December 1993

The effect of discordant, concordant and alphabetical orderings of brand quality rankings on decision making

Robert E. Widing II
W. Wayne Talarzyk
Gene Brown

Follow this and additional works at: http://epublications.bond.edu.au/discussion_papers

Recommended Citation
http://epublications.bond.edu.au/discussion_papers/46
DISCUSSION PAPERS

"The Effect of Discordant, Concordant and Alphabetical Orderings of Brand Quality Rankings on Decision Making"

by

Robert E. Widing, II
School of Business
Bond University

W. Wayne Talarzyk
College of Administrative Science
The Ohio State University

Gene Brown
College of Business Administration
Louisiana Tech University

DISCUSSION PAPER NO 46

DECEMBER 1993

University Drive, Gold Coast, QLD, 4229

AUSTRALIA
SCHOOL OF BUSINESS
DISCUSSION PAPERS

Bond University was established by Act of Parliament in 1987 as an independent, private University. The first student intake occurred in May 1989. The School of Business offers degrees in the undergraduate (BCom and Diploma) and the graduate (MBA, MAcc, MCom and PhD) levels.

The School teaches and sponsors research in accounting, economics, econometrics, finance, marketing, management, organisational behaviour and related disciplines in hospitality and real estate fields.

The Discussion Paper series is intended to foster research and comments are invited. The views expressed in the papers are the opinion of the authors and do not necessarily reflect the views of the School or the University.

Lists of available Discussion Papers and copies of the papers (which are free of charge) may be obtained from:

The Senior School Administrator
School of Business
Bond University
GOLD COAST QLD 4229

Telephone: (075) 95 2244
Fax: (075) 95 1160

Dean: Professor Ashley W. Goldsworthy AO OBE
THE EFFECT OF DISCORDANT, CONCORDANT, AND ALPHABETICAL ORDERINGS OF BRAND QUALITY RANKINGS ON DECISION MAKING

Robert E. Widing II *
School of Business
Bond University

W. Wayne Talarzyk
College Of Administrative Science
The Ohio State University

Gene Brown
College of Business Administration
Louisiana Tech University

December 9, 1993
* Robert E. Widing II is Associate Professor of Marketing, School of Business, Bond University, Gold Coast, Queensland, 4226 Australia. W. Wayne Talarzyk is Professor of Marketing, College of Business Administration, The Ohio State University, Columbus, OH 43210. Gene Brown is Associate Professor of Marketing, College of Business Administration, Louisiana Tech University, Ruston, LA. The authors thank Jan Heide and Gary Crenwelge for their helpful comments on earlier drafts of the manuscript and C. Kent Harrington for his creative programming work. The education division of Tandy Radio Shack Inc. is gratefully acknowledged for their supply of equipment.
THE EFFECT OF DISCORDANT, CONCORDANT, AND ALPHABETICAL ORDERINGS OF BRAND QUALITY RANKINGS ON DECISION MAKING

ABSTRACT

A concordant ordering of brands, generated with a computer assisted linear format, an alphabetical listing of brands, and a discordant summary rating of brands were compared on decision quality, user evaluations, and decision time. The decision task was difficult as the decision set contained a large amount of information and negative correlations among the attributes. The computer assisted format was most highly regarded by decision makers, and resulted in superior decision quality in comparison to the other two formats. While the discordant ordering was inferior to the concordant one, it was not significantly worse than an alphabetical ordering. Therefore, the likelihood of misleading independent decision makers is not as high as might be expected, even in a worst case situation. Overall, the computer assisted linear format was superior to the other two formats. It was, however, recommended that this format be positioned as a method to present information in an orderly fashion, and not as a normative decision aid. This is due to a minority of subjects using the linear aid who actually selected the top ranked brand.
INTRODUCTION

Methods to increase information search and improve buyer decision making include enhancing the accessibility and processability of information (Widing and Talarzyk 1989; Russo et al. 1986; Scammon 1977; Bettman 1979). Since the late 1970's, computer-based electronic information systems, sometimes referred to as videotex, have begun to emerge which might reduce the effort required to engage in information search and improve purchase decisions. Due to the development and diffusion of these new communications systems, a number of authors have called for the research and development of buyer information (BI) systems that capitalize on their strengths (Thorelli and Thorelli 1977; Beales et al. 1981; King and Hill 1989; Widing and Talarzyk 1989). Only limited research, however, has been conducted on computer assisted information presentation aids for buyers. This study represents a step toward correcting this deficiency.

Videotex enables information seekers with personal computers to obtain shopping and product information services from a central database. Several well-publicized regional videotex failures, such as Knight-Ridder's "Viewtron" and Times-Mirror's "Gateway," have dampened optimism about the future of these electronic systems. Less attention, however, has been paid to lower profile but comparatively successful national systems, such as CompuServe, General Electric's "GEnie," Dow Jones' "NewsRetrieval," and IBM and Sears' "Prodigy." The Prodigy system, for example, is a rapidly expanding system with over 1.1 million subscribers nationwide. Beyond offering shopping services from a variety of full and discount price "stores", product and service evaluations are available in electronic versions of Consumer Reports, Zagat's Restaurant Guide, and the Mobile Travel Guide. Hence, information on both "search" and "experience/credence" qualities is available to help guide decision making. (c.f. Darby and Karni 1973). If buyers find they can access reliable product rating information easily and inexpensively (Prodigy charges $12.95 per month for unlimited usage of all services), they may become more willing to investigate the virtues and liabilities of different brands.

Clearly, such systems can make an enormous amount of product information available to users, thereby satisfying the "availability" criterion for enhancing decision making. As buyers,
including households, institutions and organizations, have access to increasing amounts of
information, however, how can they best cope with, process, and make sense out of all the data?
Since decision makers have limited time, willingness and capabilities in analyzing information
(Russo et al. 1986; Klein and Yadav 1989; Shugan 1977), this presents a challenging task in order
to satisfy the "processability" criterion for improving decision making.

This research, therefore, examines the relative value of three formats that are available to
present information and/or aid consumers in decision making. Two of the formats are traditional,
passive formats that present brands in either (1) alphabetical order (ALPHA), or (2) in best to worst
order as determined by the rating organization (RATER). It is important to highlight that the
RATER format studied here will present information in a rank order that is not in agreement with
user preferences, since it is only in this case that summary ratings pose much of a threat to decision
making (Curry and Faulds 1986; Kopalle and Hoffman 1992). Although research has been
performed that identifies when discordant rank orders are most likely to occur, no research has been
undertaken to examine how decision makers might adapt to such an ordering. The third format
enables each user to generate a personalized best to worst rank order of brands, using a computer
assisted linear model (CAL). The CAL format rank orders brands in concordance with decision
maker values.

