Análise Psicológica (1994), 1 (XII): 79-92

Footnotes to Fishbein: The Contrast Model of Stimulus Evaluation (*)

MARC SCHOLTEN (**)
MANUELA FAIA CORREIA (***)

1. INTRODUCTION

Since its inception in social psychology, Fishbein's (1967b) multi-attribute model of attitudes has become a major tool in the practice of marketing research for the assessment of consumer perceptions of and preferences for choice alternatives in the market place. The popularity of the model is easy to comprehend: The model seems self-evident and is easy to implement. Perhaps due to its large-scale application, however, the theoretical roots of the model have generally been neglected.

In this paper, then, we wish to acquaint the reader with the theory underlying the Fishbein model and to straighten out the relationship between both. It will become apparent, however, that, at the intersection of theory and model, there arise serious problems with the Fishbein paradigm.

We will identify Fishbein's view on attribute «salience» as a major determinant of these problems. In our view, the failure to recognize dependency of attribute salience on the context in which choice alternatives are perceived is a major limitation of the Fishbein paradigm. We will extend the original theory so that a principled and more complete account of attribute salience is provided and formalize the extended theory with an adaptation of Tversky's (1977) contrast model of similarity. In concluding this paper, we will briefly discuss the strengths and weaknesses of the resulting framework for the marketing-research practice.

2. FOOTNOTES TO FISHBEIN

After a brief description of Fishbein's (1967b) learning theory of attitudes and the multi-attribute model formalizing this theory, we will discuss some major problems with the Fishbein paradigm in the study of perception and preference.

2.1. *The learning theory of attitudes*

The learning theory of attitudes is based on the general principle that attitude formation is a resultant of concept formation (Fishbein, 1967b). In the learning theory, a concept is defined as any discriminable aspect of a person's world and concept formation is defined as the association between a common response and a set of stimuli and other concepts. An attitude is defined as an evaluative response that is learned in conjunction with the formation of a concept; i.e., attitude

(*) This article is based on Faia Correia (1992) and Scholten (1993).

(**) Associate Professor, Universidade Moderna, Lisbon.

Correspondence should be sent to: Marc Scholten, Rua Diogo de Silves, 16-A, 1400 Lisboa, Portugal.

(***) Assistant, ISPA, Lisbon. UIIPOG member.

refers only to the evaluative aspect of the total meaning response.

Fishbein (1967b) reasons that, if concepts are formed in the way suggested by the learning theory, it follows that a person's responses to the concept should be some function of the learning process. Immediately after concept formation, the concept (now viewed as a stimulus) should elicit the set of stimuli and concepts (now viewed as responses) that have served to define it. The set of responses associated with the concept are thought of as a belief system or «habit family hierarchy». The higher a response in the hierarchy, the greater the probability that the response is associated with the concept, that is, the greater the belief strength. Each of these associated responses is itself a stimulus that elicits a learned evaluative response. There is always an evaluative response even though this response may be neutral (Fishbein, 1967a, 1967b). These evaluative responses are summative and, the summated evaluative response is associated with the concept. When presenting the concept as a stimulus, then, it elicits the summated evaluative response, in other words, the learned attitude. Finally, the higher the position of a response to the stimulus concept in the hierarchy (that is, the greater a belief strength), the greater is the amount of evaluative response that is available for summation.

Fishbein (1967b) formalized the learning theory of attitudes by positing the following multi-attribute model:

$$Ao = \sum_{i=1}^{n} (Boi * ai). \qquad (1)$$

In this equation, Ao is the attitude toward object o (the evaluation of object o), Boi is the probability that object o is associated with concept i (the strength of belief that object o is associated with concept i), ai is the attitude toward concept i (the evaluation of concept i) and n is the number of concepts in the system associated with object o.[1]

Importantly, Boi in equation (1) refers to the categorization of the relationship between object o and concept i along a probability dimension as judged by a person; that is, to the subjective probability of the relationship between object o and concept i. Analogously, Ao and ai refer to the categorization of object o and concept i along an evaluative dimension as judged by a person; that is, to the subjective goodness-badness of object o and concept i.

According to Fishbein (1967a, 1967b; Fishbein & Ajzen, 1972, 1975; Ajzen & Fishbein, 1980; Kaplan & Fishbein, 1969), it is only the «salient beliefs» that determine an individual's attitude toward an object at a given moment. Fishbein suggests that the first five to nine beliefs spontaneously elicited by an individual should be regarded as his «salient beliefs» about the object and should be included into the model. Analogously, one could construct a model for the entire sample (instead of a unique model for each individual in the sample) by including those beliefs most frequently elicited across individuals (referred to as the «modal salient beliefs»).

In the following three subsections, we will consider several issues related to the Fishbein paradigm. The first issue concerns the appropriateness of the conceptual definition of attitude that is proposed by Fishbein. In relation to this, we will discuss an often overlooked aspect of the learning theory of attitudes, which is that it is not only a theory of stimulus evaluation but also, implicitly, a theory of stimulus identification. The second issue concerns the operationalization of the multi-attribute model that formalizes the learning theory of attitudes. In relation to this, we will argue that Fishbein's suggestion to use salient beliefs for model construction actually interferes with an appropriate application of the model. The third and final issue concerns the adequacy of the learning theory of attitudes as a general theory of stimulus evaluation. In relation to this, we will discuss the implications and limitations of the learning theory of attitudes with regard to belief salience.

---

[1] Based on the learning theory sketched out previously, it is more appropriate to view the multi-attribute model as a formula for predicting the evaluative aspect of *concept* o (the representation of object o).

2.2. *Stimulus identification and stimulus evaluation*

According to Fishbein (1967a, 1967b), an attitude is the evaluative aspect of the total meaning response to an object. Thus, it is basically suggested that the attitude construct is a unidimensional construct. The question may be raised, however, whether it is appropriate to assume that the evaluative aspect of an object is unidimensional *if* one adheres to the learning theory of attitudes.

