



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Mansholt Paper 03 (2001)

RANGE AND NUMBER-OF-LEVELS EFFECTS IN DERIVED AND STATED MEASURES OF ATTRIBUTE IMPORTANCE*

Peeter W.J. Verlegh¹
Hendrik N.J. Schifferstein²
Dick R. Wittink³

Abstract

We study how the range of variation and the number of attribute levels affect five measures of attribute importance: full profile conjoint estimates, ranges in attribute level attractiveness ratings, regression coefficients, graded paired comparisons, and self-reported ratings. We find that all importance measures are affected by the range manipulation. The number of attribute levels affects only two measures. The results allow us to benchmark the magnitude of the number-of-levels effect against the range effect: conjoint importance estimates were approximately equally affected by a threefold increase in the range of attribute variation and by the insertion of two intermediate attribute levels. Our findings show that the number-of-levels effect is most likely due to respondents' tendencies to distribute their mental stimulus representations and their responses uniformly over the corresponding continua.

Key words: attribute importance, context effects, conjoint analysis

*** Acknowledgements:** The authors thank Zhimin Chen for her assistance with the experiment.

¹ Department of Marketing Management, Erasmus University Rotterdam, The Netherlands, and
Marketing and Consumer Behaviour Group, Wageningen University, The Netherlands
² Department of Industrial Design, Delft University of Technology, The Netherlands
³ Yale School of Management, New Haven, and
Department of Economics, University of Groningen, The Netherlands

Corresponding author: Peeter W.J. Verlegh, Department of Marketing Management, Faculty of Business, Room F1-44, Erasmus University Rotterdam, PO BOX 1738, 3000 DR Rotterdam, The Netherlands, Phone +31-10-4082732 / Fax +31-10-4089011, E-mail pverlegh@fbk.eur.nl

Introduction

Marketing researchers and practitioners have a strong interest in quantifying the importances consumers attach to variations in individual attributes. For example, products can be improved meaningfully if marketers focus on attributes for which consumers desire changes. A problem with attribute importance measures is their sensitivity to experimental contexts, such as the range and the number of attribute levels (Lehmann et al. 1998; Wittink et al. 1989). As we discuss later, importance measures should be sensitive to range manipulations and should not be sensitive to number-of-levels manipulations. By including both context effects in one experiment, we create an opportunity to benchmark the effect of one against the effect of the other. We study these effects on five different measures of importance, including those most commonly used in marketing. Through the systematic study of contextual effects we obtain insight into the perceptual and judgmental processes that underlie stimulus evaluations. This allows us to examine alternative explanations for the number-of-levels effect observed in conjoint analysis (Currim et al. 1981).

Measures Of Attribute Importance

Importance measures can be assessed through compositional or decompositional approaches. In the former, respondents rate the importance they attach to each of a set of attributes. We refer to this measure as RATE. The validity of this approach is limited because respondents may not be aware of how their preferences on product choices depend on individual attributes. Also, self-reports may be affected by social norms. These problems are reduced, however, if respondents assess the importances of specified ranges of variation. (Srinivasan and Park, 1997).

Decompositional approaches are designed to infer importances. Following the lens paradigm, importance can be derived from regressions of overall product evaluations on perceptions of products' attribute levels (e.g. Tybout and Hauser 1981). However, since consumers have different ideal points for individual attributes, it is better to regress product evaluations on affective *evaluations* (and not perceptions) of levels. Slope coefficients from linear regressions can then serve as measures of attribute importance (e.g. Moskowitz and Krieger 1993). This measure is called REG.

This approach is commonly used to investigate evaluations of products that are available on the market. A drawback is that it may suffer from reverse causality and multicollinearity. Conjoint analysis was developed in part to avoid these problems. In full-profile conjoint, respondents indicate their preferences for (hypothetical) products, where each product consists of specified positions on multiple attributes. Stimulus sets are designed by an experimenter who controls the (co)variations of attribute levels. Green and Wind (1975) proposed that attribute importance equals the range in the utility estimates for an attribute, i.e. the difference in part worths for the best and worst levels of an attribute. We refer to this widely used measure (Cattin and Wittink 1982) as CON.