Finally, it is relevant to assess the subjective reactions of users to alternative formats and
decision aids. "User friendly" is a well-worn phrase, but still represents an important task facing
developers of BI systems. Russo (1986) notes that if information seekers do not like an information
acquisition or decision aid, it is not likely that they will use it regardless of its purported benefits.
Following is a review of past research on these formats and the effects of the decision environment
on choice outcomes.
PAST RESEARCH

Even useful and readily available information will not be thoroughly evaluated if it requires more effort to process than the perceived benefits of doing so (Russo et al. 1986). A great deal of research indicates that consumers are either unwilling or unable to thoroughly evaluate product information as information loads exceed five or six brands (Payne 1976; Lussier and Olshavsky 1979; Billings and Marcus 1983), which has been observed to lead to satisficing behavior (Simon 1955) and less than optimal decision processes and choices (Keller and Staelin 1987; Klein 1983; Klein and Yadev 1989; Malhotra 1982; Jacoby, Speller and Kohn 1972). Given the voluminous amount of product information that is or could be made available on videotex systems, enhancing information processability becomes increasingly important.

The two traditional methods of enhancing information processability have been (1) to provide information in formats that facilitate intrabrand evaluations and interbrand comparisons, such as the matrix format; and (2) to summarize brand information for the user in some fashion, typically with a best to worst rank order of brands. New communications technologies allow for the development of a third method, the (3) provision of computer aids that serve to present product information based on the individual user's preferences.

The standard formats used to provide unordered information include (1) by-brand, (2) by-attribute, and (3) the brand-by-attribute matrix (cf. Bettman 1979). Studies comparing the value of these three formats indicate that the matrix format provides for quicker decision making and is the most preferred format by decision makers (Bettman and Zins 1979). This format facilitates either brand or attribute processing, allowing decision makers the flexibility to switch strategies at will. The brands can be presented in any order using a matrix format, but the unaided order would typically have brands listed alphabetically. Consumer Reports, for example, uses alphabetical listings of brands for product classes in which consumers are likely to be highly heterogeneous in their attribute preferences (Pittle 1984).

A summary brand evaluation is an overall measure of quality that is obtained by combining the scores of individually rated attributes for each brand, typically with an additive compensatory
model (Newman 1977; Pittle 1984; King and Hill 1989). The summary brand evaluations have often been used to rank brands in order from best to worst; in addition, a matrix of attribute scores for the brands is generally available for perusal by the information seeker. This format is typified by the best to worst rank order format found in Consumer Reports.

The key potential problem with summary ratings rests in the process of combining individual attribute ratings to obtain overall brand evaluations. The linear model, using "importance" weights to modify the contribution of each attribute to the overall evaluation, is usually applied by rating organizations to calculate summary ratings (Meyer and Johnson 1989; Pittle 1984; Beales et al. 1981; Newman 1977). This procedure has been controversial since people do not necessarily apply importance weights similar to those of the evaluator does and, therefore, might be misled (Widing and Talarzyk 1989; Day 1976).

Although not much of a problem when attributes are non-negatively correlated (Dawes and Corrigan 1974; Einhorn and Hogarth 1975; Wainer 1976; Newman 1977), it is a potential problem in decision environments that have systematic tradeoffs among the attributes (Newman 1977). Two conditions need to exist for consumers to be misled: first, some attributes need be negatively correlated; second, the rating organization and information seeker weights on these attributes must be in conflict and of sufficient extremes in order to meaningfully affect the resulting rank orders (Curry and Faulds 1986; Meyer and Johnson 1989; Kopalle and Hoffman 1992).

There are two possible extreme responses to rater ordered information (Pittle 1984). First, totally dependent decision makers (DM's) may simply follow the rater's recommendation by selecting the top rated brand. In the case of concordant rank orders, decision quality is not likely to be reduced and could be greatly enhanced, whereas with discordant rankings decision quality is likely to be poor. Second, DM's may elect to independently process the information and, thereby, ignore the rank order. With discordant rank orders, DM's engaging in further information processing may adapt by realizing that the top ranked brands are less suitable than those ranked further down the ratings. Indeed, even a cursory evaluation of the top ranked brands in a discordant ranking might inspire further processing, since these brands would not likely perform well on the
more important attributes for the DM. Decision quality, therefore, need not necessarily be lower than having had the information presented without the summary rankings (e.g., in alphabetical order). Of course, in the case of a concordant ranking, further information processing would simply reinforce the rater's recommendation.

In summary, rankings that are in concordance with decision maker values should enhance, or certainly not greatly hurt, decision quality regardless of the decision strategy. Indeed, such rankings should not seriously detract from decision quality even in the case of DM's blindly selecting the top ranked brand. What is of interest is to study how DM's confronted with a rank order not in concordance with their values perform. Clearly, in this case, those blindly selecting the top ranked brand would perform poorly. However, it is not clear if more independent DM's adapt to a discordant rank order by recognizing and ignoring it, or are actually misled by it.

Due to the large amount of information that can be made available electronically, as well as the inherent analytical capabilities of a computerized medium, computer assisted information acquisition and decision aids should become increasingly attractive and feasible methods for decision makers to acquire and analyze product information. One such information acquisition aid is the computer assisted linear format (CAL). Decision scientists have long advocated linear models for structuring and aiding individual and group decision making (cf. Edwards, et al. 1975; King and Hill 1989). The procedures used for applying a linear model in decision aiding are, in essence, identical to those used by rating organizations to obtain brand evaluations (Newman 1977). An important difference is that the individual's importance weights, as opposed to the rater's, are used to generate the ratings.

The compensatory processes embodied by linear models have been associated with optimization, since they assume a thorough evaluation and balancing of information, a hallmark of sound decision making (Payne, Bettman and Johnson 1988). To our knowledge no extant research, however, has empirically supported the superiority of linear decision aids in actually yielding better decisions versus unaided decision making (cf., Aldag and Powers 1986). Computer assisted linear models have also been used to aid in actual buyer decision making. For example, Software Digest,
a rater of computer programs, has made a computer assisted linear model available to its subscribers to aid in the analysis of software ratings. The firm offers a floppy disk which contains attribute scores for the tested brands and a program that enables the user to specify importance weights. These weights and scores are then combined to yield summary brand ratings and rankings. *Consumer Reports* has also experimented with a CAL format (Pittle 1984). Since a primary objection to summary ratings is the rater weighting of attribute information, this approach eliminates a point of controversy.

**Decision Environment Effects on Decision Making**

In addition to the formats used to present information to consumers, the decision environment itself has also been shown to have profound affects on decision processes and outcomes. In particular, high context complexity (i.e., negative correlations, extreme attribute values, and large numbers of non-dominated alternatives) creates the opportunity for different outcomes to be attained depending on different processes, models, and weights that are applied to decision sets. Context effects, for example, appear to create the setting for: (1) screening errors to occur with decision makers using noncompensatory processes (Widing, Burnkrant, and Talarzyk 1986; Klein 1983); (2) noncompensatory and compensatory processes to yield different outcomes (Einhorn, Kleinmuntz and Kleinmuntz 1979); (3) the opportunity for the generation of inappropriate summary rank orders by ratings organizations (Newman 1977; Curry and Faulds 1986; Meyer and Johnson 1989); and (4) the lessening of linear model robustness in discrete choice prediction (Johnson, Meyer and Ghose 1989).

