Available online at www.sciencedirect.com
ScienceDirect
Journal of Business Research 60 (2007) 1177 – 1190
JOURNAL OF BUSINESS RESEARCH

The validity of attribute-importance measurement: A review

Koert Van Ittersum [a,*], Joost M.E. Pennings [b,c,d], Brian Wansink [e], Hans C.M. van Trijp [d,f]

[a] *Georgia Institute of Technology, College of Management, 800 West Peachtree Street, NW, Atlanta, GA 30332-0520, United States*
[b] *Maastricht University, The Netherlands*
[c] *University of Illinois at Urbana-Champaign, United States*
[d] *Wageningen University, The Netherlands*
[e] *Cornell University, United States*
[d] *Wageningen University, The Netherlands*
[f] *Unilever Food & Health Research Institute, The Netherlands*

Received 30 May 2006; accepted 7 April 2007

Abstract

A critical review of the literature demonstrates a lack of validity among the ten most common methods for measuring the importance of attributes in behavioral sciences. The authors argue that one of the key determinants of this lack of validity is the multi-dimensionality of attribute importance. Building on the notable work of Myers and Alpert (1968) [Myers JH, Alpert MI. Determinant buying attitudes: Meaning and measurement. J Mark 1968;32(July):13–20], they propose that different methods measure different dimensions of attribute importance and, more specifically, what methods measure which specific dimensions. A re-examination of existing research reveals convergent and nomological validity among methods that are proposed to measure the same dimensions of attribute importance and discriminant validity between methods that are proposed to measure different dimensions of attribute importance. Acknowledging the multi-dimensionality of attribute importance substantially reduces the apparent lack of validity reported in the literature and forms an important first step enabling practitioners and scholars to improve the validity of attribute-importance measurement.

Keywords: Attribute importance; Measurement; Convergent validity; Nomological validity; Discriminant validity

Contents

* Corresponding author. Tel.: +1 404 385 4884; fax: +1 404 894 6030.
E-mail address: koert.vanittersum@mgt.gatech.edu (K. Van Ittersum).

0148-2963/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbusres.2007.04.001

1. Introduction

Identifying product attributes of importance is a key objective of marketing research. Although a wide variety of methods exists to measure the importance of attributes, the convergent validity among and nomological validity of different methods is often low (Jaccard et al., 1986). Convergent validity identifies whether different measurements reflect the same construct (i.e., are positively correlated). Nomological validity examines whether measures are related to other constructs in a theoretically meaningful way (i.e., predictive accuracy). Low levels of validity can cause serious empirical and practical problems (Sethuraman et al., 2005).

Based on a critical review of the literature, we demonstrate a lack of convergent and nomological validity exists among the ten most common methods for measuring the importance of attributes in behavioral sciences. We believe that one of the key determinants of the lack of validity is the multi-dimensionality of attribute importance. Research on attribute-importance measurement often defines attribute importance tautologically as the importance of an attribute to an individual. Consequently, it is assumed that attribute importance is a unidimensional concept that can be measured with different methods. However, what is important based on personal values and desires (e.g., safety of an airline) may be unimportant in judgment and choice (e.g., choosing which airline to travel with as all airlines are equally safe). Therefore, in line with Myers and Alpert (1968), we propose that there is no single definition of attribute importance. Instead, attribute importance is a *multidimensional* concept — i.e., different dimensions of attribute importance exist. We further propose that different methods for measuring attribute importance measure different dimensions of attribute importance, which explains the lack of validity of attribute-importance methods.

To examine our central proposition that different methods measure different dimensions of attribute importance and as such may explain the apparent lack of validity among methods for measuring the importance of attributes, we propose a *multidimensional* research framework of attribute importance that details the relationships among three dimensions of attribute importance and specifies their critical antecedents. Based on the framework, we propose which specific dimension of attribute importance is being measured by ten common attribute-importance methods. Finally, the propositions are examined through a critical and integrative review of seemingly divergent findings in the literature. Acknowledging the strengths and weaknesses of both meta-analyses and reviews (Brinberg and Jaccard, 1986), we conducted a review, as it enabled us to examine multiple methods and account for the influence of factors that may have been ignored in a more traditional meta-analysis. While we cannot statistically test our method-specific propositions, statistical evidence for our central proposition will be provided.

2. Understanding the lack of validity

It is widely recognized that the validity of available methods for measuring attribute importance remains limited. Our review of the literature (see Section 4 for a description of the methodology) confirms the reported lack of convergent validity (see Table 1). For example, Jaccard et al. (1986) find low correlations ($r < .35$; Cohen and Cohen, 1983) between the importance of attributes measured using the free-elicitation method and the importance of attributes measured using the direct-rating method. Table 1 shows many other method comparisons that also lack convergent validity.

In addition, our review reveals that the predictive accuracy of importance measures, which are elicited with different methods, varies. This suggests a lack of nomological validity of methods for measuring attribute importance. For instance, Harte and Koele (1995) test the nomological validity of the direct-rating method and the multi-attribute attitude method by using the attribute-importance measures from both methods to predict choice. They find a significant difference in the percentage correctly classified choices between both methods ($p_a = 76.5\% < p_b = 85.3\%$) (see Table 1). While Harte and Koele (1995) relate the importance measures to choice, others examine the nomological validity by relating attribute-importance measures to preference scores (Schoemaker and Waid, 1982) and perceptions (Neslin, 1981).

Many factors that cause this lack of convergent and nomological validity have been identified: familiarity with object (Park and Lessig, 1981), the number of attributes studied (Adelman et al., 1984), the number of attribute levels studied (Wittink et al., 1989), and for instance the range of attribute levels studied (Verlegh et al., 2002). Most of these factors merely explain the lack of validity among specific methods, without providing generalizable insights.

Table 1
Studies that show a lack of validity of methods for measuring attribute importance

	Major finding	Statistical basis	Source
Convergent validity	Free-elicitation method ≠ direct-rating method	$r_{a,b}$ = .19	Jaccard et al. (1986)
		$r_{a,b}$ = .33	
	Free-elicitation method ≠ information-display board	$r_{a,b}$ = .22	Jaccard et al. (1986)
		$r_{a,b}$ = .19	
	Free-elicitation method ≠ rank-based conjoint method	$r_{a,b}$ = .09	Jaccard et al. (1986)
		$r_{a,b}$ = .11	
	Direct-rating method ≠ multi-attribute attitude method	$r_{a,b}$ = ns for 7/10 tests	Wiley et al. (1977)
	Direct-rating method ≠ trade-off method	$r_{a,b}$ = .17	Fischer (1995)
		$r_{a,b}$ = .31	Hobbs (1980)
	Direct-rating method ≠ rating-based conjoint method	$r_{a,b}$ = inconsistent	Zhu and Anderson (1991)
	Direct-rating method ≠ rank-based conjoint method	$r_{a,b}$ = .10	Jaccard et al. (1986)
		$r_{a,b}$ = .11	
	Direct-ranking method ≠ multi-attribute attitude method	$r_{a,b}$ = .27	Barlas (2003)
	Direct-ranking method ≠ multi-attribute attitude method	$a_1/a_2 = 1.99 \neq b_1/b_2 = 1.28$	Van der Pligt et al. (submitted for publication)
	Direct-ranking method ≠ rating-based conjoint method	$r_{a,b}$ = inconsistent	Norman (1977)
	Point-allocation method ≠ multi-attribute attitude method	$a_1/a_2 = 1.94 \neq b_1/b_2 = 1.28$	Van der Pligt et al. (submitted for publication)
		$a_1/a_2 = 1.81 \neq b_1/b_2 = .99$	
		$a_1/a_2 = 1.74 \neq b_1/b_2 = 1.04$	
		$a_1/a_2 = 1.69 \neq b_1/b_2 = 1.09$	
		$a_1/a_2 = 1.55 \neq b_1/b_2 = 1.01$	
	Point-allocation method ≠ rating-based conjoint method	$r_{a,b}$ = inconsistent	Zhu and Anderson (1991)
	Analytical-hierarchy process ≠ trade-off method	$a_1/a_2 = 3.08 \neq b_1/b_2 = 1.44$*	Pöyhönen and Hämäläinen (2001)
	Analytical hierarchy process ≠ swing-weight method	$a_1/a_2 = 3.08 \neq b_1/b_2 = 1.37$*	
	Information-display board ≠ multi-attribute attitude method	$r_{a,b}$ = .41	Barlas (2003)
	Information-display board ≠ rank-based conjoint method	$r_{a,b}$ = .32	Heeler et al. (1979)
		$r_{a,b}$ = .11	Jaccard et al. (1986)
		$r_{a,b}$ = .09	
Nomological validity	Direct-rating method < multi-attribute attitude method	$p_a = 76.5\% < p_b = 85.3\%$*	Harte and Koele (1995)
	Direct-rating method < rating-based conjoint method	$r_{a,c} = .49 < r_{b,c} = .64$*	Neslin (1981)
	Point-allocation method > multi-attribute attitude method	$r_{a,c} = .78 > r_{b,c} = .70$*	Schoemaker and Waid (1982)
	Analytical hierarchy process < trade-off method	$r_{a,c} = .74 < r_{b,c} = .77$*	
	Information-display board ≠ multi-attribute attitude method	$r_{a,c} \neq r_{b,c}$*	Billings and Marcus (1983)

