FACULTY WORKING PAPER 92-0125

Revised - A Study of Comparative Judgments of Numerical and Verbal Attribute Labels

Madhubalan Viswanathan
Department of Business Administration
University of Illinois

Sunder Narayanan
Department of Advertising
University of Illinois

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 92-0125

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

April 1992

Revised - A Study of Comparative Judgments of Numerical and Verbal Attribute Labels

Madhubalan Viswanathan
Department of Business Administration

Sunder Narayanan
Department of Advertising

* Madhubalan Viswanathan is an Assistant Professor in the Department of Business Administration and Sunder Narayanan is an Assistant Professor in the Department of Advertising at the University of Illinois, Champaign, Illinois.

Abstract

Consumers often compare brands on product attributes using the information available to them. Such information in marketing communications is often conveyed in numerical and verbal modes. This paper examines comparative judgments of numerical and verbal labels describing product attributes. Using a comparative judgment task from cognitive psychology, comparisons of pairs of numerical and/or verbal labels are studied in two experiments. Results of the first experiment suggest that comparing a pair of numerical labels may be easier than comparing a pair of verbal labels or comparing a numerical label with a verbal label. Further, several past findings from cognitive psychology are replicated in the context of brand attribute comparisons. In addition, results also suggest that the utility property of a product attribute may influence comparisons on that attribute. A second study replicates the findings of the first study and provides support for the effect of the utility property of an attribute on comparative judgments.

A consumer at the supermarket examines a new brand of cereal on the shelf and notices from the package that the brand has 100 calories, and 4 grams of dietary fiber per serving. Further, s/he notices the claim on the package that the cereal is "very crunchy". Later that evening, while examining the brand of cereal that s/he purchases regularly, s/he notices that it has 150 calories, and 3 grams of dietary fiber per serving. Based on prior usage of the cereal, s/he rates it as being "quite crunchy". Trying to recall the information about the new brand in order to make comparisons with the regular brand on these attributes, s/he recalls that the new brand had 'low' calories (though s/he does not remember the precise numerical information), and a dietary fiber content of 4 grams per serving. S/he proceeds to compare the two brands on each of the attributes using the information available in order to choose between them. Numerical information on fiber content was available on both brands (i.e., "3 grams" versus "4 grams"), while verbal information on crunchiness was available for both brands (i.e., "quite crunchy" versus "very crunchy"). On the attribute, calorie content, verbal information was available on one brand while numerical information was available on the other (i.e., "150 calories" versus "low calories"). In this example, the consumer had to compare the two brands of breakfast cereals using numerical and/or verbal labels.

Consumers often face situations where they have to compare brands on attributes using numerical and/or verbal labels. Comparisons of brands on specific attributes may be performed by consumers in the course of making higher level decisions such as a judgment or a choice. Comparing brands on attributes may also be a means of learning about various brands in the market place. Consumer researchers have pointed out that comparisons of brands on attributes form an important element of consumer decision making and that consumers often compare various brands on attributes, particularly in tasks involving a brand choice (cf. Biehal and Chakravarti 1982). Further, the information that is used to make these comparisons is often in the form of numerical and/or verbal labels since marketing communications often involve the description of product attributes using numerical and verbal labels, such as in information conveyed on packages. In typical consumer settings such as the example above, information about brands on product attributes are conveyed using numerical labels (such as "200" calories) or verbal labels (such as "very crunchy"). In recent times, some consumer research has focussed on differences between numerical and verbal information (cf. Yalch and Yalch, 1984; Viswanathan and Narayanan, 1992).

This paper focuses on comparative judgments of numerical and verbal labels describing product attributes. Specifically, it focuses on comparisons of labels describing attributes which are in numerical and/or verbal modes or forms.[1] Using research on comparative judgments from cognitive psychology, hypotheses are generated and tested to achieve several objectives; first, to examine differences in the ease of comparisons between pairs of numerical labels, pairs of verbal labels, and pairs of numerical/verbal labels, second to assess whether some effects observed in past research from cognitive psychology generalize to consumer settings, and third, to study the effect of utility properties of product attributes on comparative judgments. Each of these objectives is important to consumer research in that it provides insights into the processes involved in making comparisons on product attributes. The first objective relates to possible differences in the ease with which consumers could use labels in numerical versus verbal forms to make comparisons. The second objective attempts to assess the generalizability of findings from cognitive psychology to consumer setting. The third objective relates to a unique aspect of consumer settings that does not hold in settings used in past research on comparative judgments, that product attributes have utility properties reflecting consumer preferences. By examining a range of effects, the aim is to provide a set of empirical findings that should be explained by models of comparative judgments in consumer settings. The rest of the paper is organized as follows. First, relevant research on comparative judgments from cognitive psychology is briefly reviewed. This is followed by a discussion leading to the generation of hypotheses about comparative judgments based on labels describing product attributes. Finally, details of two experiments conducted to test the hypotheses are presented.

