

**An Alternative Approach to Measuring Schwartz's Values:
The Best-Worst Scaling Approach**

Dr Julie Anne Lee

Graduate School of Management, University of Western Australia
julielee@gsm.uwa.edu.au

Prof Geoffrey Soutar

Graduate School of Management, University of Western Australia
gsoutar@gsm.uwa.edu.au

Prof Jordan Louviere

School of Marketing, University of Technology, Sydney
Jordan.louviere@uts.edu.au

An Alternative Approach to Measuring Schwartz's Values: The Best-Worst Scaling Approach

This paper outlines an alternative method to the more traditional way of measuring people's values. A recently developed approach, termed the "Best-Worst Scaling" technique, which has been developed by Louviere and his colleagues, offers an opportunity to collect such data in a way that may produce more desirable estimates. Schwartz's (1992, 1994) value structure was used to examine differences between this new approach and the traditional Schwartz Values Survey approach. The present paper reports the results that suggest the new technique has some advantages over the more traditional approach.

Key Words: *Rating scales, Ranking, Best-worst scaling, Schwartz values survey*

Values are guiding principles that determine what is important to us (Rokeach, 1973; Kahle, 1983). They have been defined as "enduring beliefs that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or end-state" (Rokeach, 1973, p. 5). Values are thought to be relatively stable in adults and to motivate people's behavior (e.g., Rokeach, 1973; Schwartz 1992; Kahle, Rose, & Shoham, 1999). Rokeach and other values researchers have suggested that values reflect real differences between cultures, social classes, occupations, religions and political orientations. Consequently, it is not surprising that researchers have long felt that values impact on decisions and have tried to investigate this impact (e.g., Carman 1977; Vinson, Scott, & Lamont, 1977; Prakash & Munson 1985; McCarty & Shrum, 1993; Donthu & Cherian 1994).

Several instruments have been developed to measure values, including Rokeach's (1973) Value Survey (RVS), Kahle's (1983) List of Values (LOV), Mitchell's (1983) Values and Life Styles (VALS) survey and Schwartz's (1992) Values Scale (SVS). Each of these instruments use some form of rating or ranking scales to obtain the needed data. The ranking method was used in most of the early research on values, but increasingly, rating scales have become the preferred method (see Krosnick & Alwin, 1989).

Rating scales have several advantages and disadvantages as compared to ranking for respondents and researchers. On the positive side, rating scales are easier for respondents to answer (e.g., Rokeach, 1973; Alwin & Krosnick, 1985; Schwartz, 1994), allow longer lists of values to be asked, enable the measurement of "negative" values and have better statistical properties (Schwartz, 1994). On the negative

side, people often use different parts of a scale, or “response styles,” which can affect both the mean and the variance obtained (Craig & Douglas, 2000; Baumgartner & Steenkamp, 2001). For instance, people often endorse most values as being important in their life, leading to skewed responses and so-called “end-piling” (Hood, 2003; Mattila, 1999). Indeed, Schwartz and Bardi (2001) noted that responses to values items generally range from “somewhat” to “very important” for each individual. Thus, the order of importance is often what is compared across samples. This skewing tendency can be compounded by high correlations between different value types (Krosnick & Alwin, 1989). While this should be expected for “similar” or “compatible” values, the pattern of correlations between dissimilar values should be negative, or non-significant, if there is discrimination between values items.

Optimally, the way in which values are measured should not impact on associations between values or between values and behaviour. Many studies have shown that rankings and ratings produce a similar aggregate ordering of values (e.g., Alwin & Krosnick, 1985; Feather, 1973, 1975; Moore, 1975). However, differences have been found when the correlational structures of ranking and rating data have been compared (e.g., Rankin & Grube, 1980; Reynolds & Jolly, 1980, Alwin & Krosnick, 1985, 1989). Since response bias is a systematic tendency that distorts ratings so that they do not represent the true score, many researchers standardize raw scores to reduce or eliminate unwanted differences (Fischer, 2004). The standardization of raw rating scores is often used to adjust the means and/or the standard deviations of individuals, groups or cultures, depending on the type of analysis being conducted (see Fischer, 2004 for a full discussion of the variety of standardization procedures). Many authors suggest ipsatizing, or centering, responses around an individual's overall mean or using this mean as a covariate to reduce such effects (e.g., Schwartz, 1992). The problem is, however, that this process may remove true differences from the data (e.g. Smith, 2004; Van Hemert, *et al.*, 2002).