*$p < .05$.

Building on the notable work of Myers and Alpert (1968, 1977), we propose that the multi-dimensionality of attribute importance is a critical determinant of the lack of validity. Because attribute importance tends to be defined as a unidimensional concept (as "the importance of attribute to an individual"), it is generally assumed that it can be measured with diverse methods. Myers and Alpert (1968, 1977), however, suggest that attribute importance is not a unidimensional concept. After all, what is important based on personal values and desires may not be important in judgment and choice. Instead of considering attribute importance as a unidimensional concept, Myers and Alpert suggest that it is a multidimensional concept and that different methods may measure different dimensions of attribute importance. Building on this notion, we propose a multidimensional research framework that details the relationships among the three dimensions of attribute importance and specifies their critical antecedents.

2.1. Multi-dimensionality of attribute importance

Myers and Alpert (1968, 1977) distinguish between three dimensions of attribute importance: salience, relevance, and determinance. These three dimensions are the key components of our research framework. By explicitly specifying the antecedents of these three dimensions, as well as their relationships, the multidimensional approach to attribute importance enables us to gain a better understanding of the causes of the lack of convergent validity among and nomological validity of common methods for measuring attribute importance. Fig. 1 shows the three dimensions of attribute importance, their relationships, and their antecedents.

Salience refers to the fact that "not all of a [person]'s beliefs stand out with equal prominence in his cognitive field" (Krech and Crutchfield, 1948; p. 163), and reflects the degree of ease with which attributes come to mind or are recognized when thinking about or seeing a certain object. Salient attributes are considered more important than non-salient attributes (Steenkamp and Van Trijp, 1997; Wansink et al., 2005). The salience of attributes is largely determined by the accessibility of attribute information in people's memory (i.e., *internal attribute information*, see Fig. 1) (Alba et al., 1991).

The *relevance* of attributes is largely determined by *personal values and desires* (Batra et al., 2001) and reflects the importance of attributes for individuals (Myers and Alpert, 1977) (see Fig. 1). Attributes that provide benefits that satisfy important values and desires are considered more relevant, hence more important, than attributes that do not provide such benefits (Schwer and Daneshvary, 2000). Variations in personal

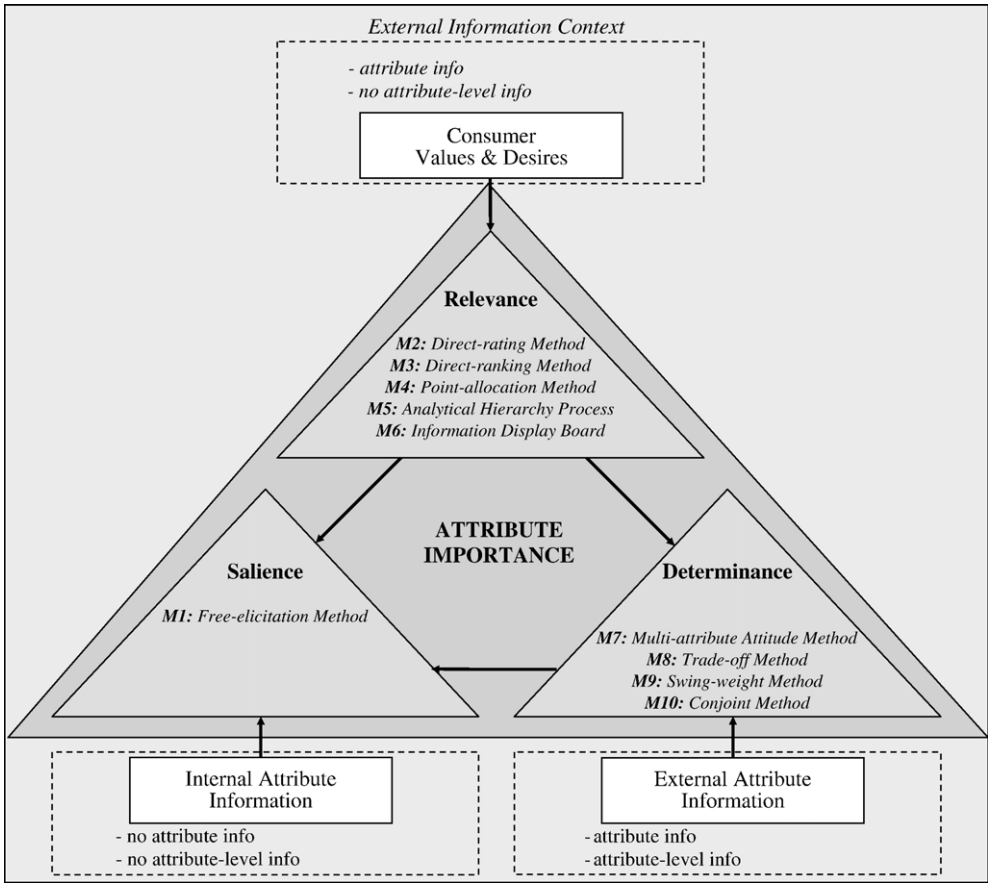

Fig. 1. The three dimensions of attribute importance, their antecedents, and examples of methods that identify attributes that are salient, relevant, or determinant. Notes: salience reflects the degree of ease with which attributes come to mind or are recognized when thinking about or seeing a certain object. The relevance of attributes is largely determined by personal values and desires and reflects the importance of attributes for individuals. The determinance of an attribute reflects the importance of an attribute in judgment and choice. The salience of attributes depends on the accessibility of attribute information in people's memory and increases with the quantity and quality of processing that information (Alba et al., 1991). As determinant and relevant attributes receive more attention than non-determinant attributes, determinant and relevant attributes will be more salient than non-determinant attributes. While relevant attributes tend to be more salient, the opposite is not expected to be true. People may use salience as a heuristic for inferring attribute relevance (Fishbein and Ajzen, 1975). The determinance of attributes is positively affected by the relevance of the attribute to the individual (Alpert, 1971). Note that an attribute may be relevant to an individual, yet not be determinant in the decision process. For instance, the determinance of the miles-per-gallon attribute is zero if an individual has to choose between two cars that have the same fuel efficiency. Irrelevant attributes may also be determinant in judgment (Carpenter et al., 1994). As the determinance of attributes is driven by the difference in (valuation of) attribute levels, the effect of the salience of the attribute on its determinance will be limited.

values and desires explain individual differences in attribute relevance.