Fishbein (1967a) acknowledges that the cognitive aspects of an object may influence the evaluative aspect of an object. «Other types of meaning which the object may have for a person (...) are excluded from the notion of attitude except insofar as (the other types of meaning) may influence the placement which the person gives the object on the evaluative dimension.» (p. 258) With «the other types of meaning» Fishbein refers to the attributes associated with an object. The first point to make in response to this acknowledgement is that the attributes associated with the object not only *may* have an influence on the evaluation of the object but that, following the learning theory and its formalization by the multi-attribute model in equation (1), these attributes necessarily *have* an influence on object evaluation. The second point to make is that the attitude toward the object cannot be accounted for without prior specification of exactly those attributes that are characteristic of the object. Within the Fishbein perspective, there *is* no object meaning (whether evaluative or not) without «the other types of meaning».

Another way to make the second point is to acknowledge that the Fishbein theory is not only a theory of object evaluation but also, implicitly, a theory of object identification. One step in Fishbein's (1967b) reasoning is that, immediately after concept formation, the concept (now viewed as a stimulus) should elicit the set of stimuli and concepts (now viewed as responses) that have served to define it. If this arbitrary step is not taken, however, it becomes more apparent that object evaluation necessarily includes the object's attributes as meaning elements. Consider the following argument.

Suppose one presents a person with attributes which one a priori assumes to be those characteristic of some specific object or some specific set of objects. The task of the person is to indicate the probability that any object can be characterized by the attributes (object identification) and to indicate the characteristicness of the attributes vis-à-vis every object that is identified (the associative strength between attributes and object). The Fishbein theory predicts that the probability of object identification is an increasing function of the associative strengths between attributes and object. Now suppose, instead, that one presents the person with the same set of attributes and, before proceeding, let the person indicate the evaluation of the various attributes. Then the task of the person is to indicate the evaluation of each object that can be characterized by the attributes (object evaluation) and to indicate the characteristicness of the attributes vis-à-vis every object that is identified (the associative strength between attributes and object). The Fishbein theory predicts that object evaluation is an increasing function of the multiplicative combination of attribute evaluations and associative strengths between attributes and object. The point is that, at a conceptual level, the attitude toward the object is a composite response: It is constituted of both an identification response and an evaluation response. The identification response is a meaning response to the object's attributes as meaning elements.

The implication of the previous argument is that only the operational definition of the attitude toward an object should be regarded as unidimensional (the categorization of the object along a good-bad dimension). By the same token, the operational definition of object identification should be regarded as unidimensional (the categorization of the object along a probability dimension).

2.3. *Stimulus dimensions and stimulus attributes*

Consider the probability distribution in an individual's representation of brand X on four binary dimensions as displayed in Figure 1.

In the individual's representation of brand X, each dimension (D) has two categories or attributes (in abstract denoted by 1 and 0) and these attributes provide mutually exclusive and collectively exhaustive characterizations of the brand. Suppose we are dealing with the representation of some car brand, characterized by the following attributes: Small and large size (D1), catalyser and no catalyser (D2), five and four doors (D3) and two side mirrors and one side mirror (D4). Consistent with the learning theory of attitudes, there is a probability between 0 and 1 that an attribute is associated with the brand and there is a probability of 1 that either the one or the other attribute on a dimension is associated with the brand. Thus, for example, the belief hierarchy represents the individual's belief that cars of brand X are small with 80% probability and large with 20% probability.

FIGURE 1

	Brand X			
	D1	D2	D3	D4
1	0.80	0.60	1.00	0.60
0	0.20	0.40	0.00	0.40

Even though it seems safe to assume that the strength of belief involving one attribute (BX1) is inversely related to the strength of belief involving the other attribute on a dimension (BX0), the assumption of probabilistic symmetry (that is, BX1=1–BX0) may be too strong (Fishbein & Ajzen, 1975). The learning theory of attitudes, however, does not account for systematic departures from probabilistic symmetry. In what follows, then, we will simply assume that the assumption is correct.

The purpose of this subsection is to demonstrate that Fishbein's suggestion to include only «salient beliefs» about the object into the multiattribute model of attitudes (see equation [1]) may, in general, lead to inaccurate prediction of the attitude toward the object. We discuss model construction based on individual salient beliefs and model construction based on modal salient beliefs in succession.

2.3.1. Inclusion of individual salient beliefs

When constructing a model unique for each individual, Ajzen & Fishbein (1980) suggest the use of bipolar scales for measuring evaluative aspects and unipolar scales for measuring belief strengths.[2] Thus, the evaluation of the object (Ao) and the evaluation of the attributes (ai) may be measured on seven-point scales ranging from -3 («extremely unattractive») to +3 («extremely attractive»); the belief strengths may be measured on continuous scales ranging from 0% to 100% likelihood that the attributes are associated with the object or, alternatively, on four-point scales ranging from 0 («not at all likely») to +3 («extremely likely»).

Prior to scale construction, however, the individual elicits his beliefs about the object. Given the belief hierarchy in the above example, the individual would certainly not elicit the belief that cars of brand X «have four doors» and it would be extremely unlikely that he would elicit the belief that cars of brand X «have a large size». In contrast, the belief that cars of brand X «have five doors» and the belief that cars of brand X «have a small size» would most probably be among the salient beliefs and would, therefore, be included into the model.