The remaining two measures are conceptually similar to CON. We use the attractiveness ratings for individual attribute levels (also used for the REG measure) to calculate the difference in attractiveness between the best and worst levels for each attribute. We refer to this importance measure as MAX. The final measure of importance is obtained from a paired comparison task, similar to its use in ACA (Johnson, 1987). We ask respondents to compare the attractiveness of two profiles, x and y . By varying x and y on just one attribute, for which x and y have the best c.q. worst levels, these attractiveness ratings

provide an importance measure we refer to as PAIR.

We provide a summary of the five importance measures in Table 1. In this table we show that OLS is used for just two measures (CON, REG), and we define whether the experimental manipulation is explicit (see next section).

TABLE 1.
CHARACTERISTICS OF IMPORTANCE MEASURES AND CONTEXTUAL MANIPULATIONS

Description of measure	Abbreviation	Input data	OLS	Explicit manipulation	
				Range	# Levels
Full-profile conjoint	CON	Full profile ratings Dummy variables	Yes	Yes	Yes
Attractiveness range	MAX	Level attractiveness ratings	No	Yes	Yes
Regression weight	REG	Full profile ratings Level attractiveness ratings	Yes	No*	No*
Paired comparison	PAIR	Graded paired comparisons	No	Yes	No
Self-report	RATE	Self-reports	No	No	No

* Since both types of input data are subject to the contextual manipulation, these effects may compensate for each other

Context Effects: Range And Number Of Levels

range

The effect of stimulus range on response behavior is well studied in psychophysics. Decreasing the physical range of attribute levels in an experiment leads to a decrease in the range of average stimulus responses because the extreme stimuli have been removed from the stimulus set. However, the range of average responses to the remaining stimuli increases. Furthermore, an identical decrease in stimulus intensity has a larger effect on responses for a narrow stimulus range than for a wide range (Parducci 1974). To illustrate, suppose a consumer judges the expensiveness of TVs. Based on experience, the consumer expects prices to vary from Dfl. 500 to 1700. To judge the expensiveness of TVs, the consumer represents prices on an internal perceived expensiveness continuum. In the simplest case, the consumer uses direct linear transformations to represent prices internally. However, if the consumer is shown only TVs priced from Dfl. 1000 to 1200, when her internal continuum runs from Dfl. 500 to 1700, then the stimuli shown initially span a small area in the middle of the continuum, taking up 1/6 of its total length. In this case she tends to restructure her internal continuum so that the two end points will be close to Dfl. 1000

and 1200, respectively, taking up almost its entire length.

Now consider what would have happened if the TV prices ranged from Dfl. 700 to 1500. Relative to her experience, the internal representations of these TVs would initially occupy 3/4 of the expensiveness continuum. Importantly, now the restructuring of her internal continuum leads only to minor shifts outward on all representations. Thus, when the internal representations are mapped onto a response continuum, they will shift outward in both cases, but the shift will be much larger for a narrow range (1000-1200) than for a wide range (700-1500), relative to expectations (500-1700). The consumer ideally uses direct linear transformations, but the process is biased by a tendency to distribute internal representations uniformly over a continuum (Parducci 1974). Thus, when the internal representations are mapped onto a response continuum, they shift in the direction of the response tendency, so that overt responses are a compromise between a 'real' value, and a 'context' value based on a uniform distribution (Schifferstein 1995).

The question remains whether differences in stimulus range have the same effects on all attribute importance measures. This should be true for CON, MAX, and PAIR, since these measures are all based on evaluations of the difference between best and worst attribute levels. In that case, an increase in attribute range should translate directly into an increase in measured importance (Von Nitzsch and Weber 1993).

We postulate that a given physical difference has a larger effect on attribute level responses in a narrow context than in a wide context. Consequently, if the evaluations of attribute levels are used as the explanatory variables in a regression analysis, the slope coefficient should be smaller in a narrow context than in a wide context. This holds even if the effect of the physical parameter on the criterion variable (overall product liking ratings) is the same in the two contexts. Therefore, REG is expected to be larger with a wider range of attribute levels.