The *determinance* of an attribute reflects the importance of an attribute in judgment and choice (Myers and Alpert, 1977) and is generally calculated based on the difference in (valuation of) different attribute levels (e.g., the difference in part-worth utilities in conjoint methods). The determinance of attributes thus depends on the difference in attribute levels in the object space examined (i.e., *external attribute information*, see Fig. 1). The larger the differences in attribute levels, the more determinant an attribute becomes (Fischer, 1995). Determinant attributes are considered more important than non-determinant attributes.

We conclude that attribute importance is a multidimensional concept consisting of three dimensions of attribute importance: the salience, relevance, and determinance of attributes. The *salience* of an attribute represents the *importance* of the attribute in memory. The *relevance* of an attribute represents the *importance* of the attribute to the individual based on personal values and desires. Finally, the *determinance* of an attribute represents the *importance* of the attribute in judgment and choice.

3. An explanation for the lack of validity: the multi-dimensionality of attribute importance

3.1. Different methods measure different dimensions of attribute importance

We proposed a framework that explicitly specifies the antecedents of the three dimensions of attribute importance (see Fig. 1). Understanding these antecedents allows us to propose more specifically what method measures which specific dimension of attribute importance. Understanding what method measures which dimension of attribute importance forms an important first step in improving the validity of attribute-importance measurement. Hence, we propose,

Central proposition. *The lack of convergent validity among and nomological validity of available methods for measuring the importance of attributes is a result of different methods measuring different dimensions of attribute importance.*

The external attribute-information availability will be an important factor determining what method measures which specific dimension of attribute importance. More generally, we propose that if a method is used without providing any specific attribute information, the method will measure the *salience* of attributes, as people have to retrieve the important attributes from memory. If a method is used that provides attributes *and* attribute-level information, the method is proposed to measure the *determinance* of attributes (Alpert, 1971; Fischer, 1995; see Fig. 1, left bottom corner). The rationale for this proposition is that the attribute-level information frames the participants when responding to the attribute-importance questions (Verlegh et al., 2002; Wittink et al., 1989). Finally, if a method is used that only provides attribute information (but no attribute-level information; see Fig. 1, top corner), individuals rely on their personal values and desires when responding to attribute-importance questions (Myers and Alpert, 1968), as a result of which these methods are proposed to measure the *relevance* of attributes.

3.2. What methods measure which attribute-importance dimension?

Using the proposed framework, we elaborate upon the role of the *attribute-information context* in attribute-importance measurement and formulate propositions regarding which attribute-importance dimension each of ten common methods for measuring the importance of attributes in behavioral science measures. Section 4 describes how these ten common methods were selected.

3.2.1. Methods measuring the salience of attributes

The first group of methods used to identify important attributes is that of the attribute-elicitation methods (Steenkamp and Van Trijp, 1997). Here we focus on the free-elicitation technique, as it is the only method measuring the salience of attributes that is compared with other methods for identifying important attributes (Jaccard et al., 1986). The *free-elicitation method* (M_1) uses an open-ended question to let individuals indicate which attributes they believe are important, for instance, when thinking about a new product or brand. The free-elicitation technique assumes that the order of elicitation reflects importance (Kaplan and Fishbein, 1969). The attribute that is elicited first is considered more important than the subsequent attributes elicited. As no attribute information is presented when using this method, it solely relies on people's ability to retrieve internal attribute information stored in memory. Hence,

Proposition 1. *The dimension of salience is measured most directly by the free-elicitation method (M_1).*

3.2.2. Methods measuring the relevance of attributes

A second group of methods discussed here identifies important attributes by directly asking people to judge the importance of attributes or by inferring it based on information search. Consider a situation in which one wants to establish the importance of n attributes. The *direct-rating method* (M_2) has individuals rate the attributes on a rating scale (e.g., 1 = "unimportant"–7 = "important"). The *direct-ranking method* (M_3) asks individuals to rank-order the n attributes. The *point-allocation method* (M_4) has individuals distribute 100 points among the attributes (important attributes receiving more points). The *analytical hierarchy process* (M_5) starts with a multi-attribute problem, which is structured into a hierarchy of products and attributes. The importance of the attributes is determined by asking individuals to compare the importance of two attributes at a time. Based on multiple comparisons, the overall importance of each attribute is derived (Saaty, 1977). Finally, the *information-display-board method* (M_6) measures attribute importance through the extent and order of information search (Ford et al., 1989). Individuals are shown an information-display board with attributes arranged by product, for instance, with the attribute-level information covered. To identify their most preferred product, individuals can gather attribute information by uncovering one cell at a time. They continue to uncover additional cells until they believe they can make a confident decision. Attribute importance is positively related to the extent and order of search. As attribute-level information is available (though not visible to the individual if not requested) when using the *information display board* (M_6), it could be argued that this method may also measure the determinance of attributes. However, as people have to decide if and how much additional attribute-level information to collect, this search for attribute information (indicator of attribute importance) is *primarily* based on personal values and desires.

When using these methods (M_2–M_6), attribute information (e.g., price) is explicitly presented to the participants. Hence, the role of the *salience* of the attributes in memory is less prominent in these methods. Second, generally speaking, none of these methods explicitly identifies the levels of the attributes studied. Hence, people's responses will be primarily driven by the importance of the attributes for them based on their personal values and desires (see Fig. 1, top corner). More formally, we propose that

Proposition 2. *The dimension of relevance is measured most directly by the direct-rating method (M_2), direct-ranking method (M_3), point-allocation method (M_4), analytical-hierarchy process (M_5), and information-display boards (M_6).*

3.2.3. Methods measuring the determinance of attributes

The last group of methods discussed infers the importance of attributes based on evaluative judgments (see Fig. 1, right bottom corner). The *multi-attribute attitude method* (M_7) (Wilkie and Pessemier, 1973) assumes that individuals value each distinct thought they have about an attitude object, and they weigh each of these thoughts according to its likelihood of occurrence. The attitude is assumed to equal the sum of the weighted evaluations. To estimate the importance of attributes, people's overall attitudes are regressed on their valuation of the attribute levels involved. The attribute importance is represented by the weight that an individual attaches to the attribute level in judgment and choice. As we study methods for

identifying important attributes, we focus on decompositional methods that infer attribute importance based on regression weights and not on compositional methods that incorporate importance measures based on for instance the direct-rating method (Fishbein and Ajzen, 1975). In *multi-attribute attitude methods* (M_7), the importance of an attribute is represented by the weight that an individual gives to the attribute level in judgment — the impact of (the valuation of) a specific attribute level on the overall liking of the object (i.e., *determinance*). Generally, these weights are determined through regression analyses (Harte and Koele, 1995; Van Ittersum et al., 2003, 2007). This method requires the presentation of attribute and attribute-level information. Therefore, the role of the *salience* of the attributes involved is limited. While personal values and desires will affect the importance of the attributes, the specific attribute levels presented when using this method will primarily determine the importance of the attribute in judgment and choice. The *trade-off method* (M_8) (Keeney and Raiffa, 1976) has individuals conduct a matching task — for instance, adjust one attribute of one product, such that the product becomes equally attractive to another product that is fully described on all available attributes, from which attribute importance is derived. By having individuals make trade-offs between attribute levels, this method also measures the importance of attributes in judgment. Since attribute and attribute-level information is provided to individuals, the role of the salience and the relevance will be limited, or at least be outweighed by the determinance of the attributes. The same holds for the *swing-weight method* (M_9), which, like the trade-off method, relies on the effect of differences in attribute levels in judgment (Von Winterfeldt and Edwards, 1986). The *swing-weight method* asks individuals to indicate which attribute they would upgrade first if they were confronted with a product that has attributes with only the worst possible levels available. This attribute receives 100 points. Next, individuals are asked to upgrade a second attribute, and indicate how many points this attribute would receive. The *conjoint method* (M_{10}) determines attribute importance based on the difference in valuations (part-worth utilities) between the most and least favorable level of an attribute (Van der Lans et al., 2001; Wansink and Van Ittersum, 2004).