What are the implications of this measurement procedure for the prediction of the individual's attitude toward brand X? We concentrate on the size dimension. If the individual indicates a +2 evaluation of «small size» cars and an 80% probability that cars of brand X are of «small size», this increases the predicted evaluation of brand X by (0.80 * (+2)) = 1.60 units. But what about the individual's belief that cars of brand X are of «large size» with a 20% probability? Since this belief is not among the salient ones, it is not included into the model. But if the individual has a nonneutral attitude toward «large size» cars, his nonsalient belief should nonetheless have an effect on the predicted evaluation of brand X. Assuming that the individual's evaluation «large size» cars is the polar opposite of his evaluation of «small size» cars (that is, a0 = -a1), the predicted evaluation of brand X should decrease by 0.40 units. This decrement in brand evaluation fails to manifest

[2] See Kaplan & Fishbein (1969), however, for measuring belief strenghs on *bipolar* scales while constructing models based on individual salient beliefs.

itself, however, if we follow Fishbein's guideline to include the salient belief only.

2.3.2. Inclusion of modal salient beliefs

When constructing a model for an entire sample of individuals, Ajzen & Fishbein (1980) suggest the use of bipolar scales for measuring evaluative aspects and belief strengths. Most typically, belief strengths are measured on seven-point scales ranging from -3 («extremely unlikely») to +3 («extremely likely»). The rationale for measuring belief strengths on bipolar scales is that an individual may actually *disbelief* a statement that he has not elicited himself. As we will see in the following, however, the measurement of belief strengths on bipolar scales raises more problems than it solves.

In an earlier work, Fishbein & Ajzen (1975) point out that responses toward the negative pole of bipolar probability scale should not be regarded as dissociations (negative associations), since the learning theory of attitudes permits only positive associations. If, in response to the belief statement «o is j», an individual chooses a position toward the negative pole of the probability scale, this means for the learning theorist that the individual rather believes that «o is (not j)». Viewed in this way, the bipolar scale is not truly a bipolar scale: It is a composite of two unipolar scales – one for attribute j and one for attribute (not j).

One problem with this solution, as also recognized by Fishbein & Ajzen, is that it works if and only if the evaluations of j and (not j) are equally polarized in opposite direction (that is, a1 = - a0). However, neither the learning theory of attitudes nor common sense dictates this evaluative symmetry. The individual in our example may evaluate cars with «fours doors» as quite attractive (as indicated by +2 on the scale) because such cars are easy to enter for passengers, whereas he may evaluate cars with «five doors» as extremely attractive (as indicated by +3 on the scale) because such cars combine ease of entry with ease of loading and unloading goods.

Evaluative asymmetry may lead to dramatically inaccurate predictions when using bipolar probability scales. Given the strengths of belief that brand X has either attribute, we should arrive at the prediction that the individual's evaluation of brand X increases by (0 * (+2)) + ((+3) * (+3)) = 9 units. That is, the attribute of which he believes that it is «not at all likely» to be associated with brand X adds nothing at all to the predicted evaluation and the attribute of which he believes that it is «extremely likely» to be associated with the brand adds 9 units to the predicted evaluation. When using bipolar scales, however, we obtain ((-3) * (+2)) = -6. Thus, we arrive at the prediction that the individual's evaluation of brand X decreases by 6 units rather than that it increases by 9 units.

Another problem with the use of bipolar probability scales, as also recognized by Fishbein & Ajzen, is that these scales provide relative rather than absolute measures of belief strengths. That is, a bipolar probability scale does not provide an unambiguous measure of either the belief that «o is j» or the belief that «o is (not j)»; rather, a response on a bipolar probability scale is best viewed as a function of the strength of the individual's two unipolar beliefs. When confronted with the statement that cars of brand X «have a small size», the individual in our example may comprise his belief that it is quite certain that cars of brand X will «have a small size» with his belief that it is slightly certain that cars of brand X will «have a large size» by responding on the bipolar scale that it is only slightly certain that cars of brand X will «have a small size». This ambiguity in measurement may further interfere with an accurate prediction of the individual's attitude toward brand X: A compromise strength of belief is paired with the evaluation of only *one* of the two attributes believed to be associated with the brand.

We can take the present argument against the use of bipolar probability scales one step further by raising the question: When will an individual choose the neutral position on such a scale? Assuming, as we have been doing throughout our discussion, that the object is represented along binary dimensions, the most reasonable answer seems to be: When the individual considers it equally probable that the object is associated *with either the one or the other attribute* on a dimension. But this runs counter to the learning theoretical view on the bipolar probability scale, which is that it is a composite of two unipolar scales. According to this view, the neutral position on the scale indicates that the individual believes that *neither the one nor the other attribute* is associated with the object. The conclusion is again that the measurement procedure may interfere with an accurate prediction of the individual's attitude toward brand X: Regardless of how attractive or unattractive the attributes are to the individual, the prediction of his attitude toward brand X remains unaffected.

A final problem with the use of bipolar probability scales is that this measurement procedure may confound attribute possession with probability of attribute possession whenever the representation of an object along a dimension is differentiated into more than two categories or attributes; that is, the procedure may confound *what* the individual believes with *how strongly* the individual believes what he does (see also Ahtola, 1975). Suppose that the individual's representation of brand X along the number of doors dimension is differentiated into four categories: Cars of brand X may have two, three, four or five doors. When confronted with the statement that cars of brand X «have four doors», there are many response strategies he could adopt. We mention only two relatively simple ones. One strategy is to reserve the «four doors» category for the positive end of the scale, to join the remaining categories on the negative end and to provide a probability rating as if the dimension had only two categories. Another strategy is to select the category that is most strongly associated with brand X and to assign it to a location on the scale that best indicates the relative number of doors. It is in such case that probability of attribute possession gets confounded with attribute possession. In either case, however, it seems doubtful whether one will arrive at an accurate prediction of the attitude toward brand X.