It is also desirable to incorporate high task complexity (e.g., a large number of brands and/or attributes or time pressure) in a study of format performances, since decision makers thoroughly process low levels of information quite readily. With larger amounts of information, problems associated with contingent (phased) processing, "pure satisficing" behavior (choose the first acceptable alternative), and information overload become increasingly likely to emerge (Payne 1976; Lussier and Oshavsky 1979; Malhotra 1982; Keller and Staelin 1987). Together, a decision
environment incorporating high task and context complexity is a difficult one that allows for
discrimination among format performances.

HYPOTHESES

Decision Quality

Summary ratings have been cited as potentially misleading decision makers. It is of interest,
therefore, to discern the impact of a discordant rank order on decision making. It needs to be
highlighted that this is a "worst case" scenario. This is appropriate, as discussed below, since a rank
order in concordance with DM values would obviously enhance decision making (Curry and Faulds
1986; Meyer and Johnson 1989); therefore, a relevant focus of analysis is how decision makers
react to a rank order not in agreement with their values.

Although it is reasonable to suspect that DM's might be misled by a discordant rank order, it is
also possible they would recognize an inappropriate rank order and ignore it. If an inappropriate
rank order is ignored by decision makers, then their decisions would not necessarily be worse than
those based upon an unbiased alphabetical format. If the discordant rank order is not ignored or
does influence choice, however, one would expect an alphabetical ordering to result in a superior
decision. It is hypothesized, however, that the RATER format will perform less well than the
ALPHA format, since it should certainly not perform better and, consistent with prior assumptions
in the literature, may perform worse.

Hypothesis 1: A discordant RATER order format will result in inferior decision quality in
comparison to an alphabetical listing of brands format (ALPHA).

Both alphabetical and rater orderings are subject to screening errors in complex decision
environments. A screening phase is induced by high task complexity (Payne 1976; Lussier and
Olshavsky 1979; Billings and Marcus 1983), whereas the opportunity for errors to be made with a
screening phase is enhanced by high context complexity (Widing, Burnkrant and Talarzyk 1986;
Klein 1983; Johnson and Payne 1985). Further, DM's faced with high information loads might also
adopt a "pure satisficing" decision rule (select the first acceptable alternative) and may be susceptible to dysfunctional outcomes attributable to information overload (such as frustration and confusion resulting, perhaps, in satisficing or random choice-like outcomes).

It is reasonable to assume that a computer assisted linear (CAL) format is less susceptible, relative to the alphabetical and rater orderings, to the potentially dysfunctional decision processes cited above. The rank order provided by the CAL format is based on the individual DM’s importance weighting policy and, at the least, should be helpful in organizing information. That is, dominated brands and lesser performers are positioned toward the bottom of the rank order, whereas the higher scoring brands on the more important attributes would be listed toward the top. Therefore, clutter ought to be reduced and the ease of making comparisons enhanced.

Hypothesis 2: The CAL format will result in superior decision quality in comparison to the ALPHA and discordant RATER formats.

An additional issue related to the CAL format also deserves investigation. That issue is to determine how often the top ranked brand is indeed selected, which would help determine if only the best (or top few) or many brands ought to be presented for a visual evaluation of the ratings. Since the rank order is based on the DM’s weighting scheme, passing over the top ranked brand would suggest that a summary rating based on a linear combination is not appropriate to recommend as a normative aid. Ironically, the CAL format might hamper decision making if positioned as a tool to select the best choice. Such a finding would also support the Meyer, Johnson and Ghose (1989) contention that discrete choice predictions are lessened in efficient/negatively correlated decision environments. This, coincidentally, also calls into question the validity of using linear combinations to establish an experimentally determined best choice with which to gauge decision quality (Meyer and Johnson 1989). This latter issue will be further addressed later in the paper.
Subjective Evaluations of the Decision Tool

Along with assessing decision quality using objective criteria, it is also desirable to examine reactions and attitudes of DM's toward the alternative information provision formats. Russo (1986) states that "the success of any information system for consumers depends upon consumer acceptance...". Other research studying format and information load effects have asked questions about format satisfaction, confusion and frustration experienced in the choice process, and certainty and confidence that the best choice was made (Bettman and Zins 1979; Jacoby, et al. 1974; Malhotra 1982; Keller and Staelin 1987; Klein and Yadav 1989). Using similar items, this study will also measure differences in preferences for the three formats.

Research Issue 1: Do differences exist on subjective reactions among the users of the CAL, RATER and ALPHA formats?

A related issue addressed in this study deals with the perceived and actual time subjects took to make a decision. It is of interest to discern if differences exist across formats, since time can be a barrier to search and has been used as an indicator of the processing effort expended in choice (Bettman and Zins 1979). Time, however, might be a misleading indicator of mental effort expended to make a decision with computerized aids. For example, a computer assisted format might take more time, due to the time required to enter weights and perform the operations necessary to generate the rank orders, but require less thinking about decision alternatives.

Research Issue 2: Do differences exist on perceived and actual decision making time among the users of the CAL, RATER and ALPHA formats.

METHODOLOGY

Pretests

The formats, choice sets, measures, and procedures used and discussed below were derived from a program of extensive pretesting. Over 100 subjects were used to assess training procedures and scripts, refine measures and formats, examine differing numbers of brands and product classes.
and generally evaluate all aspects of the research. Following is a description of the methodology employed, which evolved over the course of pretesting.

**Format Descriptions.**

The computer assisted linear format (CAL) required decision makers to specify relative attribute importance weights (using a constant sum to 100 scale) (Exhibit 1). The weights were multiplied times the corresponding rating score and the resulting weighted average was used to rank order the brands. The attribute scores for the brands were also provided in a matrix format. "Importance" weighting was used since the attributes satisfy the monotonicity assumption (i.e., higher scores on each attribute are inherently better). Therefore, alternative models, such as the ideal point or Fishbein models, were unnecessary for this decision set. The constant sum scale was used since past research indicates it performs as well as, or better than, other subjective weighting methods in predicting outcomes (Stillwell, Seaver and Edwards 1981), and has been recommended since it requires users to make explicit tradeoffs among attribute importances (Wilkie and Pessimier 1973).