Consequently, a measurement approach that overcame these issues would be a valuable addition to researchers examining values. This paper discusses Best-Worst Scaling (BWS), which is such an approach. In the paper, the performance of a traditional rating scale (the Schwartz Values Survey) is

compared with a BWS approach used to obtain similar values data. Given the importance of Schwartz's values approach to the present study, it is discussed in some detail in the next section.

SCHWARTZ'S VALUES APPROACH

While there are different ways to measure values across cultures, Schwartz has developed a set of what he terms "universal values" that have a strong theoretical underpinning. He derived ten types of values from responses to; (1) our needs as biological organisms, (2) our need to coordinate social interaction, and (3) our need for groups to function smoothly and survive (Schwartz, 1994). The value types, including power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity and security, have been confirmed in more than 100 studies undertaken in over 50 countries (Schwartz, 1992, 1994; Schwartz & Sagiv, 1995).

Schwartz (1992, 1994) developed the 57-item Schwartz Value Survey (called the SVS), which can be used to compute the ten value types. Each item in the scale is followed by a short description (e.g. CREATIVITY [uniqueness, imagination]) and respondents are asked how important the item is for them as a guiding principle in their life. They are asked to choose the one item that is most important and rate its importance. Next, they are asked to choose the value that is most opposed to their values and rate it as -1 or, if there is no such value, to rate the least important value as 0 or 1. They are then asked to rate the remaining 55 values on a 9-point scale (which ranges from -1 [opposed to my principles], through 0 [not important], 3 [important], 6 [very important] to 7 [of supreme importance]). Responses to the items associated with each of the 10 value types are averaged to obtain a score for each value type (Schwartz, 1992, 1994). A higher score for a value indicates that the respondent attaches greater importance to that value. Schwartz and Bardi (2001; p. 271) noted that, "people typically view values as desirable (and) values generally range from somewhat to very important".

THE BEST-WORST SCALING (BWS) APPROACH

The BWS technique was originally designed for use in Best-Worst Conjoint Analysis (see Louviere, Swait and Anderson 1995). While most BWS applications have been in Best-Worst Conjoint Analysis contexts, Cohen and Markowitz (2002) recently used the technique to measure the importance of product benefits. They placed a list of benefits into smaller sets through an experimental design and asked respondents to choose the most and least important attribute in each set. This procedure produces a unidimensional interval-level importance scale based on the nominal level data provided by respondents (see Marley and Louviere 2005). We used a similar method to develop our BWS task, as outlined below.

While the BWS task used a similar item-description format to the SVS (previously described), it differed in the following ways. First, a balanced incomplete block (BIB) experimental design was chosen to create subsets of Schwartz 10 value types. This resulted in 11 subsets, each consisting of six value types. The characteristics of the design meant that each respondent saw each value type six times and each pair of value types three times. This resulted in multiple measures for each of the 10 value types. In addition, the complement of the design was also used, in which each respondent saw each value type five times.

Second, the three value items with the strongest reliability across cultures (see Spini 2003), were chosen to represent each of the ten value types in the experimental design (e.g., creativity, curious and freedom for Self-Direction). Spini's (2003) study reported the number of times the various SVS items appeared in their correct value type location in 185 studies and the results were used to determine the three items used in the present study. The items chosen ranged from 100% correct location for Social Power to 81% correct location for Social Order.

Third, Schwartz short description of each value item was imbedded into the webpage, so respondents could read the description as they moved their mouse over any of the words. For example when people held their mouse over the word CREATIVITY, Schwartz's description [uniqueness, imagination] appeared. Respondents were given the following instruction:

“In this section, we will ask you to pick the most and least important values that guide your life. These values come from different cultures. While more than one may be important or unimportant, please choose the MOST and the LEAST important to YOU as a guiding principle in YOUR life. There are 11 sets of statements in this section.”

Which is the MOST and LEAST important factor to you as a guiding principle in YOUR life?