As discussed, within these methods (M_7–M_{10}), attribute importance is defined as the difference in valuations between the most and least favorable level for an attribute (e.g., Green et al., 1972). As such, this method establishes the determinance of attributes in judgment or choice.

Proposition 3. *The dimension of determinance is measured most directly by the multi-attribute attitude method (M_7), trade-off method (M_8), swing-weight method (M_9), and conjoint methods (M_{10}).*

In conclusion, there are three groups of methods for measuring the importance of attributes (see also Fig. 1). The first "group" is proposed to measure the *salience* of attributes (M_1). The second group is expected to measure the *relevance* of attributes (M_2, M_3, M_4, M_5, and M_6). The third group of methods is proposed to measure the *determinance* of attributes (M_7, M_8, M_9, and M_{10}).

4. Review methodology

To identify articles that use two or more methods to measure the importance of attributes and investigate the validity of those methods, we conducted a review of the literature. As attribute-importance measurement is a central topic in most behavioral sciences, our review focused on the major journals in marketing, organizational behavior and management, and psychology. Our search consisted of a combination of an online search of bibliographic databases and manual searches of the selected journal volumes for which no electronic versions were available. Only articles that use two or more methods to measure the importance of the *same* attributes were included in our review. Furthermore, an article must report a comparison of the outcome of the methods used to measure the importance of the same attributes to be included. From the total set of identified articles, we selected the ten most common methods for measuring the importance of attributes based on an objective count of the number of applications ($\geq$ 5 times): free-elicitation method, direct-rating method, direct-ranking method, point-allocation method, analytical hierarchy process, information display board, multi-attribute attitude method, trade-off method, swing-weight method, and conjoint method (see Fig. 1: M_1–M_{10}). This does imply that not all method-comparisons discussed in an article may have been included in our review. Examples of methods not included in our review are the equal weights method (Wainer, 1976), the unit weighting method (Schoemaker and Waid, 1982), and the verbal protocol method (Harte and Koele, 1995). Considering that the ten methods were selected based on a review of the major journals in most behavioral sciences, we consider them to adequately represent the range of available methods.

Our final selection includes 34 articles (see Appendix A), reporting 42 studies and 91 method comparisons. The studies identified examine either the convergent or the nomological validity of the methods investigated. The most common way of examining the convergent validity is to correlate the importance of attributes measured by two methods across individuals (Stillwell et al., 1983). The absolute correlation coefficient can be used to examine the convergent validity. Correlations below .35 are generally considered low, while those above .45 are considered moderate to high (Cohen and Cohen, 1983). We used a correlation of .35 as the cut-off level for concluding whether or not two methods show convergent validity. A second method to examine convergent validity is to determine the ratio of the two most important attributes measured with two methods for measuring the importance of attributes (Van der Pligt et al., submitted for publication). If this ratio is not significantly different, one may conclude that there is convergence. Nomological validity is established by using the importance measures from two methods to predict choice or to relate both measures to perceptions, preferences, or intentions (Schoemaker and Waid, 1982). Evidence for nomological validity can be obtained by comparing these relationships for both methods studied. Likewise, the correlations between attribute-importance measures and a selected dependent variable can be examined.

5. Results

When examining the convergent and nomological validity for the 91 method comparisons, ignoring the multidimensional approach to attribute importance, we would conclude that 67.0% of the method comparisons (61 out of the 91) lack validity. That is, we find low correlations ($r < .35$) between attribute-importance measures determined with different methods. Likewise, we find significant differences in nomological validity — the relationship between the importance of attributes measured with different methods and a third, dependent variable (e.g., perception, preference, intention, and choice). Our central proposition is that differentiating between methods based on which dimension of attribute importance they measure helps explain the apparent lack of validity of different methods for measuring the importance of attributes. To investigate this central proposition, we first investigate the *discriminant validity* between methods that are proposed to measure *different* dimensions of attribute importance.

5.1. Discriminant validity

Evidence for discriminant validity is reported in Table 1. The 32 apparent inconsistencies reported in Table 1 may be attributed to the fact that all comparisons involve methods that measure *different* dimensions of attribute importance. For instance, the lack of convergent validity between the free-elicitation method and the direct-rating method ($r_{a,b} = .19 < .35$), the information-display board ($r_{a,b} = .22 < .35$), and the rank-based conjoint method ($r_{a,b} = .09 < .35$), reported by Jaccard et al. (1986), may all be attributed to the fact that the free-elicitation method measures the salience of attributes (Proposition 1), while the other three methods measure the relevance (Proposition 2) and determinance of attributes (Proposition 3) (see first rows of Table 1). All inconsistencies in Table 1 can be explained based on the central proposition that different methods measure different dimensions of attribute importance. In conclusion, we find evidence for *discriminant validity* between methods that measure *different* dimensions of attribute importance. This finding substantiates the central proposition that the apparent lack of validity of methods for measuring the importance of attributes can be explained by taking a multidimensional approach to measuring attribute importance.

5.2. Convergent and nomological validity

We next examine the *convergent validity* among and *nomological validity* of methods that are proposed to measure the *same* dimension(s) of attribute importance. Table 2 provides an overview of 30 method comparisons that suggests convergent validity among and nomological validity of methods proposed here to measure the same dimensions of attribute importance. For instance, Stillwell et al. (1981) report convergence among the direct-rating and direct-ranking method ($r_{a,b} = .63 > .35$), both proposed to measure the *relevance* of attributes (Proposition 2). Fischer (1995) reports convergent validity among the trade-off and the swing-weight method ($r_{a,b} = .67, > .35$), both of which were proposed here to measure the *determinance* of attributes (Proposition 3).

These results suggest that there is *convergent validity* among and *nomological validity* of methods that measure the *same* dimension(s) of attribute importance.

5.3. Statistical test of discriminant and convergent validity

To provide a more statistical foundation for these findings, we tested the differences in the correlations used to examine convergent validity. The objective was to investigate whether the correlations are significantly larger in attribute-importance measures between methods that measure the same dimension of attribute importance (i.e., convergent validity) than the correlations in attribute importance between methods that measure different dimensions of attribute importance (i.e., discriminant validity). Please note that the differences in reported ratios (a1/a2¹b1/b2) associated with convergent validity as well as the correlations associated with nomological validity ($r_{a,c}$=$r_{b,c}$) cannot be tested statistically.