The unaided format (ALPHA) presented the ratings in a brand by attribute matrix with the brands listed in alphabetical order. The display characteristics were the same as for CAL (Exhibit 1), except the "Importance" weight column was not included. The summary rating format (RATER) ranked the brands in best to worst order, accompanied by the matrix of attribute scores. This display was also identical to CAL except for the exclusion of the "Importance" weight column. The RATER rank order, as discussed earlier, was based on a weighting policy that was not in agreement with those of the subjects, since it is only in this case that summary ratings pose much of a threat to decision making (Curry and Faulds 1986; Meyer and Johnson 1989). To verify this was attained, a manipulation check was conducted, which will be presented later in this section. The "Rater Rank Order" of the brands is indicated in Exhibit 1 in the far right-hand column.
Choice Set Characteristics

The choice set is a "real" one, containing actual performance evaluations, to preclude potential criticisms about nonrepresentative stimuli that might affect outcomes (Goldberg 1971; Billings and Marcus 1983). The stimuli consisted of 20 brands of word processing programs evaluated on 6 attributes by Software Digest. Word processing programs were selected since this data set fit the requirements of a difficult environment and other data sets were available for training that used the same attributes. Fifteen of the 20 brands were non-dominated and represent the only viable best choices. Actual brand names were used to foster task realism; however the brand names were shifted about, while keeping the six attribute scores intact to avoid problems with copyright issues. Subjects were instructed to make their decisions based only on the attribute scores and to ignore any other considerations such as brand name. While many subjects reported having used or purchased a word processing program, especially mainframe or university supplied packages, no subject cited having used any of the programs included in the study. More importantly, no subject reported having had prior experience with the brand they chose for either the initial best and second best choices or the final choice. A question asking subjects if they restricted their decision only to the attribute information was answered affirmatively by all subjects.

Software Digest evaluated the brands on six attributes: Ease of Start-up, Ease of Use, Ease of Learning, Error Handling, Performance, and Versatility. They were scored using a 0.0 to 10.0 point scale. Attribute range descriptions were Outstanding (9.0 to 10.0); Excellent (8.0 to 8.9); Very Good (7.0 to 7.9); Good (6.0 to 6.9); Fair (5.0 to 5.9); and Poor to Very Poor (less than 5.0). These attributes are abstract in nature, as opposed to concrete, technical descriptions (e.g., RAM, ROM, Megahertz, etc.), which is desirable since the subjects are not likely to be experts in the product category (Bettman and Sujan 1987). Subject familiarity with word processing programs and computers is presented in the section on data integrity. The attributes were described on a separate handout that was provided and read at the onset of the experiment, and then retained by subjects throughout the experiment.
One attribute, Versatility, was strongly negatively correlated with the other five, particularly among the 15 non-dominated brands. Since these 15 brands represent the only viable choices, correlations for these brands are presented here. Correlations among Versatility and the other attributes ranged from $r = -.36$ to $-.63$. Only the negative correlation among the three "Ease Of" attributes and Versatility were significant ($p < .05$). All other attribute intercorrelations possessed positive signs and ranged from $r = .05$ to $.91$.

The range of attribute values was fairly high. For example, the low values ranged from 3.0 to 5.2 and the high values from 8.6 to 9.3 for the six attributes. Due to the negative correlations between Versatility and the other attributes, the "higher scoring" brands had fairly low scores on Versatility and five high scores on the other attributes, whereas the "lesser" brands had a high versatility score and low scores on the other attributes. This should create meaningful conflict in choice since tradeoffs must be made. That is, the decision maker cannot "have it all" with any single brand.

**Subjects, Design, And Procedures**

The study compared the three formats using a between-subjects design. Twenty subjects were randomly assigned to RATER and CAL, and 19 to ALPHA, for a total of 59 subjects. Subjects were processed in groups of four, with the formats rotated to mitigate time of day and sequencing (e.g., experimenter learning) effects. Training and instruction scripts were used to maintain across group consistency and the same person led all sessions. Subjects were students drawn from survey level business classes who received bonus points for their participation in the study.

The experiment consisted of three distinct segments that included training, experimental, and brand comparison sessions. The training session was designed to thoroughly familiarize subjects with the format, task, necessary computer operations, attributes used, and the importance weight concepts/questionnaires. Subjects in the RATER and CAL groups were also given the same tutorial on weighted averages and rankings to enhance format understanding. In this segment,
subjects essentially participated in a "dress rehearsal" for the actual experimental task. Database management software programs were used in the training segment, which had been evaluated on the same attributes as the word processing programs.

The experimental (and training) task statement was to "choose the best brand for you based only on the attribute scores provided." Subjects were instructed to ignore considerations stemming from possible past experience and brand name. Further, they were told prices were not provided since "prices for all brands are about the same." The RATER and CAL subjects were also told "You may select the first brand, last brand, or any brand in-between. Do not feel you should take a highly ranked brand, although you certainly may." This was done to avoid having subjects feel they ought to select the top brand. Subjects were instructed to take all the time they needed and reminded that the choice process was entirely in their hands. These instructions should create independent decision makers, who would not use the assisted formats as normative aids. That is, they would not "blindly" select the top ranked brand.

During the experimental session subjects first filled out an importance weight questionnaire. They then engaged in the choice task. Upon completing this task they were given the major questionnaire that included the second importance weight question and the subjective reaction items to the format. After all group members finished the questionnaire, they left the experiment room for approximately five minutes, were monitored in a hallway, and were instructed not to speak with one another. During this time, each subject's experimentally determined best choice (EDBC) was calculated using post-choice weights (as described in the next section). The EDBC was subsequently used as a comparison brand in the following session.

The last session required subjects to make comparisons of the brand they chose with up to three other brands. The first of the three comparison brands was a low scoring brand, which enabled subjects to become comfortable with rejecting a comparison brand. The second comparison brand was the subject's second best choice. Finally, subjects were asked to compare their initial best choice with the EDBC, if it had not been selected as the first or second best choice. They were then asked, "if you had the best choice decision to make again, would you again select
the brand you chose." If not, they were asked to indicate their final choice. This portion of the experiment provided a measure of decision quality based on switching behavior.

**Dependent Measures and Operationalizations**

**Subj ective Reactions and Time.** Subjective reactions to the alternative formats were assessed by asking DM's to compare the format they used to a benchmark format which was described as "providing the same information, but with all 20 brands listed in alphabetical order on a sheet of paper (e.g., a magazine page)." This approach was used so that subjects in all groups would have a common frame of reference against which to compare their format. We assume that ALPHA subject responses would be toward the mid-point of the scale, although some pluses and minuses of a computer display were cited in pretests. For example, positives included the ability to highlight the brands with the cursor bar and the "brightness" of the display, whereas minuses included the inability to write on the screen. Since all formats use the same display characteristics, the only difference among formats is the order brands are presented in. Six questions, asked on a nine point semantic differential scale, measured DM's perceptions of their format on decision accuracy, confusion experienced, frustration experienced, confidence the best choice was made, effect on decision making time, and overall satisfaction with the format. Actual decision making time was measured surreptitiously with a stop watch.

**Decision Ouality.** This study used three methods to measure decision quality. The first two used the experimentally determined best choice (EDBC) to determine: (1) the relative value of the actual choice, and (2) discrete choice quality. The relative decision accuracy rule used here provides a measure of the proportion of value attained from the actual choice relative to that of the EDBC (Johnson and Payne 1985; Klein and Yadav 1989). Accuracy has typically been defined by the utility of the selected alternative divided by the utility of the optimal choice, with the utility of a random choice subtracted from the numerator and denominator; however, since in the current study the data set for all formats were the same, the correction for the value of a random choice is not
necessary; therefore, we simply divide the value of the actual choice by the EDBC. The discrete choice rule measures if the best choice matched the EDBC (cf., Malhotra 1982).