All problems discussed in this subsection basically arise from Fishbein's suggestion to use only salient beliefs for model construction. It seems safe to conclude, then, that his suggestion actually interferes with an appropriate application of the model. Especially the use of modal salient beliefs together with bipolar probability scales seems problematic. Given that Fishbein & Ajzen (1975) have recognized at least some of the problems discussed here, it may seem surprising that Ajzen & Fishbein (1980) adhere to this research strategy. But, in fact, it is hard to say what position these authors take on this issue. On the one hand, they contend that the most accurate predictions are likely to result when estimates of belief strength and evaluation for multiple categories on dimensions are obtained (Fishbein & Ajzen, 1975, p. 84). On the other hand, they doubt whether a consideration of multiple categories would greatly increase the predictive power of the attitude model (Fishbein & Ajzen, 1975, p. 103). A direct test of both models, reported by Ahtola (1975) and to be discussed in the following section, clearly favored a model with multiple categories on dimensions to one without.

2.4. *Attribute salience and its determinants*

According to Fishbein (1967b), it is only the «salient» beliefs that determine an individual's attitude toward an object at a given moment. As it follows from the learning theory, one forms beliefs about an object by associating it with various attributes; the set of attributes associated with the object is thought of as a belief system or «habit family hierarchy». The higher an attribute in the hierarchy, the greater the probability that it is associated with the object, that is, the greater the belief strength. Theoretically, the most salient beliefs are those uppermost in an individual's hierarchy. Operationally, Fishbein (1967b) assumes that the first five to nine beliefs spontaneously elicited by an individual are his salient beliefs about the object.

But the «magical number seven, plus or minus two» may not be the final word on the identification of salient beliefs. Fishbein & Ajzen (1975) are very clear on this point. «It is possible (...) that only the first two or three beliefs are salient for a given individual and that additional beliefs elicited beyond this point are not primary determinants of his attitude (i.e., are not salient). Unfortunately, it is impossible to determine the point at which a person starts eliciting nonsalient beliefs. Recommending the use of the first five to nine beliefs is therefore merely a rule of thumb.» (p. 218)[3] But if salient beliefs are to be

[3] See also Ajzen & Fishbein (1980, p. 64).

FIGURE 2

	Brand X			
	D1	D2	D3	D4
1	0.80	0.60	1.00	0.60
0	0.20	0.40	0.00	0.40

	Brand Y			
	D1	D2	D3	D4
1	0.25	0.50	0.25	0.25
0	0.75	0.50	0.75	0.75

identified with a rule of thumb, this means that we do not have substantive theory yet that specifies for us what beliefs determine an individual's attitude toward an object.

If Fishbein's learning theory of attitudes were an adequate theory of stimulus evaluation, there should be a one-to-one correspondence between the position of an attribute in the hierarchy and the probability that it is associated with a concept (Hackman & Anderson, 1968). It follows that there should be a continuous scale of salience; attributes with a higher position in the hierarchy (as operationalized, for example, by the ordinal position in a free elicitation task) and with a greater probability of being associated with a concept (as operationalized, for example, by a numerical belief strength in a probability rating task) should be thought of as more salient than attributes with a lower position in the hierarchy and with a smaller probability of being associated with the concept.

Fishbein (1967b), however, treats salience in an all-or-none way; attributes above some arbitrary threshold are in the hierarchy (salient) and those below the threshold are not in the hierarchy (nonsalient). The problem here is that the learning theory does not provide any rationale for treating salience in a discontinuous way and, therefore, does not provide a nonarbitrary threshold. Furthermore, Kaplan & Fishbein (1969) suggest that an attribute may have a great probability of being associated with a concept even though it is below the threshold and, therefore, nonsalient. The problem here is, of course, that any noncorrespondence between position in the hierarchy and probability of association cannot be accounted for by the learning theory. The basic question that arises is: What is salience if not another word for belief strength?

According to Ortony, Vondruska, Foss & Jones (1985), salient attributes are simply those that are readily accessible and *attribute salience may be determined by a variety of factors*. Acceptance of this proposal and a specification of factors determining attribute salience may suggest several shortcomings of the learning theory of attitudes as a general theory of stimulus evaluation. In what follows, we will discuss several determinants of salience.

Suppose that our exemplary individual actually has two car brands in his «evoked set» (Howard & Sheth, 1966) and that his representation of these brands can be schematized as in Figure 2.

Brand X is not anymore the only object in the psychological domain of car brands: Brand Y is an alternative. A context of alternatives to an object in some domain we call the *effective context* of that object. Thus, brand Y is the effective context of brand X and vice versa.

The attributes «catalyser» (category 1 on D2) and «two side mirrors» (category 1 on D4) have an equal strength of association with brand X; in other words, the *prominence* of these attributes in the representation of brand X is of equal magnitude. In a probability rating task, the individual is likely to indicate an equal strength of association with brand X for these attributes. Fishbein should be puzzled, however, if the individual would make a much earlier mention of «two side mirrors» in a free elicitation task (for example, as the third attribute) than of «catalyser» (for example, as the sixth attribute). This may be less puzzling if we take the prominence of these attributes in the representation of brand Y into account. Consider the following excerpt:

	D2			D4	
	X	Y		X	Y
1	0.60	0.50	1	0.60	0.25
0	0.40	0.50	0	0.40	0.75

It becomes apparent that the side mirror dimension (D4) is more relevant in discriminating brand X from brand Y than the catalyser dimension (D2). We refer to the relevance in effective context as the *diagnosticity* of a dimension (cf. Tversky, 1977). In a free elicitation task, the differential diagnosticity of both dimensions should increase the salience of the attribute «two side mirrors» relative to the salience of the attribute «catalyser» and this accounts for the fact that the individual mentions the former attribute much earlier than the latter one.

In general, then, an attribute may be very prominent in the representation of a brand and potentially salient; but if it is also very prominent in the representation of other brands and the attribute thus becomes rather nondiagnostic, there is possibly a decrement in the salience of that attribute (cf. Myers & Alpert, 1977).