The effect of the range on the RATE measure is difficult to anticipate. Self-reports may reflect some absolute notion of importance, unaffected by experimental context. However, we believe it is more likely that respondents understand that an attribute has a larger impact on product attractiveness if products differ more strongly on that attribute. This produces an increase in RATE, as long as respondents are aware of the ranges of variation in the attributes.

H1: A larger range of attribute variation creates a higher importance for all five measures.

number of levels

The 'number-of-levels' effect is well documented for conjoint analysis. Studies investigating this effect typically find that an increase in the number of levels, holding the range of variation constant, leads to higher attribute importance. This phenomenon is observed for preference ratings and ranks (Wittink et al. 1989), estimated with metric and nonmetric methods (Wittink et al. 1982), as well as for magnitude estimation data (Steenkamp and Wittink, 1994).

We consider two alternative explanations. The first one is the *attention hypothesis*. Green and Srinivasan (1990) suggest that increasing the number of (intermediate) levels may increase the attention given to the attribute, which increases its subjective importance. In our experiment, we facilitate the occurrence of an attention-based effect by showing respondents an overview of attributes and levels twice in the questionnaire. If the number of

levels influences the attention devoted to an attribute, it should affect all measures derived from overall product evaluations, regardless of whether they are obtained by regression (CON, REG), or paired comparisons (PAIR). Greater attention for an attribute should also affect the stated importance measure (RATE). For MAX, which is based on ratings of the best and worst levels, we expect no impact.

H2a : Under the *attention hypothesis*, a higher number of attribute levels leads to higher values for CON, REG, PAIR, and RATE, while MAX is not affected.

The *uniform distribution hypothesis* offers an alternative explanation. Wittink et al. (1989) found similar magnitudes of number-of-levels effects on rank order data and responses on a ten-point rating scale. This equivalence suggests that rankings and ratings have similar (ordinal) measurement properties, and it is consistent with the idea that respondents tend to distribute ratings of unidimensional objects uniformly over a response scale (Parducci 1974). Research suggests that such response shifts are probably due to respondents' tendency to distribute stimulus representations uniformly over a restricted internal continuum (Schifferstein 1995).

To illustrate how this tendency can account for the number-of-levels effect, consider the following. If an attribute (say, price) is presented at two levels, a consumer who uniformly distributes internal representations of price on a mental continuum, and subsequently rates these representations on a scale from 1 to 10, would rate the prices as 4 and 7. If two intermediate levels are added, however, a uniform distribution produces responses of 2.8, 4.6, 6.4, and 8.2. Thus, if stimulus responses were determined by experimental context alone, the range of responses would increase from 3 for two levels to 5.4 for four levels. Because responses depend on stimulus properties as well, actual increases in response ranges will be smaller but nevertheless substantial.

Under this hypothesis, the number of attribute levels will affect CON, consistent with extant results. A larger number of levels should also affect the attractiveness ratings for individual attribute levels, because these ratings are also subject to response shifts. Consequently, MAX should also be affected. Self-reports and paired comparisons are made without reference to the number of levels, so PAIR and RATE should not be affected. And, if judgments of attribute levels and profiles are equally affected, REG will be free of distortion.

H2b: Under the *uniform distribution hypothesis* increasing the number of attribute levels leads to increases in CON and MAX, while REG, PAIR and RATE are unaffected.

Methods And Materials

Stimuli were descriptions of color TVs on six attributes. The selection of attributes and levels was based on interviews with students and retailers. We manipulated the range and the number of levels *between subjects*, in three conditions. In each condition the same six attributes were used, and the total number of levels across attributes was the same, but for three attributes the ranges and/or the numbers of levels varied across conditions (Table 2). The *range effect* is analyzed by comparing conditions A and C, which have different ranges for 'Price' and 'Screen size', while holding the numbers of levels constant. The range for

‘Price’ is three times larger in A than in C, while for ‘Screen size’ it is three times larger in C than in A. We examine the *number-of-levels effect* by comparing conditions A and B. In A two intermediate levels exist for ‘Warranty’, whereas in B two intermediate levels exist for ‘Price’. Note that the attribute ranges do not differ between conditions A and B.