The EDBC was experimentally determined individually for each subject using a linear combination of attribute importance weights and rater provided attribute scores. This approach, or variations of it, has been used by numerous researchers (cf. Keller and Staelin 1987, Klein and Yadav 1989). A simple linear model using importance weights, as opposed to ideal point or expectancy-value models, was employed since the predictor variables have a monotonic relationship to the criterion (cf. Dawes and Corrigan 1974). That is, more of an attribute can be assumed to result in greater preference. The weights used to generate the EDBC were those taken after the choice task. Since weights were stable between the initial (pre-choice) and final (post-choice) assessments, similar results were found using either set of weights (stability results are presented in the next section). The final weights were used since they, in any case, should best reflect the most recent feelings about attribute importance.

No uncontested linear model based decision accuracy rule exists, however, and the results of these analyses need to be viewed cautiously (Meyer and Johnson 1989). Therefore, a third method measured the amount of switching from the actual choice to the EDBC. In this usage the EDBC serves, at the least, as an attractive alternative against which to compare the initial choice. While subjects might switch to brands other than the EDBC if given the opportunity, for this initial study the number of comparison brands was limited to the three already discussed. The switching measure examines the proportion of subjects switching to the EDBC in the comparison session. Since an unchosen EDBC is likely a strong candidate for best choice, it is of interest to discern whether subjects, who may not have fully evaluated the EDBC, switch to it. We argue that not switching is indicative of more certain and reasoned choice, whereas the reverse is the case for brand switchers.

In this study the measure of switching behavior to gauge choice quality was preferred, since it does not depend upon the assumption that the study accurately modeled the EDBC as best choice. Furthermore, it is in the type of environment studied here (negatively correlated) that the prediction
of preferences and, hence, a discrete best choice, is most likely to fail (Johnson, Meyer and Ghose 1989; Meyer and Johnson 1989). Finally, using a linear rule to assess the accuracy of a linear aid is tautological. Therefore, the results of the relative and discrete measures are reported, but the switching measure is relied upon most heavily. It is important to note that the relative measure does have more merit than the discrete measure, since it indicates the value received as a percentage of at least a "good" choice. This should provide some insight into the magnitude of choice quality attained, despite the difficulty in modeling the best choice for each individual.

The final decision quality measure used identifies the number of dominated alternatives selected. The selection of a dominated brand is a clear cut choice error, since another brand is superior on at least one attribute and no worse than tied on the others (Payne, Bettman and Johnson 1988). The correspondence among and validity of these measures will be addressed in the discussion of the results.

**Rater Format Manipulation and Data Integrity Checks**

Subjects' weighting policies were stable as the correlations between the first (pre choice) and second (post choice) sets of weights were high within each format. The correlations for the 18 possible pairs of weights (i.e., six attributes by three formats) ranged from Pearson's $r = .62$ to $.94$, all $p < .001$. The average correlation across the six attributes was $r = .79$ for ALPHA, $r = .80$ for RATER, and $r = .83$ for CAL. Furthermore, subjects were encouraged to use different weights if they felt they had changed during the choice task (in both verbal instructions and those written in the questionnaire); nevertheless, it appears that the intervening choice task did not materially affect weights. This is important to the extent that importance weight stability was attained and sustained throughout the choice task; that is, subjects knew what they desired in a brand.

Research questions concerning the RATER format depend upon the best to worst rank orders for subjects to not be in concordance with the RATER rank order. Providing subjects possessed attribute importance weighting policies that yield highly similar individual rank orders,
the individual rank orders can be averaged to generate a group rank order. The rater group and RATER format rank orders may then be compared to establish their level of agreement.

The appropriate statistic to measure the agreement among raters within a group is Kendall's W. The rater subjects rank orders were almost in complete agreement (W = .96, p < .001). Therefore, it is appropriate to aggregate and use the average group rank order to compare against the RATER format order. Upon comparing the rater group rank order and the RATER format rank order, it was determined that they were not at all in concordance (Spearman's $r = -.12$; Kendall's Tau-B = -.05). The RATER format was generated using a very experienced user's weighting scheme, which weighted the "power" dimensions, Performance (25 percent) and the negatively correlated Versatility attribute (50 percent), heavily, and the other attributes lightly (total of 25 percent). The subjects, however, weighted Versatility lightly (13 percent by both pre and post choice weights, which made Versatility the least important attribute). Therefore, the RATER rank order of brands was not at all in agreement with the subjects' rank orders. For example, the top ranked brand using group weights was ranked number 20 (last) in the RATER format (Exhibit 1).

Following are a series of checks assessing subject familiarity with computers and word processing software and the effectiveness of the training session. Only the grand mean is indicated, because the ANOVA F-test never approached significance (F $\leq 1.0$). All of the questions were measured using a nine point semantic differential scale. The subjects' self-reported familiarity with computers was moderate (mean = 4.2, 1 = extremely familiar, 9 = not at all familiar), with no difference among the three formats (F = .71, p > .10). Subjects were somewhat unfamiliar with word processing programs (mean = 5.7, 1 = extremely familiar, 9 = not at all familiar), with no difference among the groups (F = 1.0, p > .10). Since differential familiarity among the groups might systematically affect outcomes, the no difference finding is reassuring.

Three training effectiveness questions (asked prior to the experimental session) indicated the subjects were well trained. Responses on all items were close to the "fully understood" anchor for each format. Subjects were certain they "understood the computer operations and format description well enough to complete the best choice task" (mean = 2.5, F = .31, p > .10, 1 =
extremely certain, 9 = not at all certain). They also fully understood the technical aspects of operating the computer (mean = 1.8, F = .38, p > .10, 1 = fully understood, 9 = did not at all understand), and the order in which the brands were presented (mean = 1.5, F = 1.0, p > .10, 1 = fully understood).

RESULTS

Decision Quality

Hypothesis 1 specified that the discordant RATER format would lead to lower decision quality than the ALPHA format. Relative decision quality, as measured by subjects' attained proportion of the value of their EDBC's, was significantly lower for the RATER format relative to the ALPHA format. As shown in Table 1, RATER subjects obtain 95.1 percent of their EDBC, while ALPHA subjects obtained 98.5 percent. Tukey's Post Hoc test indicates this difference is significant (p < .05), although the magnitude of the difference is not striking. For a frame of reference, the median brand in a best to worst rank order, generated using equal weights, provided less than 83 percent of the value of the top ranked brand. The discrete accuracy measure indicated that 10 percent of the RATER subjects selected their EDBC, whereas 31.6 percent of ALPHA subjects did. This difference was significant (Z = 1.67, p < .05).

