Donthu and Garcia[57] |
| Job satisfaction | Lucas et al.[58] | Rich[59] | Yilmaz and Hunt[60] |
| Organisational commitment | Dubinsky and Hartley[61] | Sager[62] | Maltz and Kohli[63] |
| Price consciousness | Barak and Stern[64] | Lichtenstein, Ridgway and Netemeyer[65] | Huff and Alden[66] |
| Product quality | Petroshius and Monroe[67] | Keller and Aaker[68] | Grewal et al.[69] |
| Purchase intention | Kilbourne, Painton and Ridley[70] | Dodds, Monroe and Grewal[71] | Putrevu and Lord[72] |
| Role clarity | Busch[73] | Fry et al.[74] | Kohli and Jaworski[75] |
| Shopping enjoyment | O'Guinn and Faber[76] | Shim and Gehrt[77] | Beatty and Ferrell[78] |

This list was limited to three different scales for ten different constructs. A more detailed list is available upon request from the author.

transformation rules are known and are applied.[19] That assumption cannot be made with arbitrary scales since they are merely purported to measure the same construct.

When justified, covariance structure analysis can be a powerful means of assessing equivalency. Another possibility is similarity analysis.[20,21] As argued by Steenbergen,[22] the advantage of similarity analysis over covariance structure analysis is that it can be used with small samples and is more amenable to the exploration of structure. There are even ways of doing some preliminary equating of measures when few or no data from respondents are available.[23] This amounts to employing item response theory and 'collateral information' such as expert opinions or theory.

As important as equivalency is argued to be when multiple scales are used to measure the same construct in the same culture, it becomes even more important in cross–cultural research. A fundamental issue is to determine to what extent findings represent 'real' differences in cultures or, instead, are merely measurement artifacts. Douglas and Craig[24] have discussed the various types of equivalency that should be examined when conducting multinational studies, and Mullen[25] has illustrated empirical techniques for testing them.

Even though it is possible to examine equivalency as a means of justifying the use of new measures, it is rarely done. Instead, it has not been unusual for new measures of a construct to be developed even when one or more established scales exist. Calling this a 'horrendous state of affairs,' Jacoby criticised this practice and gave examples apparent in the late 1970s.[26] The practice has continued, however, as illustrated in Table 1. These appear to be some of the 'worst' cases of scale proliferation in scholarly marketing research that have occurred since Jacoby made his observations.

How is this practice defended? It has been argued that if a new scale's items have been drawn from the same semantic domain as other measures they can be assumed to have a similar amount of common core.[27] This is based on the domain sampling model in which 'parallel' scales can be developed and lead to essentially the same conclusions when used in empirical research.[28] Although conceptually appealing, the equivalency of scales depends upon the degree to which the items composing two or more scales were randomly sampled from the same well–defined domain. In reality, it is unlikely that such a process has been used much in marketing. Instead, items have been constructed which, at best, produce 'alternate form' scales with unknown

statistical equivalence. This introduces various forms of errors, notably systematic differences, that are not covered by the domain sampling model.[29]

### Justification

Justification is defined here as providing adequate reasoning for the use of a particular measure. One type of justification is evidence of psychometric quality. This is most called for when a new scale is used. Scale creators are expected to at least provide evidence of unidimensionality and internal consistency if not convergent and discriminant validity as well.[30,31] Providing this sort of evidence is also important for measures that are modified in some way. Lessons could be learned from the field of psychology where a *primary* standard of measure use (as opposed to a conditional standard) is that it is incumbent upon subsequent users of a measure who modify it in some way to revalidate it or at least explain why additional validation is unnecessary.[32]

Even when an established scale is used is is helpful to provide some limited justification. It may be as simple as citing the relevant sources where evidence of the scale's validity can be found. When such information is not provided readers must guess whether the scale is new, adapted or borrowed as well as the degree to which it is a valid measure and how it equates to previous measures. The lack of clear answers to these questions decreases the confidence that can be placed in the findings associated with the scale.