Most Important		Least Important
O	<u>Successful, capable, ambitious.</u>	●
●	<u>Protecting the environment, a world of beauty, unity with nature.</u>	O
O	<u>Helpful, honest, forgiving.</u>	O
O	<u>Devout, accepting portion in life, humble.</u>	O
O	<u>Clean, national security, social order.</u>	O

BWS scores were calculated for each individual by counting the number of times each item was chosen as the most or the least important and subtracting the number for least important from the number for most important. This number was then divided by the number of times the item appeared in the survey, creating a scale (-1 to +1) where high scores implying a value was more important to the respondent.

METHOD

The data were obtained from a large study of consumers' travel plans. The study was conducted on the Internet, using a sample from a large online consumer panel in Western Australia. The sample was chosen to reflect the age, gender, income and regional characteristics of the State's residents. The sample was limited to people aged from 18 to 75 years who do not work in advertising, marketing research or tourism.

Respondents were randomly assigned to answer a questionnaire that contained either the SVS or the BWS task. This resulted in a sample of 202 people completing the SVS and 404 people completing the BWS

task. While individual question completion time was not measured in this study, two versions differed only in the way in which values were measured. On average, the SVS questionnaire took 12 minutes longer for respondents to complete than did the BWS questionnaire.

RESULTS

As discussed earlier, values research has been notable for respondents tending to rate most items as important, leading to highly skewed responses (Hood, 2003; Mattila, 1999) and significant positive correlations between the various values items. Consequently, descriptive statistics for the 202 respondents who answered the SVS items using the traditional 9-point scale (ranging from -1 to 7) and the 404 respondents who answered the BWS task were computed and are shown in Table 1.

TABLE 1: Descriptive Statistics

Value Item	SVS				BWS			
	SVS Mean	Std Dev.	Skew	Std Error	BWS Mean	Std Dev.	Skew	Std Error
Benevolence	4.51	1.23	-.34	.17	.40	.37	.00	.12
Universalism	4.36	1.34	-.40	.17	.13	.42	.11	.12
Security	4.36	1.10	.22	.17	.06	.44	.00	.12
Self Direction	4.36	1.15	-.02	.17	.18	.36	.25	.12
Conformity	4.36	1.29	-.01	.17	.13	.41	.03	.12
Achievement	4.14	1.15	-.13	.17	-.03	.44	.61	.12
Hedonism	3.58	1.48	-.03	.17	-.02	.43	.35	.12
Stimulation	3.29	1.49	.17	.17	-.09	.40	.24	.12
Tradition	2.97	1.35	.27	.17	-.22	.43	-.11	.12
Power	1.89	1.22	.59	.17	-.55	.50	1.06	.12

As can be seen from Table 1, eight of the ten SVS means were above the scale midpoint and the range was less than 3 on the 9-point scale used. In contrast, five of the ten BWS means were above the midpoint and the range covered almost half of the BWS scale (-1 to +1). While there were differences in the

dispersion of the means for the two methods, there were also similarities. For instance, the most important value (Benevolence) and least important values (Power, Tradition and Stimulation) were the same using both methods. While the rank orders were different for some of the values, this was mainly due to the low variation in the rating scale means, with four of the ten SVS items being very similar (all scoring 4.36).

Interrelationship between the Value Types

Schwartz (1992, 1994) has argued that the value dimensions follow a circumplex structure, where adjacent values are compatible and opposite values conflict. Thus, the overall set of relationships between the ten value types can be examined through the creation of multidimensional maps, as Schwartz (1992, 1994) has noted on a number of occasions. Maps were developed for the ten values dimensions obtained from the SVS scores¹ and those obtained from the BWS scores, through the MRSCAL computer program (Roskam, 1972). As Schwartz's implied structure is two dimensional, the two-dimensional maps are shown in Figure 1 and Figure 2.

The maps obtained generally support Schwartz suggested structure. However, as can be seen from Figure 1, the SVS data provided a less coherent structure and much less discrimination between the dimensions than did the BWS data, which is shown in Figure 2. Further, some dimensions were misplaced in the SVS map in Figure 1, as achievement and security should have been found on either side of power but were found on the opposite side of the circle closer to benevolence. On the other hand, the BWS dimensions in Figure 2 only show benevolence as slightly out of position. Following Schwartz et al. (2001), the observed order of the dimensions around the circle for the two approaches were compared with the hypothesized order by ranking self-direction as 1 and assigning ranks in an appropriate direction so that universalism had a rank of 10. The Spearman rank correlation between the data and the hypothesised order was higher for BWS (0.99), than for SVS (0.73).