We followed the basic principles of conducting a meta-analysis using a parametric adjustability model (Farley et al., 1995). The parametric adjustability method starts with the identification of relevant studies (1) and relevant study characteristics within those studies (2). Next, using dummy variable regression, the model is tested (3). In this research, we investigated the effect of a dummy variable on the size of correlation coefficients. The dummy variable is one when a study examines the correlation in attribute importance measured with two methods that are proposed to measure different dimensions of attribute importance. The dummy variable is zero when a study examines the correlation in attribute importance measured with two methods that are proposed to measure the same dimension of attribute importance. In line with our central proposition, we expected to find a significant effect of this dummy variable on the dependent variable, the size of the correlation coefficient *r*. Besides this dummy variable, we included covariates for the number of respondents, the number of attributes, and the type of respondent (student vs. consumer sample), and one to correct for possible study-specific influences. None of the covariates influence the dependent variable. To achieve normality for our correlation coefficients, we used Fisher's z_r transformation. Furthermore, since correlation coefficients are inversely proportional to the number of observations used to calculate them, we estimated the parametric adjustability model through a weighted least-squares procedure with $N - 3$ as weight (Hedges and Olkin, 1985).

In line with our central proposition, we find a significant effect of the main dummy variable — i.e., the size of the correlation of attribute importance measured with two methods is dependent on whether both methods measure the same or different dimensions of attribute importance ($\beta = -.498$, $t = 2.00$, $p < .05$). The difference in correlations between methods that measure the same dimensions of attribute importance and those that measure different dimensions of attribute importance is substantial ($r_{same} = .66$ vs. $r_{different} = .20$).

5.4. Conclusions

We conclude that there is convergent validity among and nomological validity of methods that measure the same dimensions of attribute importance, while there is discriminant validity between methods that measure different attribute-importance dimensions. Distinguishing between the salience, relevance, and determinance of attributes, and acknowledging the antecedents of these dimensions of the importance of an attribute helps reduce the lack of validity from 67.0% to 31.9% of the method comparisons identified. These results suggest that taking a multidimensional approach to attribute importance, and relating these dimensions to different methods, may actually be a first step toward more valid attribute-importance measurement.

6. Other factors influencing the lack of validity

Despite the compelling evidence for the multi-dimensionality of attribute importance reported above, the results in Table 3 suggest that taking a multidimensional approach to attribute importance does not explain all of the lack of validity of methods for measuring the importance of attributes. For instance, both the direct-rating method and the point-allocation methods are proposed to measure the relevance of attributes (Proposition 2). However, among others, Doyle et al. (1997) report a lack of convergent validity among both methods (see Table 3).

6.1. Differences in information processing

Differences in information processes as a result of whether stimuli are presented separately or simultaneously may explain some of the inconsistencies reported in Table 3. Research has shown that there is a difference in abstractness activation between separate- and joint-evaluation tasks (Hsee, 1996). Because more abstract attributes are easier to evaluate, they can be evaluated at any time and in any context. Concrete attributes tend to be more difficult to evaluate, and therefore can only be evaluated in comparison. Consequently, compared to abstract attributes, concrete attributes have a greater impact in a joint-evaluation task (e.g., point-allocation method) than in a separate-evaluation task (e.g., direct-rating method). This may explain the unexpected results presented by Doyle et al. (1997) as their study involves both abstract attributes, such as team membership and motivation, and more concrete attributes, such as written and oral communication capabilities. The same may hold for other studies presented in Table 3. Future research may consider examining this notion more rigorously in relationship with the measurement of the importance of attributes.

Table 2
Studies that show validity of methods for measuring attribute importance

	Major finding	Statistical basis	Source
Measurement of the relevance of attributes			
Convergent validity	Direct-rating method=direct-ranking	$r_{a,b}=.63$	Stillwell et al. (1981)
		$r_{a,b}=.88$	Stillwell et al. (1981)
	Direct-rating method=analytical hierarchy process	$r_{a,b}=.68$	Srivastava et al. (1995)
	Direct-rating method=information-display board	$r_{a,b}=.90$	Harte et al. (1996)
		$r_{a,b}=.58$	Jaccard et al. (1986)
	Direct-ranking method=point-allocation method	$r_{a,b}=.98$	Stillwell et al. (1983)
	Direct-ranking method=point-allocation method	$a_1/a_2=1.99=b_1/b_2=1.94$	Van der Pligt et al. (submitted for publication)
	Direct-ranking method=information-display board	$r_{a,b}=.53$	Barlas (2003)
	Multi-attribute attitude method=swing-weight method	$r_{a,b}=.60$	Srivastava et al. (1995)
	Point-allocation method=information-display board	$r_{a,b}=.53$	Heeler et al. (1979)
Nomological validity	Direct-rating method=point-allocation method	$r_{a,c}=.89=r_{b,c}=.90^{ns}$	Cook and Stewart (1975)
		$r_{a,c}=.85=r_{b,c}=.83^{ns}$	
	Direct-rating method=analytical hierarchy process	$r_{a,c}=.89=r_{b,c}=.91^{ns}$	Cook and Stewart (1975)
		$r_{a,c}=.85=r_{b,c}=.84^{ns}$	
	Point-allocation method=analytical hierarchy process	$r_{a,c}=.78=r_{b,c}=.74^{ns}$	Schoemaker and Waid (1982)
		$r_{a,c}=.90=r_{b,c}=.91^{ns}$	Cook and Stewart (1975)
		$r_{a,c}=.83=r_{b,c}=.84^{ns}$	
Measurement of the determinance of attributes			
Convergent validity	Trade-off method=swing-weight method	$r_{a,b}=.67$	Fischer (1995)
		$r_{a,b}=.37$, $r_{a,b}=.46$	Borcherding et al. (1991)
		$a_1/a_2=1.44=b_1/b_2=1.37^{ns}$	Pöyhönen and Hämäläinen (2001)
	Trade-off method=rating-based conjoint method	$r_{a,b}$=consistent	Zhu and Anderson (1991)
	Multi-attribute attitude method=swing-weight method	$r_{a,b}=.78$	Srivastava et al. (1995)
Nomological validity	Trade-off method=rating-based conjoint method	$p_a=35.5\%=p_b=41.1\%^{ns}$	Leigh et al. (1984)
		$p_a=65.8\%=p_b=64.0\%^{ns}$	Srinivasan and Park (1997)
	Trade-off method=rank-based conjoint method	$r_{a,c}=.97=r_{b,c}=.85^{ns}$	Green et al. (1972)
		$p_a=36.7\%=p_b=26.8\%^{ns}$	Leigh et al. (1984)
	Multi-attribute attitude method=trade-off method	$r_{a,c}=.71=r_{b,c}=.77^{ns}$	Schoemaker and Waid (1982)
	Choice-based conjoint method=rating-based conjoint	$p'_a=.92=p'_b=.92^{ns}$	Elrod et al. (1992)[a]
		$p_a=p_b^{ns}$	Oliphant et al. (1992)

[a] p'=proportion reduction in mean square error of prediction.