Another type of justification is called for when a new scale is developed even though others are available. Indeed, some reasons for not using pre-existing scales are quite defendable. For example, the available scales may be considered to be too complex or lengthy for a usage

situation. Yet other possibilities are that available scales lack sufficient validity or claims of validity made for them are in doubt.[33] Putting exceptions such as these to one side, the rule is that researchers should use previously developed measures unless they can explain why it is not possible or appropriate.[34,35]

## AN EMPIRICAL ILLUSTRATION

As stated above, standardisation, equivalency and justification may be considered unnecessary depending upon one's understanding of the domain sampling model. Specifically, some researchers may believe that if a scale appears on the surface to tap into the same semantic domain as a pre-existing scale then it is pointless to provide confirming evidence of that judgment. The shortcoming of this assumption is demonstrated here as simply as possible by showing the potential consequences of using different scales to measure the same construct. Researchers may not be fully aware of the potential inconsistency in conclusions that can result because it is rare for them to use two or more multi-item scales in a study to measure the same construct. Instead, what is presented here is more akin to what would happen if the same relationships were being examined in studies conducted independently by researchers utilising different measures.

An experiment was designed to compare the effects of different visual backgrounds on a variety of responses to an advertisement at a website. Specifically, it was anticipated that an advertisement in the context of an elaborate, dynamic background would be associated with poorer attitude-toward-the-ad and attitude-toward-the-brand than plainer backgrounds.[36] The sample was composed of 90 college students

randomly assigned to three treatments. A little over half of the sample was male (59 per cent) and 54 per cent were between the ages of 20 and 24.

Included in the questionnaire were two mutually exclusive sets of items for measurement of attitude-toward-the-ad (Aad) and attitude-toward-the-brand (Ab). One set of items (Aad1) was derived from Mitchell and Olson[37] and has been used since the 1980s by many others. The second scale (Aad2) was taken from Muehling, Stoltman and Mishra[38] (1990), a popular nucleus of items that has been augmented with other items in various studies. Regarding attitude-toward-the-brand, version one has been used by Miniard, Bhatla and Rose[39] and others while the second version of attitude-toward-the-brand has been used by Loken and Ward[40] among others. (Items for all scales as well as their respective reliabilities are shown in the Appendix.)

The effect of webpage background on attitude-toward-the-ad and attitude-toward-the-brand was tested by means of a MANOVA as well as complementary ANOVAs. The results are presented in Table 2. A total of four MANOVAs were run, each using a different pair of attitude-toward-the-ad and attitude-toward-the-brand scales. This was intended to simulate the varying conclusions that could be drawn in four independent studies of the same relationships if different pairs of scales were used.

The results indicated that Aad1 and Ab1 were significantly influenced by the independent variable (different webpage backgrounds) whereas Aad2 and Ab2 were not. This, in turn, led to the varying results of the MANOVAs. Depending upon the set of scales used, the conclusions could range from full support for the hypothesised relationships, to partial support, to no support. More

generally, this illustration shows how the use of different scales to measure the same relationships can lead to different conclusions even though the scales may superficially appear to be acceptable alternative measures of the focal constructs.

## DISCUSSION

The primary call made here is for standardisation. The ideal way to accomplish this is to have one theoretical conceptualisation for primary constructs eg attitude-toward-the-ad, attitude-toward-the-brand, etc.) as well as a preferred measure for each of them. That is, however, an unrealistic goal for the near future. Alternative views are likely to coexist for some time as will a variety of measures related to each conceptualisation.

In the meantime, the procedure outlined below is offered as a guide for the selection and development of scales. With the spirit of standardisation in full view, the emphasis is on effort to identify and use previously developed scales. One reason researchers may have over-developed scales in the past is simply because they were unaware that acceptable scales were already available. This has become a less defendable excuse since the 1990s, however, due to the ability to easily conduct computerised searches of databases in libraries and the web.

Further, books totally devoted to the topic have been published within marketing[41,42] as well as related fields.[43-45]

### Recommended scale selection and usage process

1. Determine the latent construct to be measured.
2. Determine if a multi-item scale is an

---

appropriate type of measure for empirically operationalising a construct. A ratio level measure may be possible and preferable while in other cases multi-item scales may be not be possible due to survey length/time restrictions.

3. If a multi-item scale is appropriate, search to determine if an acceptable scale exists. Conduct computerised searches of internet or other digitised databases as well as looking in the available scales handbooks.