TABLE 2.
COLOR TV ATTRIBUTES AND LEVELS USED IN THE THREE CONDITIONS

Attribute	Condition		
	A	B	C
Warranty (years)	0.5	0.5	0.5
	1		1
	3		3
	5	5	5
Price (Dfl)	650	650	
		800	800
		950	950
	1100	1100	
Screen size (cm)			43
	52	52	
	61	61	
			70
Teletext	no	no	no
	yes	yes	yes
Picture quality	average	average	average
	good	good	good
	very good	very good	very good
Country of origin	Netherlands	Netherlands	Netherlands
	Japan	Japan	Japan

procedure

Respondents were 192 undergraduate students (72 male), varying in age from 17 to 28. We invited students to participate in a survey on TV purchases for which they would receive a financial reward. Participants followed the experimenter to a separate room, where an introduction was provided and questionnaires were handed out. The three different questionnaires were evenly distributed across sessions. Respondents took about 20 minutes to complete the task.

The questionnaire contained four parts, starting with a full-profile conjoint task. On the first page an overview was given of the six attributes and their levels, corresponding to one of the experimental conditions detailed in Table 2. We used SPSS Orthoplan to create sixteen orthogonal profiles based on a fractional factorial main-effects plan. Respondents were told that the TVs were equal on attributes not specified, and that they had the financial means to purchase a new TV. Respondents were asked to rate each of the profiles (shown on separate pages) on a 9-point scale varying from ‘not attractive at all’ to ‘extremely attractive’. To avoid order effects, we used eight different random orders of profiles within each condition (no significant differences were found).

In the second task respondents judged the attractiveness of each separate attribute level (e.g. ‘How attractive to you is a TV made in Japan?’) on the same 9-point scale. The third

task was a graded paired comparison. Respondents were presented with two profiles that differed on one attribute only. In every pair, this difference pertained to the two extreme levels of the attribute. Responses were obtained on a 150 mm line scale ranging from ‘A much more attractive than B’ to ‘B much more attractive than A’, with ‘equally attractive’ as the midpoint. To distract respondents from the fact that the pertinent profiles differed on only one attribute, we added three filler pairs that differed on multiple attributes. Responses to filler pairs were not analyzed. This task produced PAIR.

Before proceeding to task four, respondents were again presented with an overview of all attributes and levels pertaining to their experimental condition. Subsequently, they rated each attribute’s importance in terms of the attractiveness of TVs, on a 100-mm line scale ranging from ‘not important at all’ to ‘extremely important’, resulting in RATE. On the last page, respondents were invited to write any comments. Upon completion of the task, each respondent was asked for comments on the questionnaire. Several respondents mentioned practical aspects, but none of these pertained to the validity of our measures nor to the experimental design.

calculating importance measures

The five different measures of attribute importance for each respondent were obtained as follows. For CON, we used indicator variables to obtain part worth utilities for the attribute levels. For each respondent, the maximum difference in part worths for an attribute is the measure of its importance (CON). The second measure (MAX) was derived from the attractiveness ratings for the attribute levels. For each respondent, we calculated the difference in ratings for the best and worst rated attribute levels.

The third measure, REG, was obtained from individual-level regressions of the overall attractiveness of the 16 profiles (obtained in task one) as a function of the perceived attractiveness of the attribute levels (obtained in task two) pertaining to the profiles. If a respondent provided identical ratings for different levels of an attribute (i.e., zero variance on a predictor), or if the estimated regression coefficient was negative we set the slope coefficient to zero. The fourth measure (PAIR) was obtained from the difference ratings in the graded paired comparison task. RATE captures the self-reported importances.