Table 3
studies comparing methods that yield an unexpected (lack of) validity

	Major finding	Statistical basis	Source	Possible reasons
Convergent validity	Direct-rating method≠point-allocation method [a]	$a_1/a_2=.64<b_1/b_2=1.72$ *	Doyle et al. (1997)	Differences in information processing
		$a_1/a_2<b_1/b_2$	Schori (1995)	
		$a_1/a_2=1.34<b_1/b_2=2.15$ *	Zhu and Anderson (1991)	
		$a_1/a_2=1.08\neq b_1/b_2=1.81$	Van der Pligt et al. (submitted for publication)	
		$a_1/a_2=1.06\neq b_1/b_2=1.74$		
		$a_1/a_2=1.03\neq b_1/b_2=1.69$		
		$a_1/a_2=1.06\neq b_1/b_2=1.55$		
	Direct-rating method≠information-display board [a]	$r_{a,b}=.25^{\text{car data}}$	Jaccard et al. (1986)	
	Direct-rating method=multi-attribute attitude method [b]	$r_{a,b}=.56$	Srivastava et al. (1995)	
		$a_1/a_2=1.08\neq b_1/b_2=0.99$	Van der Pligt et al. (submitted for publication)	
		$a_1/a_2=1.06\neq b_1/b_2=0.99$		
		$a_1/a_2=1.03\neq b_1/b_2=1.09$		
		$a_1/a_2=1.06\neq b_1/b_2=1.01$		
	Direct-rating method=swing-weight method [b]	$r_{a,b}=.66$	Srivastava et al. (1995)	Effects of framing
	Point-allocation method=trade-off method [b]	$a_1/a_2=1.43=b_1/b_2=1.44^{ns}$	Pöyhönen and Hämäläinen (2001)	
	Point-allocation method=swing-weight method [b]	$a_1/a_2=1.43=b_1/b_2=1.37^{ns}$	Pöyhönen and Hämäläinen (2001)	
	Point-allocation method=rank-based conjoint method [b]	$r_{a,b}=.59$ *	Heeler et al. (1979)	
	Analytical hierarchy process=multi-attribute attitude method [b]	$r_{a,b}=.63$	Srivastava et al. (1995)	
	Analytical hierarchy process=swing-weight method [b]	$r_{a,b}=.84$	Srivastava et al. (1995)	
Nomological validity	Point-allocation method=multi-attribute attitude method [b]	$r_{a,c}=r_{b,c}$	Adelman et al. (1984)	
	Point-allocation method=trade-off method [b]	$r_{a,c}=r_{b,c}$	Adelman et al. (1984)	
	Point-allocation method=trade-off method [b]	$r_{a,c}=.78=r_{b,c}=.77^{ns}$	Schoemaker and Waid (1982)	
	Direct-rating method>direct-ranking method [a]	$r_{a,c}>r_{b,c}$ *	Maio et al. (1996)	Differences in information processing
	Direct-rating method<direct-ranking method [a]	$r_{a,c}<r_{b,c}$ *	Schriesheim et al. (1991)	
		$r_{a,c}=-.12<r_{b,c}=.48$ *	Krosnick and Alwin (1988)	
	Direct-rating method>point-allocation method [a]	$p_a=88.0\%>p_b=73.7\%$ *	Bottomley et al. (2000)	
	Multi-attribute attitude method<trade-off method [a]	$r_{a,c}<r_{b,c}$ *	Adelman et al. (1984)	
	Rating-based conjoint method>trade-off method [a]	$r_{a,c}=.36>r_{b,c}=.25$ *	Akaah and Korgaonkar (1983)	
	Rating-based conjoint method<choice-based conjoint [a]	$p_a=50.1\%<p_b=60.5\%$ *	Vriens et al. (1998)	

[a] Both methods are proposed to yield similar results.
[b] Both methods are proposed to yield different results.
* $p<.05$.

6.2. Effects of framing

Part of the unexpected convergent and nomological validity reported in Table 3 involves methods that are proposed to measure the relevance of attributes (Proposition 2) versus methods that are proposed to measure the determinance of attributes (Proposition 3). Since these methods are proposed to measure different dimensions of attribute importance, one would expect them to show discriminant validity. However, Table 3 shows that the some of the results of these methods converge. A possible reason for these findings may be the effects of framing as a result of the nature and timing of the direct-method-related questions (direct-rating and point-allocation questions) (Vriens et al., 1998). For instance, if the point-allocation method is used directly following a task that involves attribute-level information, this point-allocation method starts to measure the determinance instead of the relevance attributes (Fischer, 1995). While the procedures used by Heeler et al. (1979) did avoid this potential effect of framing, the procedures used by Srivastava et al. (1995) and Schoemaker and Waid (1982) may have caused the reported lack of discriminant validity. Both studies measured the importance of attributes using the point-allocation method following a judgment or choice task that involved attribute-level information.

7. General discussion

This review revealed some key issues regarding the use of different methods to measure attribute importance. First, there is an apparent lack of convergent validity among and nomological

validity of different methods for measuring the importance of attributes. Second, distinguishing between the salience, relevance, and determinance of attributes, and acknowledging the antecedents of these dimensions of the importance of an attribute helps reduce the lack of validity from 67.0% to 31.9% of the method comparisons identified. Furthermore, acknowledging the multi-dimensionality also provided the basis for explaining part of the unexpected (lack of) convergent and nomological validity reported in Table 3; the possible framing effects. To reduce the apparent lack of validity among methods for attribute-importance measurement, we recommend that future research clarifies which dimension of attribute importance it focuses on and selects the appropriate measurement method accordingly.

7.1. Managerial recommendations

From a managerial perspective, attribute importance is most appropriately defined in relation to its behavioral outcomes (Pennings and Smidts, 2003). Although it is tempting to conclude that the determinance of attributes thus should be the key component in research, all three dimensions — the salience, relevance and determinance, should play an important role in research. First, if we only focus on determinant attributes, there is a risk of overlooking non-determinant attributes that are relevant. Consequently, "we might increase the leg space in airplanes, but ignore important safety features" (Myers and Alpert, 1968). The non-determinant, but relevant attribute actually is then likely to become determinant in a negative sense. Second, the effect of the difference in attribute levels on judgment and choice increases with the relevance of the attribute. Hence, identifying determinant attributes that are relevant is more efficient than identifying attributes that are merely determinant. Third, while determinant attributes are important when deciding which one of two products to buy (joint evaluation), relevant attributes are more critical when deciding whether or not to buy at all (separate evaluation). Since it is generally unknown which decision process applies best and most often, both relevant and determinant attributes should be available. Finally, in some decision processes, attribute information is externally available, while in other decision processes attribute information is only internally available, in memory. Since it is difficult to determine which "information context" applies best and most often, both determinant and salient attributes need to be available.

7.2. Limitations and future research

While this study revealed some important insights into those factors causing a lack of validity among available methods for measuring the importance of attributes, some cautionary notes are warranted. First, the limitations of reviews need to be acknowledged (Brinberg and Jaccard, 1986) and addressed in future research. Our broad approach yielded some key insights for future research, which are embedded in propositions and, as such, form the basis for a future line of research on valid attribute-importance measurement. Explicitly testing our propositions would be a valuable first step towards developing valid methods for attribute-importance measurement. More generally, future research may take a Multi-Trait–Multi-Method approach to studying the validity of attribute-importance measurement using differences in information content as a manipulation to examine the central proposition of this paper.

The emphasis of this research has been on the lack of validity. However, the review of the literature also revealed a less salient inconsistency with respect to the outcome of different methods — a lack of test–retest reliability. For instance, Fischer (1995) shows that the importance measures determined using the direct-rating method do not correlate with those determined using the trade-off method. Jaccard et al. (1986), however, conclude the opposite. Both studies are by no means perfect replication studies. Therefore, executing perfect replications to examine the test–retest reliability of approaches and methods is an important direction for future research (Srivastava et al., 1995).

Acknowledgment

The authors thank the Unilever Research Laboratory (Vlaardingen, Netherlands) for supporting this research. Furthermore, the authors also thank Asteria Straathof for her help during the review of the literature.