A test comparing the proportion of switchers in the two groups, however, does not support the contention that RATER subjects performed worse than ALPHA subjects (Z = .636, p > .10). The direction of the results, however, did favor ALPHA (21.1 percent switched) over RATER (30 percent switched). Finally, one RATER subject selected a dominated alternative, which is a clear decision error (Payne, Bettman and Johnson 1988), whereas no ALPHA subject selected a dominated alternative. In post-choice questioning, this subject indicated he selected the dominated alternative because the top ranked brands were not to his liking and, therefore, he thought the best brands for him would be toward the bottom of the list.

These results provide mixed implications for Hypothesis 1. The relative and discrete accuracy measures provide some support for the prediction that ALPHA is superior to RATER, but
the test on switching behavior does not. We prefer the test of switching since it does not rely upon our accurately modeling the best choice experimentally. We cannot, therefore, conclude that RATER is clearly inferior to ALPHA. To summarize, while a discordant RATER format is certainly not a good way to present information, decision makers seem to be able to overcome a discordant rank order by performing at a level comparable to unassisted (ALPHA) decision makers.

Hypothesis 2 stated that the decision quality of subjects using a CAL format would be superior to those using the ALPHA and RATER formats. Based on the percentage of EDBC value achieved, the CAL format (98.3 percent) was significantly superior to the RATER format (95.1 percent), but was equivalent to the ALPHA format (98.5 percent). The discrete measure indicated that CAL (40 percent selected the EDBC) was superior to RATER (10 percent selected the EDBC, $Z = 2.19$, $p < .01$), but not to ALPHA (31.6 percent, $Z < 1.0$, $p > .10$). Contrast tests on switching behavior were also conducted using Z-tests for proportions (since no subject switched in the CAL format, one-half subject was counted as having switched). CAL was found to have significantly fewer switchers than RATER (30 percent switched, $Z = 2.36$, $p < .01$) and ALPHA (21.1 percent switched, $Z = 1.81$, $p < .05$).

In response to Hypothesis 2, the CAL format did result in superior decision quality compared to the RATER format. This is evident from the brand switching measure, which we most prefer, as well as the evidence provided by the less preferred relative and discrete choice accuracy measures. Clearly, the CAL format is superior to a discordant ordering. Compared to the ALPHA format, however, it cannot be definitively stated that CAL is superior. The test of switching behavior indicated superior decisions were made with CAL. The relative and discrete measures, however, did not yield a significant difference. This suggests that despite more switching with ALPHA, the value lost by selecting the original choice was not high. The evidence, therefore, leans toward the conclusion that CAL is superior to ALPHA. It is important to remember, furthermore, that the analysis was biased against CAL, which had one-half subject counted against it when, in actuality, none did switch.
Subjective Reactions and Decision Times

Research Question 1 asked whether the computer assisted format was evaluated more highly than the two traditional formats. Based on the results shown in Table 2, the CAL format was more highly evaluated on the subjective reactions than ALPHA and RATER. First, the direction of preferences is in favor of the CAL format on every subjective reaction question. Second, the CAL format is statistically significantly superior (Tukey's HSD Post Hoc test, $p < .05$) on all five questions when compared to the ALPHA format, and on two questions compared to the RATER format. Third, when responses to the five subjective reaction questions are summed to provide an overall measure of format performance (coefficient alpha on the five items $= .90$), subjects using the CAL format evaluated it as being significantly superior to both ALPHA and RATER (Tukey's HSD, $p < .05$).

With regard to the decision quality items (decision accuracy and confidence that the best choice was made), the CAL was perceived as being significantly superior to both RATER and ALPHA. For the questions about difficulty experienced with the choice process (frustration and confusion) and format satisfaction, however, CAL was perceived as being significantly superior to ALPHA, but only nominally better than RATER. No difference existed between the ALPHA and RATER formats on any measure (Tukey's HSD, $p > .10$). ALPHA was consistently evaluated, however, as being nominally worse than RATER.

Research Question 2 asked whether there were differences in decision making time among the formats (Table 2). The difference in decision time between the ALPHA (168 seconds) and RATER (183 seconds) formats was not significantly different ($p > .10$); however, both were significantly less than the CAL format (244 seconds, $p < .05$, Tukey's HSD post-hoc test). Perceived decision making time corresponded to the actual time results. CAL was perceived to take significantly more time than ALPHA (Tukey's HSD < .05) and marginally significantly more time than RATER (Tukey's HSD < .10). The time taken to enter the importance weights and execute the CAL format is likely responsible for the higher actual and perceived decision making time in this format. While the CAL format apparently takes longer with which to make a decision, this does not
appear to be due to any additional difficulty experienced during decision making. Indeed, as observed above, it appears that DMs experienced similar or less difficulty when compared to DMs using the other formats.

The final set of questions asked subjects to indicate the number of brands they "thoroughly evaluated", "found to be acceptable," and "seriously considered as best choice candidates" (Table 1). Although no significant differences existed among the formats on any of these three items, several observations on the number of brands thoroughly evaluated is worth noting. ALPHA subjects reported thoroughly evaluating 9.8 brands, RATER 6.5, and CAL 7.5. Importantly, the RATER and CAL subjects did not simply select the top rated brand, which is consistent with the independent decision making perspective this study intended for the subjects. It is of interest to note, however, that the CAL and RATER subjects thoroughly evaluated nominally fewer brands than ALPHA subjects. Indeed, ALPHA subjects thoroughly reviewed, on average, 50 percent more brand than did RATER subjects. This suggests that the RATER format was, perhaps, somewhat influential in limiting the number of brands evaluated. Given the lack of significant differences among the formats on this item, however, a more definitive conclusion cannot be drawn from this finding.

**DISCUSSION**

Based on the subjective reaction questions, evidence was provided that indicated decision makers preferred the CAL format over the traditional formats (ALPHA and RATER). In addition, CAL subjects made very confident and certain choices based on the fact that no subject switched in the comparison session. The relative accuracy measure also indicated that subjects gained a great deal of the value of, at the least, a "good" choice. The major drawback to CAL was that subjects took significantly longer to reach a decision than those in the passive formats, but that was likely due to the time taken to enter weights and execute the procedure and not added thinking time about the choice.
Although the CAL format appears to be a sound method in which to present ratings, only eight CAL subjects selected their EDBC and only nine selected the actual top ranked brand generated by the linear format during the choice task (this was recorded during the experiment). The number of subjects selecting the EDBC and top ranked brand (during choice) differs slightly due to an importance weight change by one subject. Therefore, subjects did not automatically select the top ranked brand, as over half preferred brands ranked from second to seventh on their choice task rank orders. It seems clear that information providers ought to present more than just the top few brands when a CAL format is used. Further, it does not seem that the CAL format is used by subjects as a normative decision tool (i.e., to indicate the actual choice). Hence, it appears to be preferred by DM's because it helps organize information.

A related issue is the poor discrete decision accuracy of the linear model in predicting choice in all formats. Indeed, as just noted, subjects provided with a linear model failed to select the top rated brand on the screen display 55 percent of the time. Further, subjects did not select the top rated brand 60 percent of the time when post-choice weights are used to indicate the best brand (Table 1). This reinforces the Johnson, Meyer and Ghose (1989) contention that linear predictability is lessened, for discrete choice, in negatively correlated environments.