4. If alternative scales are found, compare them using several criteria:

   a. face validity — does the scale appear to capture the meaning one intends? It is quite possible that a scale is called one thing by one researcher but is referred to as something else by another researcher;

   b. psychometric quality — what support is available attesting to the scale's unidimensionality, reliability, validity as well as norms for use with different samples;

   c. typicality/acceptance — researchers should have an understanding of the alternative views of the construct that may exist and the scales most accepted by the different schools of thought. Ultimately, choice of a scale should be influenced by knowledge of what scales were used by previous researchers on whose work the current researchers hope to build;

   d. once the scale has been selected and used in a study, it should be sufficiently described in any scholarly paper written about the research. A minimal amount of information about the scale should be provided: (i) source — the origin of the scale should be specified. It may also be helpful to

cite a recent use or two of the scale in the field if the scale was developed many years earlier or originated in another field; (ii) reliability — an estimate of the scale's reliability should be reported. While measuring stability (test-retest) may be popular in other fields,[79] internal consistency is overwhelmingly the most prevalent type reported in marketing;[80] (iii) indications of any substantial changes that were made. If items were added or deleted, particularly if the result of factor analysis, then that should be described. Rephrasing of items and use of a different response format (points, anchors) would be helpful to report as well.

5. If no scale is found or those found are unacceptable then a measure using widely accepted procedures[81,82] should be developed:

   a. when the research is reported, some rationale for the scale's construction should be provided, eg no known scale was available or those available were inadequate in some specific way;

   b. it is incumbent upon the scale developer to provide more information about the scale's content and psychometric quality in this case than in the case of 4d (above). Chief among the added information would be: (i) the scale items, response format (points, anchors), and evidence of unidimensionality, reliability and validity should be provided; (ii) if the new scale is being offered as an alternative to an established scale the equivalency of the two should be tested and reported; (iii) if such information cannot be included

**Table 2:** The effect of webpage background on advert effectiveness

| Variable | F-Ratio | Sig. level | Means | | |
|---|---|---|---|---|---|
| | | | Webpage1 | Webpage2 | Webpage3 |
| Multivariate test | | | | | |
| (Wilk's lambda) | 2.496 | 0.045 | | | |
| Univariate tests | | | | | |
| Aad1 | 3.474 | 0.035 | 5.525* | 5.158 | 4.650 |
| Ab1 | 3.987 | 0.022 | 4.911* | 4.700 | 4.200 |
| Multivariate test | | | | | |
| (Wilk's lambda) | 1.939 | 0.106 | | | |
| Univariate tests | | | | | |
| Aad1 | 3.474 | 0.035 | 5.525* | 5.158 | 4.650 |
| Ab2 | 2.219 | 0.115 | 4.711 | 4.611 | 4.200 |
| Multivariate test | | | | | |
| (Wilk's lambda) | 2.099 | 0.083 | | | |
| Univariate tests | | | | | |
| Aad2 | 1.662 | 0.196 | 5.411 | 5.033 | 4.744 |
| Ab1 | 3.987 | 0.022 | 4.911* | 4.700 | 4.200 |
| Multivariate test | | | | | |
| (Wilk's lambda) | 1.406 | 0.234 | | | |
| Univariate tests | | | | | |
| Aad2 | 1.662 | 0.196 | 5.411 | 5.033 | 4.744 |
| Ab2 | 2.219 | 0.115 | 4.711 | 4.611 | 4.200 |

*In *post hoc* comparisons (Tukey HSD) the means associated with Webpage1 were found to be statistically different ($p \leq 0.05$) from those associated with Webpage3. None of the means associated with Webpage1 were statistically different from those associated with Webpage2.

with the published article due to length restrictions, an appendix should be made available directly from the authors or from the publishing journal's website.

One of the tenets of the proposed paradigm is that development of new scales should be conducted only if necessary and the process ought to carry with it certain validation and reporting responsibilities. Understanding of the domain sampling model should include the realisation that the judgment of researchers will differ in the definition of a construct (the semantic domain) as well as what constitutes an adequate selection of items from that domain. The burden of evidence is on developers of a new scale, not reviewers or readers; authors must explain why existing scales are inadequate and cannot be used.

The benefits of following this guide can accrue to individual researchers as well as to the discipline as a whole. For researchers themselves the benefits come in the form of efficiency and effectiveness. They are more efficient because it probably takes less time to locate and use a previously published scale of acceptable psychometric quality than it does to create a new one of similar quality. Researchers can be more effective because when they use the same measures as were employed in the work on which they are building then they can be more confident in their conclusions than if a totally new measure is used.

There are several benefits for the field as well. If a growing proportion of researchers follow these guidelines then the unwanted variance in findings should be decreased due to the reduction in the variety of scales being used to measure the same thing (method variance). As the variance in findings is reduced then the quantity and quality of principles upon which marketing operates will be improved. This in turn is one of the

requirements for moving the discipline down the road towards being a science.