¹ While Schwartz recommends various forms of correction prior to analysis, he recommends that the raw scores be used for multidimensional scaling. Thus, the raw scores were used to produce the SVS map.

FIGURE 1: Two Dimensional Map – Ratings Data

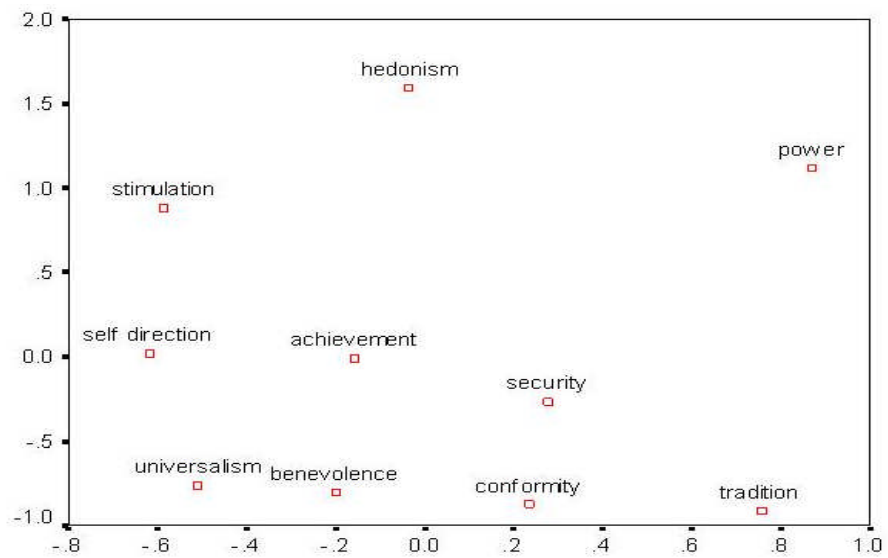
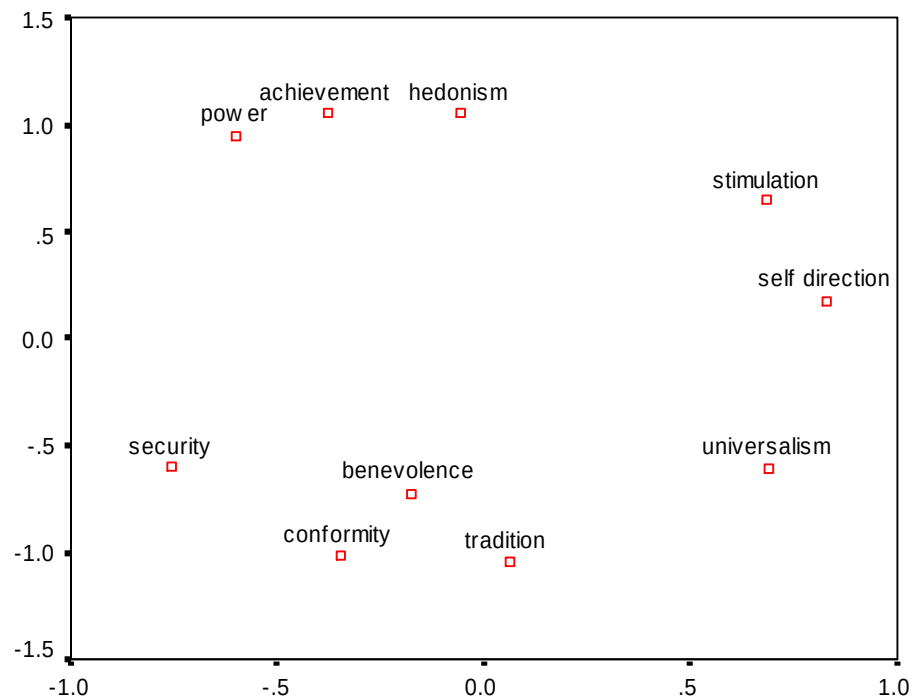


FIGURE 2: Two Dimensional Map – BWS Data



Relationship of Value Types to Behaviors

Bardi and Schwartz (2003) suggested sixty behaviours that relate to different values. We chose twenty of the most appropriate behaviors (two for each value type) and respondents were asked how frequently they engaged in these behaviors using a 5-point scale that ranged from never (1) to all the time (5) (taking into account their ability to perform the behavior). The mean scores for the behaviors, ranged from relatively low scores (e.g., networking was 2.40) to relatively high scores (e.g., making my own decision was 4.21), while the standard deviations ranged 0.70 to 1.06, suggesting there was reasonable variation.