Our findings clearly indicate that linear models should not be positioned as normative aids, especially in difficult decision environments such as the one used here. Otherwise, decision makers might be induced to simply select the top ranked brand; instead, they ought to be encouraged to evaluate many or even all efficient (i.e., non dominated) brands. In the same vein, information providers should not be tempted to display only one or a few of the top ranked brands. This has been shown in this study to be inappropriate. While it is not likely that a DM would make a poor choice by simply selecting the top ranked brand, not only is a better choice possibly available further down the list, but manufacturers could be unfairly biased against by having their possibly more preferred brand inadvertently overlooked.

ALPHA subjects performed reasonably well on the decision quality measures and made the quickest decisions, but scored the format lower on the subjective evaluation measures. Relative
decision quality was not significantly lower than the CAL format, although they did switch at significantly greater levels. Overall, the results for the unaided subjects might be considered to be acceptable, despite high task and context effects; however, the CAL was, on balance, superior.

The study provided evidence indicating that those using the RATER format made inferior decisions relative to CAL format users. However, we could not conclude that it was clearly inferior to ALPHA. Providing decision makers do not blindly follow rater recommendations, they seem to cope reasonably well with a discordant rank order. Indeed, not one subject selected the top rated brand, which would be rated last according to group weights. Further, this potential problem exists only in situations such as the one used in this study. That is, where strong negative attribute correlations exist among the attributes and when the weights used by the information provider are opposite to those of the DM (Curry and Faulds 1986). In summary, even when this occurs, evidence has been provided indicating that decision makers seem able to overcome the influence of a discordant ordering, at least relative to an alphabetical listing. Of course, totally dependent DM's who blindly select the top rated brand would make a poor choice when choosing from a decision set such as that used here.

CONCLUSION

The results of this research should be viewed as representing an initial step in the study of formats offered over buyer information systems. Although a number of issues have been addressed, the results raise a series of other questions for further research. With that in mind several limitations should be cited, which also serve as recommendations for future research. First, future research ought to be carried out by further manipulating the decision environment. Specifically, using uncorrelated sets and varying the degree of negative correlations in the decision sets would more fully document the relative efficacy of the formats. In a similar fashion, varying task complexity (e.g., number of brands, attributes and/or time pressure) might also provide important information on the comparative value of the formats.
Second, a comparison of the formats used in this study to other computer assisted formats, such as the cutoff or disjunctive formats, would prove beneficial. The cutoff model in particular seems important to include in a comparative study, as it is being offered by several videotex information providers. Fourth, an expansion of the number of brands used in the comparison session would provide for a more complete measure of switching behavior. This could more fully identify possible errors, since subjects might switch to brands other than the EDBC. Fifth, alternative formats should be studied across various market segments. For certain groups of consumers (e.g., high versus low knowledge) or in household versus organizational buying situations, format preferences and performances might vary. Likewise, the formats should be evaluated across different product categories. Certain products and types of attributes (e.g., discrete attributes such as features versus continuous attributes such as ratings) might better lend themselves to one format versus another.

Sixth, it is also of interest to study situations where consumers have been trained to use multiple computer assisted formats (e.g., CAL and cutoff formats). By observing choice behavior and preferences using a within subject design, a direct comparison on the dependent measures can be attained. Finally, it is of key importance to test the effects of training and cautions when a format might adversely interact with a decision environment. For example, training and warning about negative correlations among the attributes might prevent potential problems from arising.

As more consumers have access to and interest in electronic information services, the effect of formats on data acquisition and decision making becomes more important; however, numerous decisions for information providers also arise. For example, does one format "fit all" or should consumers be able to select from multiple formats? Should more emphasis be placed on consumers' abilities to make the best decisions or on their willingness to use and feel comfortable about using consumer information systems? Should consumer information systems be viewed as a replacement for, or as an aid to facilitate, the buyer decision making process? What training or other actions should buyer information systems provide to users. It can be concluded from this research, however, that the computer assisted linear format appears promising in enhancing decision quality,
and is highly regarded by users. We recommend that product raters not offering a CAL type aid begin offering one as an option to information seekers using videotex systems.
REFERENCES


**TABLE 1**

**Decision Quality of the ALPHA, RATER and CAL Formats**

<table>
<thead>
<tr>
<th></th>
<th>ALPHA n=19</th>
<th>RATER n=20</th>
<th>CAL n=20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Accuracy(^1)</td>
<td>98.5% a</td>
<td>95.1% ab</td>
<td>98.3% b</td>
</tr>
<tr>
<td>Discrete Accuracy(^2)</td>
<td>31.6% (6) b</td>
<td>10.0% (2)ab</td>
<td>40.0% (8)a</td>
</tr>
<tr>
<td>Switching to the Comparison Brand(^3)</td>
<td>21.1% (4) b</td>
<td>30.0% (6) a</td>
<td>2.5% (1/2) ab</td>
</tr>
<tr>
<td>Number of Times the Top Ranked Brand on the CAL formats and RATER was selected as the best choice</td>
<td>N/A</td>
<td>0</td>
<td>45.0% (9)</td>
</tr>
<tr>
<td>Selection of a Dominated Alternative</td>
<td>0</td>
<td>5.0% (1)</td>
<td>0</td>
</tr>
<tr>
<td>Number of brands(^4) thoroughly evaluated</td>
<td>9.84</td>
<td>6.45</td>
<td>7.50</td>
</tr>
<tr>
<td>Number of brands found(^4) acceptable</td>
<td>5.00</td>
<td>5.10</td>
<td>4.65</td>
</tr>
<tr>
<td>Number of brands seriously considered as a best choice</td>
<td>2.95</td>
<td>2.90</td>
<td>4.15</td>
</tr>
</tbody>
</table>

1 Relative accuracy indicates the percentage of EDBC obtained for each format. Formats sharing the same letter are statistically different (p < .05) using Tukey's HSD post-hoc test.

2 Number of times the actual choice matched the EDBC best choice. Formats sharing the letter "a" were significantly different (p < .05) using a 2-Tail Z-Test. Formats sharing the letter "b" were marginally significantly different (p < .10). Note that 1/2 subject was added to CAL to carry out the contrasts.

3 Formats sharing the letter "a" were significantly different (p < .05) using a 2-Tail Z-test. Formats sharing the letter "b" were marginally significantly different (p < .10). Note that 1/2 subjects was added to CAL to carry out the contrasts.