It is acknowledged that typical users of scales are unlikely to go to the effort of providing ample evidence of reliability, validity and equivalency (with previous measures) for every scale they use. They would rather get on with testing relationships. Given this, the field would benefit from more journal articles that critically compare competing scales that have been developed for popularly measured constructs. Good examples are the work done on innovativeness,[46] consumer values[47] and market orientation.[48] In some cases, the reviews might only judge the relative superiority of a scale over others based upon a review of the evidence from secondary sources. Ideally, however, there will be many more studies conducted that gather primary data to simultaneously test alternative measures with analyses that are rigorous enough to gauge the equivalency of the scales. The value of these articles is that other researchers can refer to them to make more informed choices among available measures. Such articles could also note areas for future research such as when all available scales have serious limitations and the development of better measures is justified.

Short of having the results of such reviews and comparisons as guides, there should be a presumption of difference between scales. This is in contrast to what seems to be occurring currently, where as long as a researcher uses the same or similar name for a measure as used by previous researchers and provides some limited evidence of its reliability then the scale and its associated findings are accepted with little question. Given this, it is recommended that literature reviews, meta-analyses and syntheses of findings across studies be conducted more cautiously. If the scales used in a set of studies are the same or very similar

then comparison of results may be safe. In contrast, comparison of findings across studies with different scales should be sensitive to the strong possibility that disparate conclusions may have been reached by the various researchers due to the method variance of the type described in this paper.

## SUMMARY AND CONCLUSIONS

The purpose of this paper has been to describe a problem that is afflicting marketing as it did its parent disciplines: ie scale proliferation. Not only is this practice wasteful but it is probably injecting large amounts of undesirable variance into results and conclusions. The primary solution advocated here is standardisation, use of the same scale by all who want to contribute to a stream of research. Short of that, use of another scale may be necessary and/or desirable at times but only if justification is provided. Part of the justification that is called for when an alternative scale is used rather than an established one is evidence of equivalency. Thus, these three qualities of scale usage, standardisation, justification and equivalency deserve much greater attention than given in the past if marketing is to advance beyond an art to become a science.

### References
1 Bruner II, G. C. (1998) 'Standardization and justification: Do ad scales measure up?', *Journal of Current Issues & Research in Advertising*, Vol. 20, Spring, pp. 1–18.
2 Nunnally, J. and Bernstein, I. (1994) 'Psychometric theory', McGraw-Hill, New York, p. 6.
3 Cronbach, L. (1971) 'Test validation', in Thorndike, R. (ed.) 'Educational measurement', American Council on Education, Washington, DC, pp. 443–507.
4 Peter, J. (1981) 'Construct validity: A review of basic issues and marketing practices', *Journal of Marketing Research*, Vol. 18, May, pp. 133–145.
5 Rentz, J. (1987) 'Generalizability theory: A comprehensive method for assessing and improving

the dependability of marketing measures', *Journal of Marketing Research*, Vol. 24, February, pp. 19–28.

6 Finn, A. and Kayandé, U. (1997) 'Reliability assessment and optimization of marketing measurement', *Journal of Marketing Research*, Vol. 34, May, pp. 262–275.

7 Churchill, G., Jr. (1979) 'A paradigm for developing better measures of marketing constructs', *Journal of Marketing Research*, Vol. 16, February, pp. 64–73.

8 Furse, D. and Stewart, D. (1982) 'Standards for advertising copytesting: A psychometric interpretation', *Journal of Advertising*, Vol. 11, Winter, pp. 30–38, 76.

9 Chun, K., Cobb, S. and French, J. (1975) 'Measures for psychological assessment', Institute for Social Research, University of Michigan, Ann Arbor, Michigan, p. x.

10 Bonjean, C., Hill, R. and McLemore, S. (1967) 'Sociological measurement: An inventory of scales and indices', Chandler Publishing Company, San Francisco.

11 Chun, K., Barnowe, J., Cobb, S. and French, J. (1974) 'Publication and uses of psychological measures in the 1960s', Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

12 King, M. F. and Bruner II, G. C. (1993) 'Evaluation of consumer-related scales: A structured schemata approach', Office of Scale Research, Technical Report, No. 9301, Southern Illinois University, www.siu.edu/departments/coba/mktg/osr/

13 Burton, S. and Lichtenstein, D. (1988) 'The effects of ad claims and ad content on attitude toward the advertisement', *Journal of Advertising*, Vol. 17, Spring, pp. 3–11.