REVIEW OF RELEVANT RESEARCH AND HYPOTHESES

Research on Comparative Judgments

Research in psychology has studied comparative judgments of stimuli along various dimensions (Banks et al. 1976; Holyoak 1978; Jaffe-Katz et al. 1989; Moyer and Landauer 1967). The comparative judgment task requires individuals to compare stimuli on a dimension and make judgments based on the magnitudes of the stimuli along that dimension. For example, subjects may be required to identify the larger (or smaller) of two stimuli (such as an elephant and a mouse) along a dimension such as size. Past research in comparative judgments has studied comparative judgments across a range of dimensions such as magnitudes of digits, size of objects, and pleasantness of stimuli. These studies have included comparisons based on numbers (Foltz et al., 1984: Jamieson and Petrusic, 1975) as well as verbal labels (Holyoak and Walker, 1976; Parkman, 1971). Two important effects, referred to as the symbolic distance effect and the semantic congruity effect (cf. Banks and Flora 1977), have been observed in several studies.

The symbolic distance effect is the finding that as the distance between two stimuli along a dimension increases, it is easier to make comparisons between the two stimuli (as evidenced by lower response times and/or higher accuracy of comparisons). For example, in a comparison task involving digits, a comparison between '1' and '100' is made faster (and/or more accurately) than a comparison between '1' and '3'. Studies have demonstrated the symbolic distance effect for comparisons involving numerical as well as verbal labels (Banks et al. 1976). The semantic congruity effect refers to the finding that comparisons are easier when instructions are congruent with the magnitudes of the stimuli being compared than when they are incongruent. For example, if a task requires a choice of the 'larger' item on a dimension such as magnitude of digits, decisions are made faster by subjects for a pair of large stimuli (such as '101' and '99') than for a pair of small stimuli (such as '3' and '1'), even though the difference in magnitude between the pairs of stimuli is the same. The semantic congruity effect has been obtained in several studies involving comparisons of numerical as well as verbal labels (Banks et al. 1976; Jaffe-Katz et al. 1989).

Several models have been developed to explain findings about comparative judgments. Some researchers (cf. Moyer and Landauer 1967) suggest that size comparisons of stimuli are made using analog representations of magnitudes. Hence, when comparing numbers, representations of their magnitudes are argued to have as many discriminable positions as there are numbers to be compared (i.e., if numbers from 1 to 100 are compared, then representations of a 100 discriminable positions are used). Others (cf. Banks 1977) suggest that comparisons are made using coarse-grained magnitudes (such as ‘large’ and ‘small’). One early attempt to explain the distance and congruity effects was by Holyoak (1978), who put forward a 'reference point' explanation for findings about comparative judgments. According to this explanation, in addition to the effect of distance, variations in discriminability may be produced due to reference points. Holyoak (1978) used a comparative judgment task where subjects were required to decide which of a pair of digits was closer to a third, which was the reference point. The findings supported the argument that discriminability is increased as the distance from the reference point decreases, as indicated by smaller response times for comparisons of pairs that were closer to the reference point. The semantic congruity effect is explained by arguing that instructions to "choose the higher" (or lower) of a pair of stimuli serve to provide a reference point (i.e., the high end of a continuum). Therefore, the congruity effect occurs since discriminability is increased for magnitudes that are closer to the reference point (i.e., "higher" magnitudes). The reference point may be implied in the question that is put forward by the experimenter (using an example adapted to consumer research, if the question is 'which brand has lower calorie content,' then the reference point for size may be the lower end of the continuum).

Jaffe-Katz et al. (1989) focused specifically on comparisons of numerical and/or verbal labels. They studied comparisons of pairs of probability expressions which were both verbal (VV), both numerical (NN), and one expression numerical and the other verbal (NV). The authors argued and found that faster comparisons are made of pairs of numerical when compared to verbal probability expressions since the relatively precise nature of numerical expressions leads to lesser overlap between a pair of numerical expressions. They argued that verbal probability expressions, due to their imprecision, tend to overlap with each other more than numerical expressions, thereby necessitating more repeated observations to arrive at a comparative judgment. Comparisons of numerical/verbal (NV) pairs were also argued to be slower than comparisons of numerical pairs due to the overlap between the numerical label and the verbal label. They found that, as hypothesized, comparison times for verbal/verbal (VV) and numerical/verbal (NV) pairs were greater than for numerical/numerical (NN) pairs. Other studies of verbal probability expressions have found a high variation in the magnitude values assigned to verbal expressions as well as a high degree of overlap (cf. Beyth-Marom 1982), thereby pointing to the relatively imprecise nature of verbal labels. Jaffe-Katz et al. (1989) suggest that their findings could be explained by a modification of Holyoak's (1978) models to allow for the effect of the precision of labels being compared.

Comparative Judgments of Numerical and Verbal Labels of Product Attributes

Past research on comparisons provides a basis to understand comparisons of labels describing product attributes by consumers. This section develops several hypotheses about comparative judgments based on numerical and/or verbal labels describing product attributes. The objectives of this research are threefold; first, to examine whether comparisons of pairs of numerical labels (NN) are easier than comparisons of pairs of verbal (VV) or numerical/verbal (NV) labels, second to assess whether the distance and congruity effects are observed for attribute labels, and third, to see how the utility property of a product attribute affects comparative judgments on that attribute. The first two objectives were addressed in Experiment 1 and the results were replicated in Experiment 2. The data from the two experiments was analyzed together to address the third objective.