In a similar vein, the above logic can account for the fact that the individual indicates an equal strength of association with brand Y for the attribute «four doors» (category 0 on D3) and the attribute «large size» (category 0 on D1) while eliciting the former attribute much earlier than the latter one. Even though these attributes are equally prominent in the representation of the brands, the number of doors dimension (D3) is more diagnostic than the size dimension (D1). This becomes apparent from the following excerpt:

	D1			D3	
	X	Y		X	Y
1	0.80	0.25	1	1.00	0.25
0	0.20	0.75	0	0.00	0.75

But the diagnosticity principle may not reach far enough in accounting for differential salience of attributes. During attribute elicitation, the individual may also be influenced by some internally generated context like «suitability for city traffic»: The attributes that are most characteristic of brand X (small size, catalyser, five doors and two side mirrors) make a car more suitable for city traffic than those that are characteristic of brand Y (large size, four doors and one side mirror).[4] Such context outside the effective domain, whether generated internally or presented externally, we call *general context*. Now suppose that the individual's representation of a car that is suitable for city traffic is as follows:

	Hypothetical car			
	D1	D2	D3	D4
1	1.00	0.50	0.60	0.75
0	0.00	0.50	0.40	0.25

We see that the probability distribution on the size dimension (D1) departs much more from a chance level distribution than the probability distribution on the number of doors dimension (D3). In this sense, D1 has a much greater relevance in the general context than D3. We refer to the relevance in general context as the *selectivity* of a dimension (cf. Smith, Osherson, Rips & Kean, 1988).

The diagnosticity principle and the selectivity principle may influence the relative salience of attributes in opposite directions. In our example, diagnosticity boosts the salience of «four doors» relative to the salience of «large size» and selectivity does the opposite. This becomes apparent from the following excerpt:

	D1		
	HC	X	Y
1	1.00	0.80	0.25
0	0.00	0.20	0.75
	D3		
	HC		
1	0.60	1.00	0.25
0	0.40	0.00	0.75

More specifically, the excerpt suggests that, in the present case, the influence of the selectivity

[4] Along the catalyser dimension, none of the attributes is most characteristic of brand Y.

principle is likely to overcome the influence of the diagnosticity principle and that, judged from free elicitation data, the salience of «large size» is considerably greater than the salience of «four doors».

Clearly, the overall conclusion may be that, to the extent that attribute salience is determined by factors other than feature prominence, the learning theory of attitudes is inadequate. The multi-attribute perspective on stimulus evaluation may profit from a joint consideration of the prominence of attributes, the relevance of dimensions in effective context (diagnositicity) and the relevance of dimensions in general context (selectivity) in determining what attributes will become salient and, in effect, what meaning the stimulus will have to the individual.

The purpose of the preceding discussion was to point out three basic difficulties with the Fishbein paradigm in the study of perception and preference. Our discussion has supplied us with sufficient ingredients to suggest some fundamental improvements of the multi-attribute perspective. In the following section, we will discuss modifications of the Fishbein model that have already been put forward by other investigators. None of the existing modifications, however, will offer a solution to *all* difficulties pointed out previously.

3. MODIFICATIONS OF THE FISHBEIN MODEL

3.1. *Modification 1: The Ahtola model*

Ahtola (1975) has modified the Fishbein model by incorporating multiple attributes on stimulus dimensions. Thus, the multi-attribute model in equation (1) is transformed into the following model:

$$Ao = \sum_{i=1}^{n} \sum_{j=1}^{ni} (Boij * aij),$$

where Ao is the evaluation of object o, Boij is the probability that object o is associated with attribute ij (the strength of belief that object o is associated with the jth category on dimension i), aij is the evaluation of attribute ij (the jth category on dimension i), n is the number of dimensions and ni is the number of categories (attributes) on dimension i.

The Ahtola model is the straightforward solution to all difficulties with the Fishbein model pointed out in 2.3., in that it identifies a set of mutually exclusive and collectively exhaustive (and, therefore, not necessarily salient) attributes on stimulus dimensions. As already noted, Fishbein & Ajzen (1975) doubt whether such modification would greatly increase the predictive power of the attitude model. The results obtained by Ahtola (1975) clearly suggest otherwise: The Ahtola model outperformed the Fishbein model with an average fit of 0.714 to 0.332.

Apart from this considerable improvement, the Ahtola model inherits all limitations of the Fishbein model with regard to the determination of attribute salience (see 2.4.). Only prominence is specified as a determinant of salience in the prediction of attitudes; additional factors, such as diagnosticity and selectivity, are not incorporated into the model. Both the Fishbein model and the Ahtola model, then, are adequate to the extent that stimulus evaluation can be considered context-independent.

3.2. *Modification 2: The Wyer model*

Wyer (1970) advances a model that incorporates multiple attributes on stimulus dimensions *and* a component representing the relevance of the dimensions to the evaluation of the stimulus:

$$Ao = \sum_{i=1}^{n} \sum_{j=1}^{ni} (Roi * Boij * aij),$$

where Ao is the evaluation of object o, Roi is the relevance of dimension i to the evaluation of object o, Boij is the probability that object o is associated with attribute ij, aij is the evaluation of attribute ij, n is the number of dimensions and ni is the number of attributes on dimension i.

What Wyer means by «relevance to the evaluation of the object» can be clarified with his suggestion that attributes on a dimension of «morality» may be more relevant to the evaluation of a priest than to the evaluation of a window-cleaner. That is, whether «moral» or «immoral» is specified as an attribute has a greater effect on the evaluation of a priest than on the evaluation of a window-cleaner. This differential relevance to the evaluation of an object is represented by the Roi component.