We focus on how these importance measures are affected by manipulations of the range and the number of levels (see Table 1 for details on whether the measures are subject to explicit contextual manipulations). The range manipulation is explicit if an importance measure is based on evaluations that involve the attribute’s stated levels, which applies to CON, MAX and PAIR. For RATE range effects are caused by respondents’ awareness of the attribute range, without this range being presented directly during the task. The number of attribute levels is manipulated explicitly in tasks where respondents are shown all possible levels of an attribute. This occurs in the conjoint task (CON), and it also applies when respondents are asked to rate the attractiveness of attribute levels (MAX). For the PAIR measure only the extreme levels are used. Thus, there is no explicit manipulation of the number of levels for PAIR nor for RATE. For REG we show in Table 1 that both manipulations are ‘not explicit’, because they pertain to both the criterion and the predictor variables, implying that they may cancel out.

TABLE 3.
GROUP AVERAGES OF THE FIVE IMPORTANCE ESTIMATES AND THEIR
NORMALIZED EQUIVALENTS FOR THE THREE EXPERIMENTAL CONDITIONS.

Attributes & Measures	Normalized measures					Raw measures		
	Effect of manipulations on normalized measures		Condition			Condition		
	Range (A-C)	# Levels (A-B)	A	B	C	A	B	C
CON								
Price	+ .085 **	- .088 **	.17	.26	.08	1.12	1.89	0.57
Screen size	- .069 **		.07	.08	.14	0.49	0.55	0.95
Warranty		+ .083 **	.21	.13	.22	1.41	0.96	1.41
MAX								
Price	+ .068 **	- .046 **	.19	.24	.12	4.34	5.35	3.13
Screen size	- .053 **		.10	.10	.15	2.31	2.42	3.81
Warranty		+ .043 *	.23	.19	.23	5.42	4.45	5.58
REG								
Price	+ .093 **	+ .019	.21	.19	.12	0.27	0.25	0.13
Screen size	- .109 **		.09	.10	.20	0.12	0.15	0.23
Warranty		+ .017	.16	.15	.16	0.17	0.19	0.18
PAIR								
Price	+ .060 *	+ .015	.25	.24	.19	59.6	63.7	53.5
Screen size	- .051 **		.12	.14	.17	30.0	39.6	46.6
Warranty		+ .013	.19	.18	.18	47.3	50.6	49.7
RATE								
Price	+ .020	- .007	.20	.21	.18	74.8	79.6	71.4
Screen size	- .036 **		.14	.15	.17	52.2	58.6	69.1
Warranty		+ .001	.19	.19	.17	71.6	73.1	68.3

** Significant difference between normalized and raw measures in test conditions [$p < 0.01$ (two-tailed)]

* Significant difference between normalized measures in test conditions, not significant for raw measures [$p < 0.01$ (two-tailed)]

Results

We show importances for the manipulated attributes, based on means of raw and normalized measures, in Table 3. Since raw measures are difficult to compare, we normalized the ratings within respondents and within measures by dividing each importance estimate by the sum of the estimates for all attributes (Tybout and Hauser 1981). We also show the statistical significance of each manipulation on the difference in normalized measures for the relevant attributes for each measure. The normalized values show four significant number-of-levels effects and nine significant range effects. The only nonsignificant range effect

occurs for the 'Price' attribute in RATE. This result is almost entirely consistent with H1. On the other hand, a number-of-levels effect occurs, in the expected direction, for CON and MAX only. This is perfectly consistent with H2b and inconsistent with H2a, clearly favoring the uniform-distribution hypothesis over the attention-based explanation.

It is interesting to note that the number-of-levels effect on CON is about equally large as the range effect. The conjoint task thus produces an artificial effect that has roughly the same magnitude as a *threefold* increase in the range of variation. By benchmarking the number-of-levels effect against the range effect, we determine its enormous magnitude in a traditional conjoint task. However, the magnitude of this effect, while still statistically significant, is much smaller on MAX. Importantly, for these two measures the number of levels was explicitly manipulated (Table 1).