Using a setting where pairs of numerical and/or verbal labels (i.e., NN, VV, and NV pairs) describing a range of values on a particular product attribute are presented with instructions to choose the labels conveying the higher (or lower) value, several hypotheses were generated and assessed. As suggested in past research, numerical labels appear to be relatively precise when compared to verbal labels. To the extent that the imprecision of verbal labels leads to a higher likelihood of overlap with other labels, it may be more difficult make comparisons involving verbal labels. In line with this discussion, the following hypothesis was tested where speed of comparisons was used as an indicator of the ease of making comparisons.

H1: Comparative judgments of attribute labels will be faster for pairs of numerical labels (NN), than for pairs of verbal labels (VV) or numerical/verbal label pairs (NV).

The symbolic distance effect has been observed in a wide range of studies involving comparative judgments (cf. Banks, 1977). A demonstration of this effect for product attributes would point to the need to develop models of comparative judgments or adapt models from past research to explain findings in consumer research. The distance effect points to the role of distances between stimuli in influencing comparative judgments. In light of the goal here to examine numerical and/or verbal labels, the role of distances between labels in comparative judgments can be assessed for each combination of numerical and/or verbal labels. In order to test whether this effect generalizes to comparisons on product attributes, the following hypothesis was assessed.

H2: The symbolic distance effect will be observed for comparisons of labels describing product attributes.

Similarly, in order to assess whether the semantic congruity effect generalizes to comparisons on product attributes, the following hypothesis was tested. The occurrence of this effect would point to the importance of explicit or implicit task instructions on comparative judgments involving product attributes. Specifically, instructions may serve to provide a reference point that may be used while making comparative judgments.

H3: The semantic congruity effect will be observed for comparisons of labels describing product attributes.

EXPERIMENT 1

Overview

In order to test the three hypotheses, a setting was used where pairs of labels describing a product attribute were presented to subjects who had to choose the higher (or lower) of each pair. Therefore, the experimental task was a speeded comparative judgment task. Given the nature of the three hypotheses, three factors were manipulated, namely the modes in which a label pair was presented (i.e., NN, NV, and VV), the distance between pairs of labels in terms of the magnitudes or values conveyed by them, and the instructions provided for comparisons (i.e., either to choose the higher or to choose the lower of a label pair). Within-subject rather than between-subject manipulations of mode (NN, NV, and NN), instructions ('choose higher' and 'choose lower'), and distance between stimuli were used in the study to provide a more efficient design, particularly in light of the need to have a sufficient number of trials by using all three combinations of mode and all levels of distance. Comparisons based on three combinations of stimuli (NN, NV, and NN) were used to test hypothesis 1. The distance between labels in terms of the magnitudes or values conveyed by them was manipulated to test the hypothesis on distance effect (hypothesis 2). Finally, the instructions ('choose higher' versus 'choose lower') was also manipulated to test hypothesis 3 about the semantic congruity effect.

Materials

Product Category and Attribute.

Subjects were required to perform comparisons along a single product attribute. Two criteria were important in the choice of the product attribute. First, subjects (university students) were required to have sufficient knowledge of the attribute magnitudes of the chosen product category in order to be able to make comparisons. Second, the product attribute should be easy to manipulate in numerical and verbal modes. Calculators appeared to be a suitable product category since students are familiar with this category and own calculators (Biehal and Chakravarti 1983). The attribute chosen for the first experiment was "display width" of calculators, an attribute that has been used in past research (cf., Childers et al. 1992).

Numerical and Verbal Labels.

The numerical and verbal labels used to describe this attribute were determined on the basis of a pretest. The aim of the pretest was to determine the number of levels of magnitudes to use to describe the product attribute and to generate pairs of numerical and verbal labels that were equivalent to each other in terms the magnitudes conveyed by them. The approach taken in the pretest was to use an open ended procedure to elicit impressions of magnitudes conveyed by a range of verbal and numerical labels. The range of numerical labels, and the verbal anchors to be used for the product attribute (i.e, "wide" and "narrow") were chosen on the basis of previous work (Childers et al. 1992). Thirteen verbal labels were chosen for the attribute by attaching a range of descriptors (such as "extremely") from previous research (Wildt and Mazis 1978) to the anchors used to describe display width. (Wildt and Mazis (1978) provide a list of 50 adverb-adjective combinations that were rated by subjects using a 21 point scale labeled at the end points as "the best (worst) thing I can say about a product".)

The pretest employed a magnitude scaling procedure and required subjects to estimate the magnitudes represented by a range of 13 verbal and 13 numerical labels for display width of calculators by drawing lines (or producing numbers) such that the length of the line (or the magnitude of the numbers) varied with the magnitude that was represented by the label (see Viswanathan and Childers (1992) and Viswanathan and Narayanan (1992) for more details on pretesting and use of stimulus materials). The pretesting procedure was used to determine the number of levels of labels that were used (by identifying clusters of verbal labels) as well as equivalent numerical and verbal labels (by identifying numerical labels that were equivalent to the chosen verbal labels).

On the basis of the pretest, five levels of magnitudes were chosen to describe the attribute. Five verbal labels and five equivalent numerical labels were chosen to represent the five levels of magnitudes to describe the product attribute. Therefore equivalent numerical and verbal labels that were chosen on the basis of the pretest formed a five point scale (the five levels in order of magnitude were "Extremely Narrow", "Narrow", Neither Narrow nor Wide", "Wide", and "Extremely Wide" for verbal labels, and "3 digits", "5 digits", "8 digits", "12 digits", and "18 digits", for equivalent numerical labels, respectively).