The raw means for the SVS method led to positive correlations between almost all of the value types. Consequently, it was not surprising that, when one value correlated significantly with a behavior, most of the other values also tended to do so. Thus, the unstandardized SVS method, led to 111 of the potential 200 correlations being significant ($p < .05$). Of these, 86 were found with behaviors that should have corresponded to other values. This number was reduced to 55, when the correlations with “neighboring” behaviors were discounted. Both the standardized SVS method (i.e. the partial correlations controlling for individuals SVS and behavior means) and the BWS method performed better. The standardized SVS method produced 20 significant correlations with non-corresponding behaviors and only seven when “neighboring” behaviors were discounted. Similarly, the BWS method produced 28 significant correlations with non-corresponding behaviors and only eight when “neighboring” behaviors were discounted. Thus, BWS method performed about as well as the standardized SVS method when behaviors were examined.

GENERAL DISCUSSION AND CONCLUSIONS

This study compared the results obtained from the BWS approach with those obtained from the traditional ratings approach to measuring Schwartz’s values. First, it appears that the influence of “response styles” is reduced in the BWS data. The BWS measures were evenly distributed across almost half of the scale without standardization, while the SVS raw score means were distributed over only one third of the scale

and eight of the 10 means were above the scale mid-point. After standardization (i.e., controlling for individual respondent's means as recommended by Schwartz) the SVS means performed about as well as the BWS means, but the BWS means did this without any loss of information. Second, the BWS data better represented Schwartz value circumplex, than the SVS data when multidimensional maps were created, based on the correlations between the ten dimensions. Third, the BWS data produced the hypothesized relationships between the values and predicted behaviours (suggested by Bardi and Schwartz 2003), better than the SVS raw scores and as well as the SVS standardized scores, without any loss of information.

Another significant benefit of the BWS method, over the SVS, is the reduction in task time for respondents. On average the BWS task took 12 minutes less than the SVS task for respondents to complete. This offers a significant benefit to researchers who may be considering the inclusion of Schwartz values into their surveys.

The BWS approach also offers advantages to cross-cultural researchers in terms of cross-cultural equivalence. Where the SVS includes multiple verbal scale terms (-1= opposed to my values, 0= not important, 3=important, 6=very important, and 7=of supreme importance), the BWS method only has two verbal scale terms (most important and least important). This reduces the problems associated with lexical equivalence as it is easier to find equivalent terms for "most" and "least" in different languages, than the finer grained verbal terms employed in the SVS method. The BWS method also eliminates problems that may be found when numbers have meanings in certain cultures, as it does not include numerical anchors.

It would seem that the BWS approach is a real alternative for researchers interested in values and the roles they play in people's lives. While further research is necessary to see whether the present results are generalizable to other contexts and to other values approaches (e.g. Kahle's (1983) List of Values approach), the results obtained in the present study suggest such research is desirable and that values researchers should think seriously about collecting BWS data, rather than ratings data.