Discussion

CON, MAX and PAIR are explicitly based on the range, so that the occurrence of *range* effects for these measures has strong face validity. The range effect in REG could be due to a larger effect of the manipulated attribute in the full-profile ratings than in the attribute-level ratings, to a smaller slope of the unobservable function that relates attribute level attractiveness to its physical counterparts, or to both. It is noteworthy that self-reported importance (RATE) is also affected by the range of attribute levels. RATE is only implicitly subject to the manipulation. Yet the range effect is always in the predicted direction, once significantly. The results on the range effect can be related to results obtained by Mellers and Cooke (1994). Both for single-attribute and multi-attribute judgments, they found that the effect of a given difference in attribute levels on a stimulus profile's rated attractiveness was larger in a narrow range than in a wide range. Mellers and Cooke (1994) argue that their results are consistent with the hypothesis that a range effect is due to changes in the internal representations of the attribute levels. In our study, the attribute level attractiveness ratings suggest that stimulus range affects internal representations. However, we cannot perform a similar analysis on the full-profile evaluations, because we used a fractional instead of a full factorial design. Given that we also observe effects on self-reported attribute importances, we propose that stimulus range affects measures based on differences in internal representations of attribute levels, but also self-report measures of attribute importance. Apparently, consumers judging the importance of an attribute take into account the range of levels for this attribute, so that an increase in range leads to an increase in self-rated importance. This is consistent with a decision-making point of view, which supposes that an attribute with a wider range will produce more variation in product attractiveness.

Large *number-of-levels* effects were obtained for MAX and especially for CON. For the other measures, the effects are not statistically significant and the direction is often contrary to expectation. Attribute-level attractiveness ratings show that the difference in average responses is greater with more levels (MAX, table 3). The fact that the same effect occurs for the utility estimates from the conjoint task suggests that this mechanism has a cognitive counterpart: with more levels, the range of positions on the corresponding internal continuum is expanded. The absence of significant effects for the number-of-levels manipulation on the other three measures indicates that an attention-based explanation is not applicable. Noteworthy is the pattern of results for PAIR, based on a paired comparison task in which respondents are presented with the best and worst levels for an attribute. The explicit consideration of the extreme values of an attribute explains this measure's

sensitivity to changes in the range of attribute levels. And since respondents in this task are not presented with intermediate attribute levels they do not represent those levels internally. This explains the absence of a number-of-levels effect on PAIR. The absence of an effect on RATE indicates that respondents do not attend to the number of levels even though they do attend to the range of variation.

We find that number-of-levels effects occur because internal representations and responses tend to be distributed uniformly over restricted continua. The question arises if and how conjoint analysis designs can be adjusted to avoid this artificial effect. One possibility is to adapt the design in such a way that respondents' tendencies toward uniform distributions will not distort preference judgments. For example, a conjoint exercise might be constructed so that the profiles' *predicted* utilities (and yielded responses) are distributed uniformly. Since respondents exhibit a high degree of preference heterogeneity, this requires individual-level customization of designs. The difficulty is that one would have to know precisely the information the conjoint task is designed to generate.

Alternatively, one could use the same number of levels across the attributes. Equalizing the number of levels across attributes (e.g. Hair et al. 1995), however, obtains uniform distributions of internal representations of attribute *levels*, but does not consider *response* distributions. Importantly, self-explicated measures do not suffer from the problem, and Srinivasan and Park (1997) find that self-explicated data outperform traditional and hybrid conjoint approaches. We recommend therefore Srinivasan and Wyner's (1989) CASEMAP, a computer-assisted self-explicated method.

The experimental design of our study does not accommodate order effects associated with the tasks. Although we do not expect order effects to confound the primary conclusion, it is useful to have this verified. In addition, it would have been of interest to see how alternative importance measures compare in terms of predictive validity. To do this meaningfully, the validation task should be sensitive to range and number-of-levels effects and provide externally valid results.