Based on the pretest, distances between pairs of labels in terms of the magnitudes conveyed by them were manipulated at four levels (distance between pairs were 1, 2, 3, or 4 units) since the there were five levels of magnitudes. The distance of 0 units was excluded for NN and VV since it would lead to the use of identical labels for these two combinations. All possible combinations of the five numerical and verbal labels of the attribute display width (except identical NN and identical VV labels), were used in the study resulting in a total of 45 pairs of labels; 10 NN pairs, 10 VV pairs, and 25 (5x5) NV pairs. To summarize, a three (mode; NN, NV, and VV) by two (instruction; choose higher versus choose lower attribute values) by four (distance between pairs; 1, 2, 3, and 4 units) factorial within-subject design was employed in the study.

Procedure

Thirty two undergraduate students at a midwestern university participated in the study for course credit. The experiment was conducted using Macintosh computers. Subjects performed several tasks to familiarize them with (i) the use of a mouse to provide responses, (ii) comparative judgments, (iii) and the use of labels to describe display width of calculators. Then, the actual experimental task was administered where subjects were familiarized with the set of labels used in the study and provided with instructions. The instructions required subjects to choose the higher (or lower) of a pair of labels. Subjects were also instructed to provide as fast a response as possible without compromising on accuracy in order to prevent them from performing the task at different points along the speed-accuracy curve, an important concern in using response time as the dependent variable. Each trial consisted of the presentation of a pair of labels on the left and right sides of the computer screen, respectively. Beneath each label was a square with the text "higher" (or "lower") depending on the instructions provided to the subject. Subjects were required to indicate their response (i.e, their choice of the "higher" or "lower" label) by clicking a mouse on the square below each label. Each trial was separated by a masked screen for three seconds to mark the end of a trial. The order of trials was randomized across all subjects. Subjects performed comparisons for two sets of forty five trials each, one involving the 'choose higher' decision, and the other involving the 'choose lower' decision. The order of these two sets of trials was counterbalanced across two groups consisting of an approximately equal number of subjects.

Results

Accurate responses were identified for the NN and VV conditions since these two conditions had objective criteria for determining accuracy. The accuracy levels for these two conditions were nearly 100%. The mean response times for accurate responses for NN and VV, and mean response times for NV were computed for each subject for each level of distance and instruction. Several ANOVAs were run on the data and are described below.

Response Times of Comparisons.

A two (instruction; choose higher versus choose lower) by three (mode; NN, NV, and VV) factorial ANOVA led to a significant main effect for mode ($F(2,62) = 19.41$; $p < 0.001$), and a significant interaction between instruction and mode ($F(2,62) = 4.12$; $p < 0.05$). The mean response time for the NN condition was found to be significantly less than for NV ($F(1,62) = 33.42$; $p < 0.001$) and VV ($F(1,62) = 24.04$; $p < 0.001$) conditions, with the difference between NV and VV being non-significant ($F(1,62) = 0.77$; $p>0.35$). Means for NN, NV, and VV were 3.87s, 4.34s, and 4.27s, respectively. These results support hypothesis 1.

Looking at the effect of instructions, the 'choose higher' decisions were made faster than 'choose lower' decisions, though the difference was not significant (Means = 4.08s and 4.24s, respectively for choose higher versus choose lower instructions). An examination of the interaction between instruction and mode suggested that the mean response time for 'choose higher' condition was significantly less than the 'choose lower' condition ($F(1,31) = 5.69$; $p < 0.05$), when both stimuli were verbal (VV), but not significantly less for NN and NV conditions.

Investigation of the Symbolic Distance Effect.

A two (instruction) by two (mode) by four (distance) factorial ANOVA led to a significant main effect for distance ($F(3,93) = 7.11$; $p < 0.001$), a significant interaction between mode and distance ($F(6,186) = 2.52$; $p < 0.05$), and a significant three-way interaction ($F(6,186) = 2.69$; $p < 0.05$). The mean response times for distances of 1, 2, 3, and 4 units were 4.22s, 3.94s, 4.08s, and 3.81s, respectively. In past research, the occurrence of the distance effect has been demonstrated by a significant main effect of distance. However, stronger support can be demonstrated by a monotonic decrease in response times with an increase in distance, thereby suggesting that comparisons are easier as distances between pairs of labels increase. As is evident from the means, such a monotonic decrease in response times was not obtained, thereby suggesting weak support for H2.

An examination of the interaction between distance and mode revealed significant distance effects for NV ($F(3,93) = 7.27$; $p < 0.001$) and VV ($F(3,93) = 5.70$; $p < 0.001$), but not for NN ($F(3,93 = 1.64$; $p > 0.18$). An examination of the interaction between distance and instruction revealed a significant distance effect for the 'choose higher' instruction ($F(3,93) = 7.36$; $p < 0.001$), but not for the 'choose lower' instruction ($F(3,93) = 1.94$; $p > .10$). On visual examination, none of the trends were monotonically decreasing except for VV in the 'choose higher' condition. It appears that H2 found support for the VV and NV conditions and the ‘choose higher’ condition.

Investigation of the Semantic Congruity Effect.