REFERENCES

- Alwin, G. W., & Krosnick J. A. (1985). The measurement of values in surveys: A comparison of ratings and rankings. *Public Opinion Quarterly*, 49, 535-552.
- Alwin, G. W., & Krosnick J. A. (1989). A test of the form-resistant correlation hypothesis. *Public Opinion Quarterly*, 52, 526-538.
- Bardi, A., & Schwartz S.H. (2003). Values and Behavior: Strength and Structure of Relations. *Personality and Social Psychology Bulletin*, 29(10), 1207-1220.
- Baumgartner, H., & Steenkamp, J-B. (2001). Response Styles in Marketing Research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143-156.
- Carman, J. M. (1977). Values and consumption patterns: A closed loop. In H. K. Hunt (Ed.), *Advances in Consumer Research*, Volume 5, Ann Arbor: Association for Consumer Research.
- Cohen, S., & Markowitz, P. (2002). Renewing Market Segmentation: Some new tools to correct old problems, presented at the 54th ESOMAR Congress and Exhibition, Barcelona, Spain.
- Craig, C. S., & Douglas, S. P. (2000). *International Marketing Research*. New York: Wiley.
- Donthu, N., & Cherian, J. (1994). Impact of strength of ethnic identification on Hispanic shopping behaviour. *Journal of Retailing*, 70(4), 383-394.
- Feather, N. T. (1973). The measurement of values: Effects of different assessment procedures. *Australian Journal of Psychology*, 25, 221-231.
- Feather, N. T. (1975). *Values in education and society*. New York: Free Press.
- Fischer, R. (2004). Standardization to Account for Cross-cultural Response Bias: A Classification of Score Adjustment Procedures and Review of Research in JCCP. *Journal of Cross-Cultural Psychology*, 35, 263-282.
- Hood, J.N. (2003). The Relationship of Leadership Style and CEO values to Ethical Practices in Organizations,' *Journal of Business Ethics*, 43(4), 263.
- Kahle, L.R. (1983). *Social Values and Social Change: Adaptation to Life in America*, New York: Praeger Publishers.
- Kahle, L.R., Rose, G., & Shoham, A. (1999). Findings of LOV throughout the World and Other Evidence of Cross-national Consumer Psychographics, *Journal of Euromarketing*, 8(1/2), 1-13.
- Krosnick, J.A., & Alwin, D.F. (1989). A Test of the Form-resistant Correlation Hypothesis, *Public Opinion Quarterly*, 52, 526-538.
- Louviere, J.J., Swait, J. & Anderson, D. (1995). Best-worst Conjoint: A new Preference Elicitation Method to Simultaneously Identify Overall Attribute Importance and Attribute Level Part Worths, Working paper, University of Florida, Gainesville, FL.
- Marley A. A. J., & Louviere J. J. (2005). Some Probabilistic Models of Best, Worst, and Best-Worst Choices, *Journal of Mathematical Psychology*.
- Mattila, A. (1999). Consumers' value judgements. *Cornell Hotel and Restaurant Administration Quarterly*, 40(1), 40-45.
- Mitchell, A. (1983). *The nine American life-styles*, New York: MacMillan.
- Moore, M. (1975). Rating versus ranking in the Rokeach Value Survey: an Israeli comparison. *European Journal of Social Psychology*, 5, 405-408.

The NewMDSX Project (2004) *NewMDSX for Windows*, University of Edinburgh, School of Social and Political Studies: Edinburgh.

Prakash, V. & Munson, J. M. (1985). Values, expectations from the marketing system and product expectations. *Psychology and Marketing*, 2(4): 279-295.

Rankin, W. L., & Grube J. W. (1980). A comparison of ranking and rating procedures for the value system measurement. *European Journal of Social Psychology*, 10, 233-246.

Reynolds, T J., & Jolly J. P. (1980). Measuring personal values: An evolution of alternative methods. *Journal of Marketing Research*, 17, 531-536.

Rokeach, M. (1973). *The Nature of Human Values*. New York: The Free Press.

Roskam, E.E. (1972). MDS by Metric Transformation of Data. *Psychologie*, 27, 486-508.

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1-65.

Schwartz, S. H. (1994). Are there Universal Aspects in the Structure and Content of Human Values? *Journal of Social Issues*, 50(4), 19-45.

Schwartz, S. H., & Bardi, A. (2001). Value Hierarchies across Cultures: Taking a Similarities Perspective. *Journal of Cross Cultural Psychology*, 32, 268-290.

Schwartz, S. H., & Sagiv, L. (1995). Identifying Culture Specifics in the Content and Structure of Values. *Journal of Cross Cultural Psychology*, 26, 92-116.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, 35, 50-61.

Spini, D. (2003). Measurement Equivalence of 10 Value Types from the Schwartz Value Survey across 21 countries. *Journal of Cross-Cultural Psychology*, 34(1), 2-23.

Van Hemert, D.A., Van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structure and score levels of the Eysenck Personality Questionnaire across individuals and countries. *Personality and Individual Differences*, 33, 1229-1249.

Vinson, D. E., Scott, J. E. & Lamont, L. (1977). The role of personal values in marketing and consumer behaviour, *Journal of Marketing*, 41(2): 44-50.