When may the object-specific relevance component add to the predictive power of the multi-attribute model? A possibility is that the usefulness of such component goes up when the dimensional commonality among the objects in the effective domain goes down. That is, the relevance component may be useful to the extent that the objects in the effective domain cannot be identified along the same dimensions. Viewed in this way, the Roi component regulates the entry of dimensions to the model. In case of multiple objects, Fishbein would select out dimensions for each object separately according to an all-or-none criterion and prior to model testing. With the Wyer model it is possible to select out dimensions according to a continuous criterion and during model testing.

Apart from enhancing the flexibility of the model in the prediction of attitudes toward multiple objects, however, the Wyer model exhibits the same limitations with regard to the determination of attribute salience as the Ahtola model.

3.3. *Modification 3: The Hackman & Anderson model*

According to Hackman & Anderson (1968), the evaluation of an object may not be stable across contexts. In order to obtain flexibility across contexts, these investigators suggest to add a component to the Fishbein formula representing the relevance of a belief to an external criterion. Here, belief relevance is defined as the «degree to which a particular belief is important to a subject in evaluating an attitude object when a *specific criterion* is used» (p. 56). Thus, the multi-attribute model in equation (1) is transformed into the following model:

$$Aco = \sum_{i=1}^{n} (Rci * Boi * ai),$$

where Aco is the evaluation of object o with regard to a specific criterion c, Rci is the relevance of attribute i to the specific criterion c, Boi is the probability that object o is associated with attribute i, ai is the evaluation of attribute i and n is the number of attributes associated with object o.

A strength of the Hackman & Anderson model is that it incorporates a component for, what we would call, belief relevance in general context or selectivity of beliefs. It is recognized that the attributes associated with an object may be of differential relevance to some specific criterion and that the evaluation of the object with regard to that criterion may be generated accordingly.

In spite of this theoretical improvement, our perspective suggests three limitations of the Hackman & Anderson model. Two of these limitations are directly inherited from the Fishbein paradigm. The first limitation is that the model fails to consider multiple categories on dimensions. The second limitation is that, although the model deals with belief selectivity (relevance in general context), it fails to consider belief diagnosticity (relevance in effective context). One may circumvent these limitations rightaway by advancing the following model:

$$Aco = \sum_{i=1}^{n} \sum_{j=1}^{ni} (Ri * Boij * aij), \qquad (2)$$

where Aco is the evaluation of object o with regard to general context c, Ri is the relevance of dimension i, Boij is the probability that object o is associated with attribute ij, aij is the evaluation of attribute ij, n is the number of dimensions and ni is the number of attributes on dimension i. The relevance component Ri is assumed to incorporate an aspect of relevance in effective context (a diagnosticity weight Di attached to dimension i) and an aspect of relevance in general context (a selectivity weight Si attached to dimension i), perhaps to be specified according to the formula Ri = (Di * Si).

This proposal, however, would leave the third, and final, limitation of the Hackman & Anderson model untouched. This limitation follows directly from the recognition that object evaluation implies object identification if one assumes that the learning theory is correct (see 2.2.): According to the learning theory of attitudes, no attitude toward an object can be formed without identifying the object on the basis of the attribute characteristic of that object. This applies, however, when the criterion is the evaluation of the object (Ao); when accounting for a meaning response such as the evaluation of the object with regard to some general context (Aco), the model in equation (2) is overly simplistic. The model only incorporates identification of the object, whereas it should incorporate (1) identification of the object, (2) identification of the general context and (3) an assessment of the similarity between object and general context. Without simultaneously including these aspects of context-bound stimulus evaluation into the model, a straightforward extension of the learning theory to the study of such meaning responses fails. In the following section, then, we will discuss a model that may be more appropriate for dealing with context-bound stimulus evaluation.

4. THE CONTRAST MODEL OF STIMULUS EVALUATION

Tversky's (1977) contrast model of stimulus similarity accounts for the psychological similarity of two stimuli by assuming an attribute-matching process. Similarity is expressed as a contrast between those attributes common to both stimuli, on the one hand, and those attributes distinctive of either one of the stimuli, on the other hand. By assuming that common attributes increase and distinctive attributes decrease the psychological similarity between two stimuli, the model accounts for observations like «an apple is more similar to a watermelon than to an olive».

Tversky's contrast model may account for context-bound stimulus evaluation given that two major modifications have been carried through. First, the model should include attributes *and* attribute evaluations. Second, the model should set a contrast between those attributes distinctive of the evaluated stimulus and those attributes common to the evaluated stimulus and contextual stimulus, on the one hand, and those attributes distinctive of the contextual stimulus, on the other hand.

Let us examine the resulting model with an example. Suppose that mister K. evaluates persons *only* in terms of the type of body that they have and that he distinguishes three types of body: Picnic, athletic and leptosomic. Suppose also that mister K. finds an athletic body very attractive (a=+3) but picnic and leptosomic bodies very unattractive (a=-3). If mister K. would encounter a single person on the street, he would probably evaluate that person in close correspondence with his evaluation of the person's body type; i.e., an athletic person would receive an evaluation of +3, whereas a picnic or leptosomic person would receive an evaluation of -3.

But what if mister K. would encounter *two* persons on the street? According to the previously proposed model of context-bound stimulus evaluation, mister K.'s evaluation of one person (the «focal person») in the context of the other person (the «contextual person») is a function of the evaluated attributes distinctive of the focal person *plus* the evaluated attributes common to the focal person and the contextual person *minus* the evaluated attributes distinctive of the contextual person. If mister K. evaluates a picnic or leptosomic person in the context of an athletic person, then, the focal person will receive an evaluation of -3+0-(+3)=-6. Conversely, if mister K. evaluates an athletic person in the context of a picnic or leptosomic person, the focal person will receive an evaluation of +3+0-(-3)=+6. Relative to the context-free evaluations, then, the evaluations become *more polarized*: something bad makes something good look even better by comparison, while something good makes something bad look even worse.