14 Miniard, P., Bhatla, S. and Rose, R. (1990) 'On the formation and relationship of ad and brand attitudes: An experimental and causal analysis', *Journal of Marketing Research*, Vol. 27, August, pp. 290–303.

15 Olney, T., Holbrook, M. and Batra, R. (1991) 'Consumer responses to advertising: The effects of ad content, emotions, and attitude toward the ad on viewing time', *Journal of Consumer Research*, Vol. 17, March, pp. 440–453.

16 Stafford, M. (1998) 'Advertising sex-typed services: The effects of sex, service type, and employee type on consumer attitudes', *Journal of Advertising*, Vol. 27, No. 2, pp. 65–81.

17 Brown, S. and Stayman, D. (1992) 'Antecedents and consequences of attitude toward the ad: A meta-analysis', *Journal of Consumer Research*, Vol. 19, June, pp. 34–51.

18 Mislevy, R. (2000) 'Technical considerations in marketbasket reporting', presentation at the National Research Council's Workshop on Marketbasket Reporting for the National Assessment of Education Progress, Washington, DC, February 7–8, 2000.

19 Traub, R. (1994) 'Reliability for the social sciences: Theory and applications', Sage Publications, Thousand Oaks, CA, p. 58.

20 Anderson, J. and Gerbing, D. (1982) 'Some methods for respecifying measurement models to obtain unidimensional construct measurement', *Journal of Marketing Research*, Vol. 19, November, pp. 453–460.

21 Hunter, J. (1973) 'Methods of reordering the correlation matrix to facilitate visual inspection and preliminary cluster analysis', *Journal of Educational Measurement*, Vol. 10, Spring, pp. 51–61.

22 Steenbergen, M. (2000) 'Item similarity in scale analysis', *Political Analysis*, Vol. 8, March, pp. 261–283.

23 Mislevy, R., Sheehan, K. and Wingersky, M. (1993) 'How to equate tests with little or no data', *Journal of Educational Measurement*, Vol. 30, No. 1, pp. 55–78.

24 Douglas, S. and Craig, C. (1983) 'International marketing research', Prentice-Hall, Engelwood Cliffs, New Jersey.

25 Mullen, M. (1995) 'Diagnosing measurement equivalencies in cross-national research', *Journal of International Business Studies*, Vol. 26, Third Quarter, pp. 573–596.

26 Jacoby, J. (1978) 'Consumer research: A state of the art review', *Journal of Marketing*, Vol. 42, April, p. 91.

27 Contact the author for quotes from reviewers at top marketing journals that illustrate the view representative of those who place blind faith in the domain sampling model.

28 Nunnally and Bernstein (1994) *op. cit*, Ch. 6.

29 Nunnally and Bernstein (1994) *op. cit*, pp. 250–251.

30 Churchill (1979) *op. cit.*

31 Gerbing, D. and Anderson, J. (1988) 'An updated paradigm for scale development incorporating unidimensionality and its assessment', *Journal of Marketing Research*, Vol. 25, May, pp. 186–192.

32 American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement in Education (NCME) (1985) 'Standards for Educational and Psychological Testing', American Psychological Association, Washington, DC, p. 41.

33 Summers, J. (2001) 'Guidelines for conducting research and publishing in marketing: From conceptualization through the review process', *Journal of the Academy of Marketing Science*, Vol. 29, Fall, pp. 405–415.

34 Churchill (1979) *op. cit.*

35 Varadarajan, P. (1996) 'From the editor: Reflections on research and publishing', *Journal of Marketing*, Vol. 60, October, pp. 3–6.

36 Morrison, B. and Dainoff, M. (1972) 'Advertisement complexity and looking time', *Journal of Marketing Research*, Vol. 9, November, pp. 396–400.

37 Mitchell, A. and Olson, J. (1981) 'Are product attribute beliefs the only mediator of advertising effects on brand attitude?', *Journal of Marketing Research*, Vol. 18, August, pp. 318–332.

38 Muehling, D., Stoltman, J. and Mishra, S. (1990) 'An examination of the cognitive antecedents of attitude-toward-the-ad', *Current Issues and Research in Advertising*, Vol. 12, Spring and Fall, pp. 95–117.