The semantic congruity effect relates to an interaction between the magnitudes of labels being compared and the instructions provided. Therefore, labels that were both high (i.e.., 4 or 5 on the 5 point scale based on the pretest) or both low (i.e., 1 or 2 on the 5 point scale based on the pretest) were selected for further analysis. A two (instruction) by three (mode) by two (size; both labels high, both labels low) factorial ANOVA led to a significant interaction between instruction and size ($F(1,31) = 6.02$; $p < 0.05$). 'Choose higher' judgments were made faster than 'choose lower' when both labels were high, whereas 'choose lower' judgments were made faster when both labels were low (Figure 2). Thus, hypothesis 3 was supported and the semantic congruity effect was found here.

Insert Figure 1 about here

Discussion of Results

Several interesting findings emerged from the analysis. First, the lower response times for NN labels indicated that they may be easier to compare than NV or VV labels. Support was also found for the occurrence of the semantic congruity effect. Only weak support was found for the symbolic distance effect. Further investigation suggested that support for the effect was found only for the VV and NV conditions and for the "choose higher" instructions. This findings is interesting since the presence of the distance effect would imply that distances between attribute values are utilized by subjects in making comparisons. An absence of the distance effect would, therefore, imply that comparison are made without taking into consideration the distances between the stimuli. Therefore, the lack of a distance effect for NN suggests that a subjective interpretation of distances between labels may have occurred to a lesser degree for NN when compared to NV or VV. Comparisons for NN labels may have been based on a surface level analysis rather than access to the meaning of the magnitudes conveyed by labels. The distance effect was found for 'choose higher' judgments, but not for 'choose lower' judgments, thereby suggesting greater subjective interpretation of magnitudes for the 'choose higher' judgments.

Faster comparisons and the detection of the symbolic distance effect for the "choose higher" instructions when compared to the "choose lower" instructions were interesting findings which could be interpreted as indicating the effect of the utility property of the attribute, display width. Assuming that higher magnitudes of display width (i.e., wider display width) are preferred over lower magnitudes, it could be argued that faster comparisons for the "choose higher" condition are found because the instructions are directionally consistent with the utility property of the attribute. Such an interpretation would point to the effect of the utility property of the attribute in influencing comparative judgments, and this possibility was investigated in experiment 2.

EXPERIMENT 2

A second experiment was conducted to replicate the findings of the first experiment and to further investigate the effect of the utility property of an attribute. In examining comparisons in the context of product attributes, it is important to note differences between product attributes and the dimensions typically used in past research. The typical dimensions used in past studies have been digits, size of animals, probability expressions, etc., for which subjects may not have had clear preferences for either higher or lower values on these dimensions. However, in natural consumer settings, product attributes have utility properties based on consumer preferences for higher or lower values on them. For example, in the case of calorie content of breakfast cereals, health conscious consumers may have a preference for lower calories. Thus, comparisons of labels describing product attributes may may be influenced by natural preferences about magnitudes of a product attribute.

In the first experiment, subjects may have made faster comparisons for the 'choose higher' condition when compared to the 'choose lower' condition since the comparison in the 'choose higher' condition was directionally consistent with the utility property of display width (on the assumption that utility for display width increases with increasing display width). The results obtained point to a potential interaction between the utility properties of attributes (preference for higher versus lower values) with instructions ('choose higher' or 'choose lower'). This possible interaction can be explained in terms of the directional consistency between instructions and utility properties. It is argued that comparisons are facilitated when instructions match preferences that consumers have about magnitudes on an attribute in terms of directionality. The premise here is that the natural preferences of consumers about magnitudes on an attribute make it easier to choose the more preferred rather than the less preferred label. If the explanation provided above is true, then an interaction should exist between the utility property of the attribute and the task. More formally, the following hypothesis was tested.

H4: Utility properties of attributes will interact with instructions such that, for attributes with positive utility properties, comparisons will be made faster for the 'choose higher' instruction whereas for attributes with negative utility properties, comparisons will be made faster for the 'choose lower' instruction.

Method

The basic procedure was similar to the procedure in experiment 1, except that the attribute, weight of a calculator, was chosen to represent an attribute with a negative utility on the assumption that heavier calculators were preferred to a lesser extent than lighter calculators. The instructions and materials were modified to be appropriate for the product attribute, weight. As a part of the same pretest described earlier, five verbal labels and five equivalent numerical labels were also generated for the attribute, weight. The chosen levels in order of magnitude were "Extremely Light", "Light", Neither Light nor Heavy", "Heavy", and "Extremely Heavy" for verbal labels, and "3 ounces", "9 ounces", "24 ounces", "60 ounces", and "96 ounces", for equivalent numerical labels, respectively). Twenty four undergraduates participated in the study. The results are described below.

Results

Response Times of Comparisons.

A two (task; 'choose higher' versus 'choose lower') by three (mode; NN, NV, VV) factorial ANOVA was performed on the reaction times. A significant main effect was found for mode ($F(2,46) = 41.13$; $p < 0.001$). The mean response time for the NN condition was significantly less than that for the VV ($F(1,46) = 24.44$; $p < 0.01$) and the NV ($F(1,46) = 16.92$; $p < 0.001$) conditions, thereby replicating the results of Experiment 1 and providing support for hypothesis 1. The mean response time for the VV condition was significantly less than that for the NV condition ($F(1,46) = 16.92$; $p < 0.10$), a result that was not observed in Experiment 1. The mean reaction times for the NN, NV and VV conditions were 3.67s, 4.45s, and 4.09s respectively. While hypothesis 1 was supported, some evidence was found that VV comparisons may be easier than NV comparisons.