What if mister K. would encounter a picnic and a leptosomic person? The focal person, whether the picnic or the leptosomic one, would receive an evaluation of -3+0-(-3)=0. Relative to the context-free evaluations, then, the evaluations would become *more compromised*: two different bad things make each other look better by comparison. If, however, mister K. would encounter two persons with the same body type, his evaluation of either person would be the same as the context-free evaluation: *no contrast, no effect*.

Smith et al. (1988) have adapted the contrast model in order to account for typicality relationships between stimuli. Typicality relationships can be seen as similarity relationships

$$A_{co} = \sum_{i=1}^{n} R_i * \left(\sum_{j=1}^{n_i} a_{ij} * [(B_{oij} \cdot B_{cij}) + \min(B_{cij}, B_{oij}) - (B_{cij} \cdot B_{oij})] \right)$$

between two stimuli represented at different levels of abstraction; a typicality relationship is a similarity relationship between the representation of an instance of a concept and the representation of the concept. Thus, an apple is more or less «similar» to a watermelon and a more or less «typical» example of fruit. In the formulation proposed by Smith et al., the contrast model accounts for the typicality of an instance with regard to its concept by assuming a process in which frequency distributions on dimensions are matched with one another. The greater the commonality and the smaller the distinctiveness in frequency distributions when matching the representation of the instance and the representation of the concept, the more typical is the instance with regard to its concept. By assuming this, the model accounts for observations like «an apple is a more typical example of fruit than is a watermelon».

We may adapt Smith et al.'s version of the contrast model in order to account for evaluative responses to objects in some psychological domain (effective context) with regard to some general context.[5] First, we replace the frequency distributions by probability distributions. Second, we replace the contrast rule for stimulus similarity/typicality by the previously examined contrast rule for stimulus evaluation. Thus, we propose the model of context-bound stimulus evaluation that is specified at the top of this page.[6]

In our model, Aco is the evaluation of object o with regard to general context c, Ri is the relevance of dimension i (specified as in equation [2]), aij is the evaluation of attribute ij, Bcij is the probability that general context c is associated with attribute ij, Boij is the probability that object o is associated with attribute ij, n is the number of dimensions and ni is the number of attributes on dimension i. Furthermore, (Boij · Bcij) is the associative strength that is distinctive of o, min (Bcij, Boij) is the associative strength that is common to c and o, and (Bcij · Boij) is the associative strength that is distinctive of c.

To clarify the notion of commonality and distinctiveness in probability distributions, suppose that dimension i is a binary dimension: There is a probability of 0.60 that attribute i1 is associated with object o (Boi1=0.60) and a probability of 0.80 that this attribute is associated with the general context c (Bci1=0.80); By probabilistic symmetry, there is a probability of 0.40 that attribute i0 is associated with object o (Boi0 = 0.40) and a probability of 0.20 that this attribute is associated with the general context c (Bci0 = 0.20). Given these values, (Boij · Bcij) is 0.00 for attribute i1 and 0.20 for attribute i0; min (Bcij, Boij) is 0.60 for attribute i1 and 0.20 for attribute i0; and (Bcij · Boij) is 0.20 for attribute i1 and 0.00 for attribute i0.

Faia Correia (1992) and Scholten (1993) present an empirical comparison of Fishbein's learning theory of attitudes and the extended framework developed in this paper. Though obtained with a contrast rule theoretically inferior to the one proposed here, the results clearly favor the extended framework over Fishbein's learning theory as an account of context-bound stimulus evaluation.

5. CONCLUDING OBSERVATIONS

The academic interest in the effects of context on consumer perception, preference and choice is growing (e.g., Chakravarti & Lynch, 1983; Huber, Payne & Puto, 1982; Huber & Puto, 1983; Mishra, Umesh & Stem, 1993; Simonson

[5] Here, as elsewhere in this paper (see 2.4.), general context is conceived of as some hypothetical object corresponding to the psychological domain.

[6] For expository simplicity, we have omitted free parameters for the relative weight of the three probability distributions in setting up the contrast (cf. Tversky, 1977; Smith et al., 1988).

& Tversky, 1992; Tversky & Simonson, 1993). In this paper, we have expanded the Fishbein model, a popular tool for the assessment of consumer perceptions and preferences in marketing-research practice, so as to incorporate context dependency. Although this seems a major improvement to us, some preliminary reservations are in place.

To begin with, our theoretical analysis has been based on the assumption that the general context can be characterized along the same dimensions as the objects in the effective context. We have discussed general context in terms of a «hypothetical object» added to an effective context of «real objects». This has been convenient but equally limited in scope. Consider, for example, the possibility of adding a general context like «safari» to an effective context of cigarette brands. A dimension along which this general context might differentiate the various brands in the effective context (for example, a «purity» dimension) might not have been included in the representation of any of the brands before exposure to the «safari» context. This amounts to the well-known feature addition problem (e.g., Hampton, 1987; Smith & Medin, 1981; Smith et al., 1988), which is one of the guidelines in current research on concept formation and artificial intelligence.

Another limitation of our analysis, which it shares with the Fishbein paradigm, is the strictly elementaristic perspective. Attributes are treated as meaning elements in the formation of object meaning and, in the formation of an attitude toward the object, the evaluation of attributes appears as a major independent variable. However, the evaluation of attributes may not be that invariant. As an example, the evaluation of «sweet» may be different when associated with soft drinks than when associated with liquors (cf. Ahtola, 1975; Lutz & Bettman, 1977). As another example, the evaluation of «sweet» may be different when encountered in a «refreshment» context than when encountered in a «cocktail» context (cf. Ahtola, 1975; Lutz & Bettman, 1977). Yet as another example, the evaluation of «sweet» may be different when it co-occurs with «carbonated» than when it co-occurs with «uncarbonated». Clearly, the dependency of attribute evaluation on the object in the effective context with which it is associated, on the general context in which it is encountered and on the attributes with which it tends to go together curtails the appropriateness of the elementaristic perspective adopted in multi-attribute research.