39 Miniard, Bhatla and Rose (1990) *op. cit.*

40 Loken, B. and Ward J. (1990) 'Alternative approaches to understanding the determinants of

typicality', *Journal of Consumer Research*, Vol. 17, September, pp. 111–126.

41 Bearden, W. and Netemeyer, R. (1998) 'Handbook of marketing scales', 2nd edition, Sage Publications, Thousand Oaks, CA.

42 Bruner II, G., James, K. and Hensel, P. (2001) 'Marketing scales handbook — V. III', American Marketing Association, Chicago.

43 McDowell, I. and Claire, N. (1996) 'Measuring health: A guide to rating scales and questionnaires', Oxford University Press, New York.

44 Robinson, J., Shaver, P. and Wrightsman, L. (1999) 'Measures of political attitudes', Academic Press, San Diego, CA.

45 Impara, J. (2001) 'Mental measurements yearbook', 14th edition, Buros Institute, University of Nebraska Press, Lincoln, NE.

46 Flynn, L., Goldsmith, R. and Eastman, J. (1996) 'Opinion leaders and opinion seekers: Two new measurement scales', *Journal of the Academy of Marketing Science*, Vol. 24, Spring, pp. 137–147.

47 Kahle, L., Beatty, S. and Homer, P. (1986) 'Alternative approaches to consumer values: The list of values (LOV) and values and life style (VALS)', *Journal of Consumer Research*, Vol. 13, December, pp. 405–409.

48 Matsuno, K., Mentzer, J. and Rentz, J. (2000) 'A refinement and validation of the MARKOR scale', *Journal of the Academy of Marketing Science*, Vol. 28, Fall, pp. 527–539.

49 Homer, P. and Kahle, L. (1990) 'Source expertise, time of source identification and involvement in persuasion: An elaborate processing perspective', *Journal of Advertising*, Vol. 19, No. 1, pp. 30–39.

50 Neese, W. and Taylor, R. (1994) 'Verbal strategies for indirect comparative advertising', *Journal of Advertising Research*, Vol. 34, March, pp. 56–70.

51 Rosbergen, E., Pieters, R. and Wedel, M. (1997) 'Visual attention to advertising: A segment-level analysis', *Journal of Consumer Research*, Vol. 24, December, pp. 305–314.

52 Mitchell and Olson (1981) *op. cit.*

53 Peterson, R., Wilson, W. and Brown, P. (1992) 'Effects of advertised customers satisfactions claims on consumer attitudes and purchase intentions', *Journal of Advertising Research*, Vol. 32, No. 2, pp. 34–40.

54 Bhat, S., Leigh, T. and Wardlow, D. (1998) 'The effect of consumer prejudices on ad processing: Heterosexual consumers' responses to homosexual imagery in ads', Vol. 27, No. 4, pp. 9–28.

55 Darden, W. and Perreault, W. Jr. (1976) 'Identifying interurban shoppers: Multiproduct purchase patterns and segmentation profiles', *Journal of Marketing Research*, Vol. 13, February, pp. 51–60.

56 Goldsmith, R. and Hofacker, C. (1991) 'Measuring consumer innovativeness', *Journal of the Academy of Marketing Science*, Vol. 19, Summer, pp. 209–221.

57 Donthu, N. and Garcia, A. (1999) 'The internet shopper', *Journal of Advertising Research*, Vol. 39, May/June, pp. 52–58.

58 Lucas Jr., G., Parasuraman, A., Davis, R. and Enis, B. (1987) 'An empirical study of salesforce turnover', *Journal of Marketing*, Vol. 51, July, pp. 34–59.

59 Rich, G. (1997) 'The sales manager as a role model: Effects on trust, job satisfaction, and performance of salespeople', *Journal of the Academy of Marketing Science*, Vol. 25, No. 4, pp. 319–328.

60 Yilmaz, C. and Hunt, S. (2001) 'Salesperson cooperation: The influence of relational, task, organizational, and personal factors', *Journal of the Academy of Marketing Science*, Vol. 29, No. 4, pp. 335–357.

61 Dubinsky, A. and Hartley, S. (1986) 'A path-analytic study of a model of salesperson performance', *Journal of the Academy of Marketing Science*, Vol. 14, Spring, pp. 36–46.

62 Sager, J. (1994) 'A structural model depicting salespeople's job stress', *Journal of the Academy of Marketing Science*, Vol. 22, No. 1, pp. 74–84.