The mean response time for the 'choose lower' task was marginally lower than for the 'choose higher' condition ($F(1,23) = 3.12$; $p < 0.10$). This is in contrast to the results of the previous experiment where, directionally, it was found that the response time for the 'choose lower' task was higher than that for the 'choose higher' task. This difference is important since it suggests that the utility property of the attribute may influence the comparisons, as is assessed subsequently.

Investigation of the Symbolic Distance Effect.

A two (task) by three (mode) by four (distance) factorial ANOVA did not result in a significant main effect of distance ($F(3,69) = 1.83$; $p < 0.15$). Therefore, hypothesis 2 was not supported here. However, there was a significant interaction between mode and distance ($F(6,138) = 3.38$; $p < 0.05$). A significant distance effect was found for the 'choose lower' task ($F(3,69) = 4.38$; $p < 0.01$) and but not for the 'choose higher' task condition ($F(3,69) = 2.35$; $p < 0.09$). A significant distance effect was found for the VV condition ($F(3,69) = 3.57$; $p < 0.05$), and but not for NN ($F(3,69) = 2.29$; $p < 0.09$) and NV ($F(3,69) = 2.01$; $p < 0.13$) conditions. Further, none of the distance effects for the various levels of mode appeared to be monotonically decreasing with increasing distances. Therefore, hypothesis 2 was not supported except for the VV condition and the “choose lower” instructions.

Investigation of the Semantic Congruity Effect.

A two (instruction) by three (mode) by two (size; both labels high, both labels low) factorial

ANOVA led to a marginally significant interaction between task and size ($F(1,23) = 3.03$; $p < 0.10$). An examination showed that the mean response times for the 'choose lower' task was less than that for the 'choose higher' task, when both labels were low and vice-versa. Thus, the findings replicate the results of experiment 1 and provide marginal support for the occurrence of the semantic congruity effect, as predicted in hypothesis 3.

Investigation of the Effect of Utility Properties of Attributes.

The data from the two experiments were combined in the next stage of analysis in order to assess hypothesis 4. A two (attribute type; positive (display width) versus negative (weight) utility property) by two (instructions) by three (mode) factorial ANOVA led to a significant interaction between attribute type and instructions ($F(1,54) = 4.29$; $p < 0.05$). For the attribute with the positive utility property, faster judgments were made in the 'choose higher' condition when compared to the 'choose lower' condition, whereas, for the attribute with the negative utility property, faster judgments were made in the 'choose lower' condition when compared to the 'choose higher' condition, thereby providing support for hypothesis 4 (see Figure 2).

Insert Figure 2 about here

In addition, a significant interaction was also obtained between attribute type and mode ($F(2,108) = 3.97$; $p < 0.05$). While for the NN and the VV conditions, the mean response times for judging the positive utility attribute was higher, for the NV condition, the mean response time for judging the negative utility attribute was higher.

Discussion of Results

The second experiment replicated the results of the first experiment in that NN comparisons are faster than NV or VV comparisons. It was also found that VV comparisons were made faster than NV comparisons, a result that was not observed in experiment 1. The second experiment replicated the results of the first experiment in demonstrating the semantic congruity effect. However, the symbolic distance effect was not found except for the VV condition and the "choose lower" instructions. In addition, hypothesis 4 was supported, providing evidence for the effect of the utility properties of attributes on which comparisons are performed. Therefore, it appears that the natural preferences of consumers on attributes influence the ease of making comparisons.

In looking at the results of the two experiments, the symbolic distance effect was found for the 'choose higher' instructions for the attribute with positive utility property, while the symbolic distance effect was found for the 'choose lower' instructions for the attribute with negative utility property. This finding could be interpreted as suggesting that when the instructions match the utility property of the attribute, it is more likely that distances between labels are used in making comparisons (i.e, that the meaning of the magnitude conveyed by a label is being used in making comparisons). Further, the distance effect was found for the VV condition in both experiments, thereby suggesting that the meaning of magnitudes conveyed by labels may be accessed to a greater extent for verbal pairs than for the other combinations.

GENERAL DISCUSSION

This paper examined comparative judgments based on numerical and verbal labels. Using a comparative judgment task from cognitive psychology, comparisons of pairs of numerical and/or verbal attribute labels were studied in two experiments. In both experiments, comparisons involving a pair of numerical labels were found to take less time than comparisons involving a pair of verbal or a numerical/verbal pair. This result suggests that comparisons involving a pair of numerical labels may be easier than for a pair of verbal labels or a numerical/verbal pair, probably due to the more precise nature of the information in the case of numerical labels when compared to the verbal labels. The results of experiment 2 also suggested that it may be easier to compare verbal pairs when compared to verbal/numerical pairs, a finding that was not observed in experiment 1.