Conclusion

Our study offers the first direct comparison of the effects of manipulations of range and number-of-levels. For conjoint analysis we show in Table 3 that inserting two intermediate levels can have the same effect as increasing an attribute's range threefold! Thus, an artificial manipulation in the number of levels that should have no consequence is as effectual as a substantively meaningful change in the range. We demonstrate the dramatic impact of stimulus context on product judgments as follows. With two price levels, Dfl. 800 and 950, the relative importance of Price is 8 percent. Using Dfl. 650 and 1100 instead, this importance becomes 17 percent, an increase of 9 percentage points attributable to a threefold increase in the range. However, adding the two intermediate levels creates another increase of 9 percentage points: the relative importance of Dfl. 650 versus 1100 is 17 percent without but 26 percent with Dfl. 800 and 950 as intermediate levels.

Our results suggest that the number-of-levels effect is tied to respondents' tendencies to distribute responses and internal representations uniformly over the corresponding continua. It is not obvious how this tendency can be accommodated or how it can be modified so that artificial effects are minimized. However, self-explicated methods do not suffer from the number-of-levels effect (because respondents judge the importance of an attribute based on the difference between the best and worst levels). Self-explicated methods may outperform

traditional approaches (Srinivasan and Park 1997), perhaps partly because these methods do not suffer from this artificial effect.

Each of the five importance measures responds to variations in the *range* of attribute levels. This is a desirable property, because the attribute range should affect importance. Sensitivity to the *number* of levels is undesirable. It applies to just two measures, and we find that it is caused by a bias toward evenly distributed internal representations and responses, which threatens the external validity of these measures.

REFERENCES

- Cattin, P. and Wittink, D.R. (1982). Commercial use of conjoint analysis: A survey. *Journal of Marketing*, 46 (Summer), 44-53.
- Currim, I.S., Weinberg C.B., and Wittink, D.R. (1981). The design of subscription programs for a performing arts series. *Journal of Consumer Research*, 8, 67-75.
- Green, P.E., and Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, 54 (October), 3-19.
- Green, P.E., and Wind, Y. (1975). New way to measure consumers' judgments. *Harvard Business Review*, 53 (July), 107-117.
- Johnson, R.M. (1987). Adaptive conjoint analysis. *Sawtooth Conference Proceedings*. Ketchum, ID: Sawtooth Software, 253-265.
- Lehmann, D.R., Gupta, S., and Steckel, J. (1998), *Marketing Research*. Reading: AWL.
- Mellers, B., and Cooke, A.D.J. (1994) Trade-offs depend on attribute range. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1055-1067.
- Moskowitz, H., and Krieger, B. (1993). What sensory characteristics drive product quality? An assessment of individual differences. *Journal of Sensory Studies*, 8, 271-282.
- Parducci, A. (1974). Contextual effects: A range-frequency analysis. In E.C. Carterette and M.P. Friedman (Eds.), *Handbook of Perception*, Vol. 2. New York: Academic Press.
- Schiffstein, H.N.J. (1995). Contextual effects in difference judgments. *Perception and Psychophysics*, 57, 56-70.
- Srinivasan, V. and Park, C.S. (1997). Surprising robustness of the self-explicated approach to customer preference structure measurement. *Journal of Marketing Research*, 34, 286-291.
- Srinivasan, V. and Wyner, G.A. (1989). CASEMAP: Computer-assisted self-explication of multiattributed preferences, in W. Henry et al. *New Product Development and Testing*, MA: Lexington Books, 91-111.
- Steenkamp, J.-B.E.M., and Wittink, D.R. (1994). The metric quality of full-profile judgments and the number-of-attribute-levels effect in conjoint analysis. *International Journal of Research in Marketing*, 11, 275-286.
- Tybout, A.M. and Hauser, J.R. (1981). A marketing audit using a conceptual model of consumer behavior: Application and evaluation. *Journal of Marketing*, 45 (2), 82-101.
- Von Nitzsch, R. and Weber, M. (1993). The effect of attribute ranges on weights in multiattribute utility measurements. *Management Science*, 39(8), 937-943.
- Wittink, D.R., Krishnamurthi, L. and Nutter, J.B. (1982). Comparing derived importance weights across attributes. *Journal of Consumer Research*, 8, 471-474.
- Wittink, D.R., Krishnamurthi, L., and Reibstein, D.J. (1989). The effect of differences in the number of attribute levels on conjoint results. *Marketing Letters*, 1, 113-123.