Appendix A. Studies used to examine the validity of methods for measuring attribute importance

Source	# respondents	Object	# and examples of types of attributes
Adelman et al. (1984)	20 lieutenants	Military actions	(5/9) undisclosed
Akaah and Korgaonkar (1983)	80 consumers	HMO plans	(6) deductions, clinic location, out-of-pocket payments, reputation, dependent eligibility, availability physicians
Barlas (2003)	40 students	Contraceptives	(10) e.g., complications, ability to later conceive children, prevents venereal diseases, pregnancy, main ingredient
Billings and Marcus (1983)	48 students	Apartments	(4) cleanliness, distance, rent, size

Appendix A (*continued*)

Source	# respondents	Object	# and examples of types of attributes
Borcherding et al. (1991)	200 students	Nuclear waste site	(11) e.g., workers–residents dying due to radiological exposure, workers–residents dying of an accident, impacts
Bottomley et al. (2000)	113 students	Car	(9) e.g., comfort, fuel consumption, safety and security features, ride and road handling, performance, reliability
Cook and Stewart (1975)	141 faculty and students	Financial aid	(3) need, GPA, faculty rating
		Admission rating	(7) verbal GRE, quantitative GRE, advanced test field, overall GPA, major field GPA, references, quality of education
Doyle et al. (1997)	70 students	Job applicants	(9) e.g., written & oral communication, leadership, team membership, decision making, motivation
Elrod et al. (1992)	115 students	Apartments	(4) rent level, one/two bedrooms, distance, neighborhood safety
Fischer (1995)	45/52 stud.	Job offers	(2) annual salary, # paid vacation days
Green et al. (1972)	43 consumers	Discount cards	(2) size of discount, # of cooperating stores in shopping area
Harte and Koele (1995)	22 counselors	Addicted clients	(9) e.g., duration of addiction, social support, goals, preferences, psychiatric disorders, # clinical treatments
Harte et al. (1996)	40 students	Banks	(6) credit facility, customer friendliness, # of cash dispensers, interest rate, familiarity with bank, distance
Heeler et al. (1979)	98 students	Electric blenders	(10) # speeds, safety, quietness, wattage, brand, material, ice crushing option, warranty, price, ease of cleaning
Hobbs (1980)	5 experts	Plant location	(18) e.g., surface water quality, transportation and transmission, state & federal lands, mineral resource
Jaccard et al. (1986)	110 students	Car	(9) e.g., cost, warranty, # doors, gas mileage, roominess, manufacturer repair record, friend's (dis)approval
		Birth control	(9) e.g., health consequences, side effects, attention, convenience of use, convenience obtaining product
Krosnick and Alwin (1988)	489 consumers	Child qualities	(13) e.g., manners, tries to succeed, honest, neat and clean, good judgment, self control, acts like a boy/girl
Leigh et al. (1984)	122 students	Calculators	(5) algebraic parentheses to assist calculations, rechargeability, financial functions, statistical functions, warranty

(continued on next page)

Appendix A (*continued*)

Source	# respondents	Object	# and examples of types of attributes
Maio et al. (1996)	211 students	Values	(36) Rokeach values
Neslin (1981)	112 consumers	Health service	(3) quality, personalness, convenience
Norman (1977)	24 students	Bus trip	(4) fare, walking distance to/from bus stop, # intervening stops en route, time of service during day
Oliphant et al. (1992)	300 consumers	Service package	(3) brand name, price, service features
Pöyhönen and Hämäläinen (2001)	407 students	Job alternatives	(11) e.g., salary, development opportunities, career opportunities, social environment, job security
Schoemaker and Waid (1982)	70 students	College. admission	(4) VSAT, CUM, EC, QSAT
Schori (1995)	1150 consumers	Insurance	(43)/(50) "insurance attributes" (undisclosed)
Schriesheim et al. (1991)	158 employers	Power bases	(5) reward, coercive, legitimate, referent, expert
Srinivasan and Park (1997)	121 students	Job offers	(8) travel, location, company growth, advancement opportunity, functional activity, environment, salary, people
Srivastava et al. (1995)	83 students	Apartments	(9) size, repair, noise, rent, utilities, deposit, distance from university, security, proximity to shops
Stillwell et al. (1981)	31 customers	Automobiles	(11) e.g., fuel economy, interior size, passing/acceleration ability, interior noise, ease of entry, maneuverability
	10 experts	Nuclear waste site	(7) e.g., public attitude, distance from city, geospheric path length, proximity to natural resources
Stillwell et al. (1983)	22 bank officers	Credit application	(7) undisclosed
Van der Pligt et al. (submitted for publication)	485 students	Safe sex	(3) of: e.g., decreases pleasure, protects against HIV, leads to interruptions, reduces sensitivity, is relaxing
	176 consumers	Smoking	(3) of: e.g., reduces fitness, reduces nervousness, increases coughing, helps to relax, increases one's popularity
	197 consumers	Student selection	(3) of: e.g., among others selection provides clear guidance after one year, selection after one year is unfair
Vriens et al. (1998)	185 consumers	Coffee makers	(5) brand, price, capacity, turning filter, thermos flask
Wiley et al. (1977)	32 students	Soft drinks	(3) calories, taste, carbonation
Zhu and Anderson (1991)	24 students	Job satisfaction	(3) co-worker, pay, work enjoyment

References

Adelman L, Sticha PJ, Donnell ML. The role of task properties in determining the relative effectiveness of multi-attribute weighting techniques. Organ Behav Hum Perform 1984;33:243–62.

Akaah IP, Korgaonkar PK. An empirical comparison of predictive validity for the self-explicated, Huber-hybrid, traditional conjoint and hybrid conjoint models. J Mark Res 1983;20:187–97 [May].

Alba JJ, Hutchinson JW, Lynch Jr JG. Memory and decision making. In: Robertson TS, Kassarjian HH, editors. Handbook of consumer behavior. New Jersey: Prentice-Hall; 1991.

Alpert MI. Identification of determinant attributes: a comparison of methods. J Mark Res 1971;11:65–74 [May].

Barlas S. When choices give in to temptations: explaining disagreement among Importance measures. Org Behav Hum Decis Process 2003;91:310–21.

Batra R, Homer PM, Kahle LR. Values, susceptibility to normative influence, and attribute importance weights: a nomological analysis. J Consult Psychol 2001;11(2):115–28.

Billings RS, Marcus SA. Measures of compensatory and noncompensatory models of decision behavior: process tracing versus policy capturing. Organ Behav Hum Perform 1983;31:331–52.

Borcherding K, Eppel T, Von Winterfeldt D. Comparison of weighting judgments in multiattribute utility measurements. Manage Sci 1991;37(12):1603–19.

Bottomley PA, Doyle JR, Green RH. Testing the reliability of weight elicitation methods: direct rating versus point allocation. J Mark Res 2000;27:508–13 [November].

Brinberg D, Jaccard J. A review of meta-analytic techniques. Adv Consum Res 1986;13:606–11.

Carpenter GS, Glazer R, Nakamoto K. Meaningful brands from meaningless differentiation — the dependence on irrelevant attributes. J Mark Res 1994;31(3):339–50 [August].

Cohen J, Cohen P. Applied multiple regression/correlation analysis for the behavioral sciences. 2nd edition. Hillsdale: Lawrence Erlbaum Associates, Inc.; 1983.

Cook RL, Stewart TR. A comparison of seven methods for obtaining subjective descriptions of judgmental policy. Organ Behav Hum Perform 1975;13:31–45.

Doyle JR, Green RH, Bottomley PA. Judging relative importance: direct rating and point allocation are not equivalent. Organ Behav Hum Decis Process 1997;70:65–72.

Elrod T, Louviere JJ, Davey KS. An empirical comparison of ratings-based and choice-based conjoint models. J Mark Res 1992;29:368–77 [August].

Farley JU, Lehmann DR, Sawyer A. Empirical marketing generalization using meta-analysis. Mark Sci 1995;14(3):G36–46.

Fischer GW. Range sensitivity of attribute weights in multiattribute value models. Organ Behav Hum Decis Process 1995;62:252–66.