In addition, several findings from cognitive psychology were tested in the context of comparison of product attributes. The results showed evidence for the existence of the semantic congruity effect, thereby suggesting the interaction between task instructions and the magnitudes of the labels being compared. The reference point explanation (Holyoak, 1978) may be applicable to comparisons in consumer settings in that instructions provide an explicit reference point which is used to make comparisons. The results did not provide consistent support for the symbolic distance effect in contrast to past research in cognitive psychology. Further, the results suggested that an important factor unique to consumer settings affecting comparisons may be utility property of the attribute on which comparisons were being made. It appears that comparisons are facilitated when the utility property of an attribute is consistent with the instructions provided, due to the natural preferences that consumers have about magnitudes on product attributes. This finding extends research on comparative judgments to include effects due to the directional consistency between instructions and the properties of the attribute on which comparisons are being performed. While findings such as the semantic congruity effect relate to congruency between magnitude of the labels being compared and instructions, perhaps due to the provision of reference points, this finding brings out the importance of directional consistency in influencing comparative judgments. The argument made here is that there may be certain natural preferences which facilitate comparisons in one direction (i.e., higher or lower) on an attribute. Therefore, models of comparative judgments should allow for baseline differences in comparisons in one direction versus the other due to the effect of factors such as natural preferences on an attribute.

This paper contributes to consumer research in several ways. First, it provides a set of empirical findings that should be explained by models of comparative judgments in consumer settings. Second, it provides insights into the process of comparative judgments in consumer settings, by studying the impact of various factors form past research such as information mode, the nature of instructions, and the distance between the labels. Third, it brings out the effect of utility properties of product attributes, a factor unique to consumer settings. In terms of methodological implications, this research uses the comparative judgment task from cognitive psychology. This methodology could be used to study several issues of importance in consumer research such as encoding and retrieval of attribute information.

Limitations

Several limitations of this study should be noted. First, given the aim of this paper to study comparisons under controlled settings, subjects performed the artificial task of comparing labels presented to them. In realistic settings, information may not be available to consumers in such a convenient form. Comparative judgments may often be based on both information that is immediately available and information from memory. Experimental control was essential given the lack of research in this area, but future efforts should focus on settings more reflective of typical situations where consumers make comparisons. Second, the basis for assessing correspondence between verbal and numerical labels as well as the number of levels of magnitude to be utilized was a cross-modal magnitude scaling procedure. However, the results of the pretest were used in aggregation rather than at the individual level to create the labels. Hence, individual differences may exist in the perception of distances between labels and, therefore, the manipulation of distance may not have applied for all subjects in the experiments. Third, this research used response time as the dependent variable and made inferences based on findings on this variable. Future efforts involving comparisons of brands based on information presented could focus on other dependent variables such as accuracy of comparisons. This is of particular relevance for memory-based comparisons in order to assess the impact of factors such information mode on accuracy of comparisons. Finally, in terms of focus, this paper studied comparative judgments based on information that was presented to subjects in the form of labels. Future efforts should focus on providing information on brands and assessing brand comparisons on specific attributes.

Several avenues of future research are suggested here. One line of research could focus on the effect of reference points on comparative judgments. For example, product familiarity may result in prior knowledge of the mean value or the range of values on an attribute which may serve as reference points for comparative judgments. The provision of explicit reference information may also influence comparisons. Another line of research can focus the effects of information presented in different modes such as graphical and pictorial information in facilitating comparisons. The occurrence of the distance and the congruity effects for comparisons of labels presented in different forms (for example, nutritional information preprocessed to different degrees and presented either as raw information or in relation to some standard, such as USRDA) can be examined. As suggested earlier, another line of research should focus on memory-based comparisons.

In broad terms, consumer research should focus on developing and testing models that explain comparative judgments in consumer settings. While comparisons may be one of several operations that consumers perform during a choice or judgment task, comparisons of brand information either to other brands or to internal or external reference points may form the primary means of evaluating a brand on an attribute. In line with models from past research, direct comparisons of information of brands as well as the impact of implicit and explicit reference points should be studied. In closing, this study brings out the importance of studying comparative judgments on product attributes, an important facet of consumer decision making.

References

Banks, W.P., and J. Flora (1977). "Semantic and perceptual processes in symbolic comparisons." Journal of Experimental Psychology: Human Perception and Performance, 3, 278-290.

Banks, W.P., M. Fuji, and F. Kayra-Stuart (1976), "Semantic congruity effects in comparative judgments of magnitude of digits," Journal of Experimental Psychology: Human Perception & Performance, 2, 435-447.

Biehal, Gabriel and Dipankar Chakravarti (1982), "Information Presentation Format and Task Goals as Determinants of Consumers Memory Retrieval and Choice Processes," Journal of Consumer Research, 8, 431-41.

Foltz, G., S. Poltrock, and G. Potts (1984). "Mental comparison of size and magnitude: Size congruity effects," Journal of Experimental Psychology: Learning, Memory & Cognition, 10, 442-453.

Jamieson, D.G., and W.M. Petrusic (1975). "Relational judgments with remembered stimuli," Perceptions and Psychophysics, 18, 373-378.

Holyoak, K.J. (1978). "Comparative judgments with numerical reference points," Cognitive Psychology, 10, 203-243.

Holyoak, K.J., and J.H. Walker (1976). "Subjective magnitude information in semantic orderings," Journal of Verbal Learning & Verbal Behavior, 15, 287-299.

Jaffe-Katz, A., D. Budescu, and T.S. Wallsten (1989). "Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty." Memory and Cognition, 17 (3), 249-264.

Moyer, R.S., and T.K. Landauer (1967). "Time required for judgments of numerical inequality," Nature, 215, 1519-1520.