63 Maltz, E. and Kohli, A. (1996) 'Market intelligence dissemination across functional boundaries', *Journal of Marketing Research*, Vol. 33, February, pp. 47–61.

64 Barak, B. and Stern, B. (1985) 'Women's age in advertising: An examination of two consumer age profiles', *Journal of Advertising Research*, Vol. 25, No. 6, pp. 38–47.

65 Lichtenstein, D., Ridgway, N. and Netemeyer, R. (1993) 'Price perceptions and consumer shopping behavior: A field study', *Journal of Marketing Research*, Vol. 30, May, pp. 234–245.

66 Huff, L. and Alden, D. (1998) 'An investigation of consumer response to sales promotions in developing markets: A three-country analysis', *Journal of Advertising Research*, Vol. 38, May/June, pp. 47–56.

67 Petroshius, S. and Monroe, K. (1987) 'Effect of product-line pricing on product evaluations', *Journal of Consumer Research*, Vol. 13, March, pp. 511–519.

68 Keller, K. and Aaker, D. (1992) 'The effects of sequential introduction of brand extensions', *Journal of Marketing Research*, Vol. 29, February, pp. 35–50.

69 Grewal, D., Krishnan, R., Baker, J. and Borin, N. (1998) 'The effect of store name, brand name and price discounts on consumers' evaluations and purchase intentions', *Journal of Retailing*, Vol. 74, No. 3, pp. 331–352.

70 Kilbourne, W., Painton, S. and Ridley, D. (1985) 'The effect of sexual embedding on responses to magazine advertisements', *Journal of Advertising*, Vol. 14, No. 2, pp. 48–56.

71 Dodds, W., Monroe, K. and Grewal, D. (1991) 'Effects of price, brand, and store information on buyers' product evaluations', *Journal of Marketing Research*, Vol. 28, August, pp. 307–319.

72 Putrevu, S. and Lord, K. (1994) 'Comparative and noncomparative advertising: Attitudinal effects under cognitive and affective involvement conditions', *Journal of Advertising*, Vol. 23, June, pp. 77–90.

73 Busch, P. (1980) 'The sales manager's badges of social power and influence upon the sales force', *Journal of Marketing*, Vol. 44, Summer, pp. 91–101.

74 Fry, L., Futrell, C., Parasuraman, A. and Chimelewski, M. (1986) 'An analysis of alternative causal models of salesperson role perceptions and

work-related attitudes', *Journal of Marketing Research*, Vol. 23, May, pp. 153–163.

75 Kohli, A. and Jaworski, B. (1994) 'The influence of coworker feedback on salespeople', *Journal of Marketing*, Vol. 58, October, pp. 82–94.

76 O'Guinn, T. and Faber, R. (1989) 'Compulsive buying: A phenomenological exploration', *Journal of Consumer Research*, Vol. 16, No. 2, pp. 147–157.

77 Shim, S. and Gehrt, K. (1996) 'Hispanic and Native American adolescents: An exploratory study of their approach to shopping', *Journal of Retailing*, Vol. 72, No. 3, pp. 307–324.

78 Beatty, S. and Ferrell, M. (1998) 'Impulse buying: Modeling its precursors', *Journal of Retailing*, Vol. 74, No. 2, pp. 169– 191.

79 Litwin, M. (1995) 'How to measure survey reliability and validity', Sage Publications, Thousand Oaks, CA, p. 8.

80 Peterson, R. (1994) 'A meta-analysis of Cronbach's coefficient alpha', *Journal of Consumer Research*, Vol. 21, September, pp. 381–391.

81 Churchill (1979) *op. cit.*

82 Gerbing and Anderson (1988) *op. cit.*

# APPENDIX

## Characteristics of scales examined

Attitude–toward–the–ad (Aad1), $\alpha = 0.91$

1 good/bad
2 like/dislike
3 irritating/not irritating
4 interesting/uninteresting

Attitude–toward–the–ad (Aad2), $\alpha = 0.91$

1 unappealing/appealing
2 unattractive/attractive
3 unpleasant/pleasant

Attitude–toward–the–brand (Ab1), $\alpha = 0.89$

1 dislike/like
2 unfavourable/favourable
3 negative/positive

Attitude–toward–the–brand (Ab2), $\alpha = 0.91$

1 bad/good
2 poor quality/high quality
3 unsatisfactory/satisfactory