While recognizing its inherent limitations, the framework presented here may improve the quality of multi-attribute modeling in marketing research. With regard to effective context, our approach may be relevant in that a brand is presumably not perceived and evaluated in isolation but with reference to its competitors in the product class. With regard to general context, our approach may be relevant in that a brand is often presented in the context of some persuasive communication with a selling proposition (i.e., advertising or other promotional activities). For marketing researchers, the contrast model may be a helpful tool in pretesting various communications against one another with regard to the perception and evaluation of the target brand that is provoked.

REFERENCES

Ahtola, O.T. (1975). The vector model of preferences: An alternative to the Fishbein model. *Journal of Marketing Research*, 12: 52-59.

Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs: Prentice-Hall.

Chakravarti, D. & Lynch, J.G. (1983). A framework for exploring context effects on consumer judgment and choice. In *Advances in consumer research*, Volume 10 (R. Bagozzi & A. Tybout, Eds.), pp. 289-297, Association for Consumer Research.

Faia Correia, M. (1992). *Fishbein in context: Models of stimulus typicality and attractiveness.* Monografia de Fim de Curso. Lisboa: ISPA

Fishbein, M. (1967a). A consideration of beliefs, and their role in attitude measurement. In *Readings in attitude theory and measurement* (M. Fishbein, Ed.), pp. 257-266, New York: John Wiley and Sons.

Fishbein, M. (1967b). A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In *Readings in attitude theory and measurement* (M. Fishbein, Ed.), pp. 389-400, New York: John Wiley and Sons.

Fishbein, M. & Ajzen, I. (1972). Attitudes and opinions. *Annual Review of Psychology*, 23: 487-544.

Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading: Addison Wesley Publishing Company.

Hackman, J.R.& Anderson, L.R. (1968). The strength, relevance, and source of beliefs about an object in Fishbein's attitude theory. *Journal of Social Psychology*, 76: 55-67.

Hampton, J.A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, 15(1): 55-71.

Howard, J.A. & Sheth, J.N. (1969). The theory of buyer behavior. New York: John Wiley and Sons.

Huber, J., Payne, J.W. & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9: 90-98.

Huber, J. & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, 10: 31-44.

Kaplan, K.J. & Fishbein, M. (1969). The source of beliefs, their saliency, and prediction of attitude. *Journal of Social Psychology*, 78: 63-74.

Lutz, R.J. & Bettman, J.R. (1977). Multiattribute models in marketing: A bicentennial review. In *Consumer and industrial buying behavior* (A.G. Woodside, J.N. Sheth & P.D. Bennett, Eds.), pp. 137-149, New York: Elsevier North-Holland.

Mishra, S., Umesh, U.N. & Stem, D.E. (1993). Antecedents of the attraction effect: An information-processing approach. *Journal of Marketing Research*, 30: 331-349.

Myers, J.H., Alpert, M.I. (1977). Semantic confusion in attitude research: Salience vs. importance vs. determinance. In *Advances in consumer research*, Volume 4 (W.D. Perreault, Ed.), pp. 106-110, Atlanta: Association for Consumer Research.

Ortony, A., Vondruska, R.J., Foss, M.A. & Jones, L.E. (1985). Salience, similes, and the asymmetry of similarity. *Journal of Memory and Language*, 24: 569-594.

Scholten, M. (1993). *The meaning of choice alternatives: Attitudes and images*. Tilburg: Unpublished Ph.D. Thesis.

Simonson, I. & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29: 281-295.

Smith, E.E. & Medin, D.L. (1981). *Categories and concepts*. Cambridge: Harvard University Press.

Smith, E.E., Osherson, D.N., Rips, L.J. & Kean, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, 18: 485-527.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4): 327-352.

Wyer, R.S. (1970). The prediction of evaluations of social role occupants as a function of the favorableness, relevance and probability associated with attributes of these occupants. *Sociometry*, 33: 79-96.

Tversky, A., Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39(10): 1179-1189.

ABSTRACT

Fishbein's (1967b) multi-attribute model of attitudes is a popular tool for the assessment of consumer perceptions and preferences in marketing research. This paper focuses on Fishbein's learning theory of attitudes, the multi-attribute model that he has offered as a formalization of his theory and on the relationship between both. Problems that arise at the intersection of theory and model are discussed. Fishbein's view on attribute «salience» is identified as a major determinant of these problems. Furthermore, the failure to recognize dependency of attribute salience on the context in which choice alternatives are perceived is argued to be a major limitation of the Fishbein paradigm. The original theory is extended and the extended theory is formalized with an adaptation of Tversky's (1977) contrast model of similarity. The strengths and weaknesses of the resulting framework for the marketing-research practice are briefly discussed.

RESUMO

O modelo de multi-atributos de atitudes de Fishbein (1967b) é uma ferramenta comum para aceder as percepções e preferências do consumidor em estudos de mercado. Este artigo centra-se na teoria da aprendizagem de atitudes de Fishbein, o modelo de multi-atributos que ele ofereceu como uma formalização da sua teoria e na relação entre eles. São discutidos os problemas que surgem na intersecção da teoria e o modelo. A perspectiva de Fishbein sobre a «saliência» dos atributos é identificada como a maior determinante destes problemas. Em seguida, é argumentado que a maior limitação do paradigma de Fishbein se deve à sua falha em reconhecer a dependência da saliência dos atributos do contexto em que as alternativas de escolha são percebidas. A teoria original é alargada e esta é formalizada com uma adaptação do modelo de contrastes de similaridade de Tversky (1977). São resumidamente discutidos os pontos fortes e fracos do enquadramento resultante para a prática de estudos de mercado.