Parkman, J.M. (1971), "Temporal aspects of digit and letter inequality judgments," Journal of Experimental Psychology, 91, 191-205.

Viswanathan, M. and T. Childers (1992), "The Encoding and Utilization of Magnitudes of Product Attributes: An Investigation Using Numerical and Verbal Information," Working Paper, University of Illinois.

Viswananthan, M. and S. Narayanan (1992), "The Processing of Numerical Versus Verbal Attribute Information: A Study Using Information Acquisition Patterns," Forthcoming, Marketing

Letters.

Yalch, Richard F., and Rebecca Elmore Yalch (1984), "The Effect of Numbers on the Route to Persuasion," Journal of Consumer Research, 11, 522-527.

Footnotes

1 While comparative judgments could be based on information recalled from memory or on information that is immediately available or a combination of the two, the focus here is on comparative judgments based on labels that are presented to consumers. Comparisons in natural consumer settings may involve either of the two or some combination that has elements of both and therefore, both types of comparisons are of importance to consumer research. Given past research on comparisons based on presented information in cognitive psychology, and the lack of research on comparisons in consumer behavior, such a focus provides a logical starting point for consumer research.

FIGURE 1

RESULTS SHOWING THE SEMANTIC CONGRUITY EFFECT

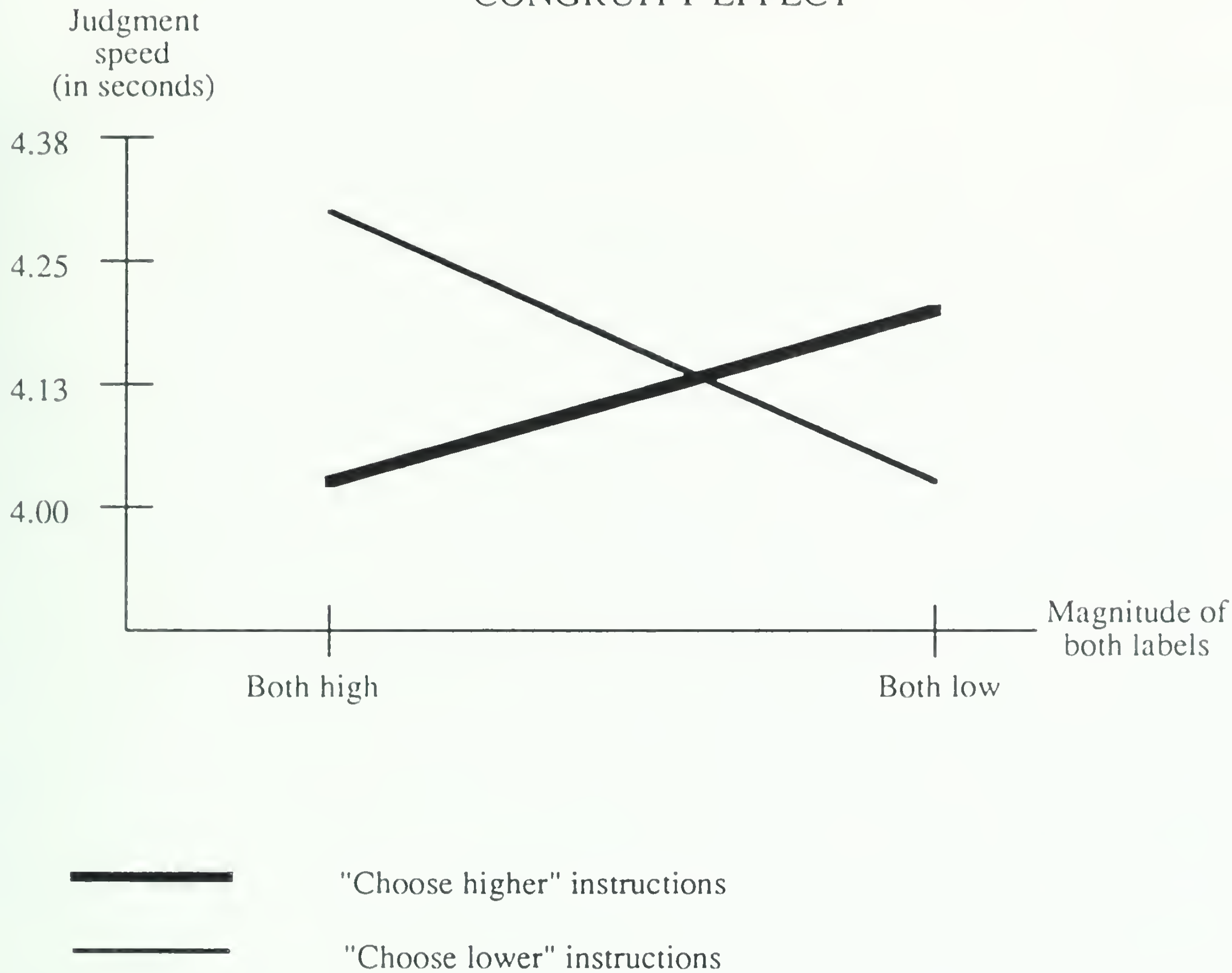

FIGURE 2

RESULTS SHOWING THE EFFECT OF UTILITY PROPERTIES OF AN ATTRIBUTE