Fishbein M, Ajzen I. Beliefs, attitude, intention, and behavior: an introduction to theory and research. Reading, MA: Addison-Wesley Publishing; 1975.

Ford JK, Schmitt N, Schechtman SL, Hults BM, Doherty ML. Process tracing methods: contributions, problems, and neglected research questions. Organ Behav Hum Decis Process 1989;43:75-117.

Green PE, Carmone FJ, Wind Y. Subjective evaluation models and conjoint measurement. Behav Sci 1972;17:288–99.

Harte JM, Koele P. A comparison of different methods for the elicitation of attribute weights: structural modeling, process tracing, and self-reports. Organ Behav Hum Decis Process 1995;64:49–64.

Harte JM, Koele P, Van Engelenburg G. Estimation of attribute weights in a multiattribute choice situation. Acta Psychol 1996;93:37–55.

Hedges LV, Olkin I. Statistical methods for meta-analysis. Orlando, FL: Academic Press; 1985.

Heeler RG, Okechuku C, Reid S. Attribute importance: contrasting measurements. J Mark Res 1979;16:60–3 [February].

Hobbs BF. A comparison of weight methods in power plant siting. Decis Sci 1980;11:725–77.

Hsee CK. The evaluability hypothesis: an explanation for preference reversals between joint and separate evaluations of alternatives. Organ Behav Hum Decis Process 1996;67:247–57.

Jaccard J, Brinberg D, Ackerman LJ. Assessing attribute importance: a comparison of six methods. J Consum Res 1986;12:463–8 [March].

Kaplan KJ, Fishbein M. The source of beliefs, their salience and predictions of attitude. J Soc Psychol 1969;78:63–74.

Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value tradeoffs. New York: Wiley and Sons; 1976.

Krech D, Crutchfield RS. Theory and problems of social psychology. New York: McGraw-Hill Book Co.; 1948.

Krosnick JA, Alwin DF. A test of the form-resistant correlation hypothesis: ratings, rankings, and the measurement of values. Public Opin Q 1988;52:537–67.

Leigh TW, MacKay DB, Summers JO. Reliability and validity of conjoint analysis and self-explicated weights: a comparison. J Mark Res 1984;21:456–62 [November].

Maio GR, Roese NJ, Seligman C, Katz A. Rankings, rating, and the measurement of values: evidence for the superior validity of ratings. Basic Appl Soc Psychol 1996;18(2):171–81.

Myers JH, Alpert MI. Determinant buying attitudes: meaning and measurement. J Mark 1968;32:13–20 [July].

Myers JH, Alpert MI. Semantic confusion in attitude research: salience vs. importance vs. determinance. Adv Consum Res 1977;4:106–10.

Neslin SA. Linking product features to perceptions: self-stated versus statistically revealed importance weights. J Mark Res 1981;18:80–6 [February].

Norman KL. Attributes in bus transportation: importance depends on trip purpose. J Appl Psychol 1977;62(2):164–70.

Oliphant K, Eagle TC, Louviere JJ, Anderson D. Cross-task comparison of ratings-based and choice-based conjoint. In: Metegrano M, editor. Proceedings of the sawtooth software conference; 1992.

Park CW, Lessig VP. Familiarity and its impact on consumer decisions biases and heuristics. J Consum Res 1981;8:223–30 [December].

Pennings JME, Smidts A. The shape of utility functions and organizational behavior. Manage Sci 2003;49:1251–63 [September].

Pöyhönen M, Hämäläinen RP. On the convergence of multiattribute weighting methods. Eur J Oper Res 2001;129:569–85.

Saaty SL. A scaling method for priorities in hierarchical structures. J Math Psychol 1977;15(3):234–81.

Schoemaker PJH, Waid CC. An experimental comparison of different approaches to determining weights in additive utility models. Manage Sci 1982;28:182–96 [February].

Schori TR. Note on attribute importance. Percept Motor Skills 1995;80 (3):1129–30.

Schriesheim CA, Hinkin TR, Podsakoff PM. Can ipsative and single-item measures produce erroneous results in field studies of French and Raven's (1959) five bases of power? An empirical investigation. J Appl Psychol 1991;76(1):106–14.

Schwer RK, Daneshvary R. Keeping up one's appearance: Its importance and the choice of type of hair-grooming establishment. J Econ Psychol 2000;21 (2):207–22.

Sethuraman R, Kerin RA, Cron WL. A field study comparing online and offline data collection methods for identifying product attribute preferences using conjoint analysis. J Bus Res 2005;58:602–10.

Srinivasan V, Park CS. Surprising robustness of the self-explicated approach to customer preference structure measurement. J Mark Res 1997;34:286–91 [May].

Srivastava J, Connolly T, Beach LR. Do Ranks Suffice? A comparison of alternative weighting approaches in value elicitation. Organ Behav Hum Decis Process 1995;63:112–6.

Steenkamp J-BEM, Van Trijp HCM. Attribute elicitation in marketing research: a comparison of three approaches. Mark Lett 1997;8(2):153–65.

Stillwell WG, Seaver DA, Edwards W. A comparison of weight approximation techniques in multiattribute utility decision-making. Organ Behav Human Perform 1981;28(1):62–77.

Stillwell WG, Barron FH, Edwards W. Evaluating credit applications: a validation of multiattribute utility weight elicitation techniques. Organ Behav Hum Perform 1983;32(1):87-108.

Van der Lans IA, Van Ittersum K, DeCicco A, Loseby M. The role of the region of origin and EU certificates of origin in consumer evaluation of food products. Europ Rev Agr Econ 2001;28(4):451–77.

Van der Pligt J, De Vries NK, Van Harreveld F, Judging attribute importance: a comparison of elicited weights, rank-based approximate weights, and statistically derived weights. Manuscript submitted for publication.

Van Ittersum K, Candel MJJM, Meulenberg MTG. The influence of the image of a product's region of origin on product evaluation. J Bus Res 2003;56 (3):215–26.

Van Ittersum K, Meulenberg MTG, Van Trijp HCM, Candel MJJM. Consumers' appreciation of regional certification labels: a pan-European study. J Agric Econ 2007;58(1):1-23.

Verlegh PWJ, Schifferstein HMJ, Wittink DR. Range and number-of-levels effects in derived and stated attribute importances. Mark Lett 2002;13(1):41–52.

Von Winterfeldt D, Edwards W. Decision analysis and behavioral research. New York: Cambridge University Press; 1986.

Vriens M, Oppewal H, Wedel M. Ratings-based versus choice-based latent class conjoint models. J Mark Res Soc 1998;40(3):237–48.

Wainer H. Estimating coefficients in linear models: it don't make no nevermind. Psychol Bull 1976;2:213–7.

Wansink B, Van Ittersum K. Stopping decisions of travelers. Tour Manage 2004;25:319–30.

Wansink B, Van Ittersum K, Painter JE. How descriptive food names bias sensory perceptions in restaurants. Food Qual Prefer 2005;16(5):393–400.

Wiley JB, MacLachlan DL, Moinpour R. Comparison of stated and inferred parameter values in additive models: an illustration of a paradigm. Adv Consum Res 1977;4:98-105.

Wilkie WL, Pessemier EA. Issues in marketing's use of multi-attribute attitude models. J Mark Res 1973;10:428–41 [November].

Wittink DR, Krishnamurthi L, Reibstein DJ. The effect of differences in the number of attribute levels on conjoint results. Mark Lett 1989;1(2):113–23.

Zhu S-H, Anderson NH. Self-estimation of weight parameter in multiattribute analysis. Organ Behav Hum Decis Process 1991;48:36–54.