4 There were no significant differences among the formats on this item using Tukey's HSD post-hoc test (p>.10).
# TABLE 2
Subjective Reactions and Decision Time

<table>
<thead>
<tr>
<th></th>
<th>Format Means¹, ², ³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALPHA (n = 19)</td>
</tr>
<tr>
<td>SUBJECTIVE REACTIONS</td>
<td></td>
</tr>
<tr>
<td>Decision Accuracy</td>
<td>5.74 a</td>
</tr>
<tr>
<td>(1 = greatly worsened, 9 = greatly improved)</td>
<td></td>
</tr>
<tr>
<td>Confusion Experienced</td>
<td>5.63 a</td>
</tr>
<tr>
<td>(1 = greatly increased, 9 = greatly decreased)</td>
<td></td>
</tr>
<tr>
<td>Frustration Experienced</td>
<td>5.37 a</td>
</tr>
<tr>
<td>(1 = greatly increased, 9 = greatly decreased)</td>
<td></td>
</tr>
<tr>
<td>Confidence Best Choice Made</td>
<td>5.37 a</td>
</tr>
<tr>
<td>(1 = greatly decreased, 9 = greatly increased)</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with Format</td>
<td>5.90 a</td>
</tr>
<tr>
<td>(1 = greatly decreased, 9 = greatly increased)</td>
<td></td>
</tr>
<tr>
<td>Composite Score</td>
<td>28.01 a</td>
</tr>
<tr>
<td>(5 item coefficient alpha = .90)</td>
<td></td>
</tr>
<tr>
<td>Decision Time (sec)</td>
<td>168 a</td>
</tr>
<tr>
<td>Perceived Decision Time</td>
<td>5.58c</td>
</tr>
<tr>
<td>(1 = Greatly Increased, 9 = Greatly Reduced)</td>
<td></td>
</tr>
</tbody>
</table>

1 Questions were slightly modified/reduced to fit in the Table.

2 Means sharing the letters "a" and "b" are significantly different at p < .05 using Tukey's HSD post hoc test.

3 Means sharing the letter "c" and significantly different at p > .05 and p < .10 using Tukey's HSD post hoc test.
Exhibit 1

EXAMPLE OF THE CAL FORMAT USING GROUP WEIGHTS

Requirement Specification

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Importance</th>
<th>SOFTWARE REVIEW RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EASE OF START-UP</td>
<td>13.4%</td>
<td>Press 'Alt' and 'S' to Search</td>
</tr>
<tr>
<td>EASE OF LEARNING</td>
<td>18.7%</td>
<td>Press Down Cursor to View Entire List</td>
</tr>
<tr>
<td>EASE OF USE</td>
<td>21.4%</td>
<td></td>
</tr>
<tr>
<td>ERROR HANDLING</td>
<td>17.3%</td>
<td></td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>16.3%</td>
<td></td>
</tr>
<tr>
<td>VERSATILITY</td>
<td>12.9%</td>
<td></td>
</tr>
</tbody>
</table>

100.0% < MAKE SURE THAT THIS EQUALS 100%

<table>
<thead>
<tr>
<th>Product Name</th>
<th>Ease Start</th>
<th>Ease Learn</th>
<th>Ease Use</th>
<th>Error Hand</th>
<th>Perf</th>
<th>Versa</th>
<th>RATER Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Einstein Writer</td>
<td>9.0</td>
<td>9.0</td>
<td>8.5</td>
<td>7.8</td>
<td>8.2</td>
<td>5.5</td>
<td>6</td>
</tr>
<tr>
<td>b. Multimate</td>
<td>8.9</td>
<td>9.3</td>
<td>8.5</td>
<td>7.6</td>
<td>6.8</td>
<td>6.0</td>
<td>8</td>
</tr>
<tr>
<td>c. Samna Word II</td>
<td>8.5</td>
<td>9.0</td>
<td>9.1</td>
<td>7.5</td>
<td>8.6</td>
<td>3.5</td>
<td>15</td>
</tr>
<tr>
<td>d. Leading Edge</td>
<td>8.9</td>
<td>9.0</td>
<td>8.7</td>
<td>8.1</td>
<td>6.8</td>
<td>5.0</td>
<td>11</td>
</tr>
<tr>
<td>e. WordPerfect</td>
<td>6.3</td>
<td>7.0</td>
<td>8.6</td>
<td>7.8</td>
<td>7.6</td>
<td>6.5</td>
<td>2</td>
</tr>
<tr>
<td>f. EasyWriter</td>
<td>8.0</td>
<td>7.3</td>
<td>8.2</td>
<td>8.0</td>
<td>7.3</td>
<td>5.0</td>
<td>12</td>
</tr>
<tr>
<td>g. Super-Text</td>
<td>6.9</td>
<td>7.7</td>
<td>8.0</td>
<td>6.1</td>
<td>7.7</td>
<td>6.5</td>
<td>4</td>
</tr>
<tr>
<td>h. EasyWriter</td>
<td>5.9</td>
<td>6.7</td>
<td>7.7</td>
<td>7.5</td>
<td>7.7</td>
<td>6.0</td>
<td>7</td>
</tr>
<tr>
<td>i. Pfs:Write</td>
<td>8.3</td>
<td>5.0</td>
<td>6.8</td>
<td>7.5</td>
<td>7.5</td>
<td>6.0</td>
<td>9</td>
</tr>
<tr>
<td>j. Volkswriter Deluxe</td>
<td>7.1</td>
<td>5.7</td>
<td>6.8</td>
<td>7.5</td>
<td>7.5</td>
<td>5.0</td>
<td>14</td>
</tr>
<tr>
<td>k. Office Writer</td>
<td>7.2</td>
<td>5.7</td>
<td>6.8</td>
<td>7.5</td>
<td>7.5</td>
<td>5.0</td>
<td>18</td>
</tr>
<tr>
<td>l. Electric Pencil PC</td>
<td>6.1</td>
<td>5.0</td>
<td>6.8</td>
<td>8.1</td>
<td>7.6</td>
<td>4.0</td>
<td>17</td>
</tr>
<tr>
<td>m. VisoWord Plus</td>
<td>6.6</td>
<td>5.7</td>
<td>6.6</td>
<td>7.2</td>
<td>6.5</td>
<td>5.0</td>
<td>16</td>
</tr>
<tr>
<td>n. WordVision</td>
<td>7.2</td>
<td>4.7</td>
<td>5.7</td>
<td>6.7</td>
<td>6.1</td>
<td>8.0</td>
<td>3</td>
</tr>
<tr>
<td>o. Qwerty</td>
<td>8.0</td>
<td>6.0</td>
<td>6.2</td>
<td>7.0</td>
<td>6.2</td>
<td>3.0</td>
<td>19</td>
</tr>
<tr>
<td>p. Select</td>
<td>7.0</td>
<td>5.3</td>
<td>5.2</td>
<td>7.0</td>
<td>5.7</td>
<td>7.0</td>
<td>10</td>
</tr>
<tr>
<td>q. Palantir</td>
<td>5.0</td>
<td>4.3</td>
<td>6.6</td>
<td>7.4</td>
<td>7.1</td>
<td>5.5</td>
<td>13</td>
</tr>
<tr>
<td>r. Peach Text</td>
<td>9.3</td>
<td>4.3</td>
<td>5.7</td>
<td>6.9</td>
<td>7.0</td>
<td>7.5</td>
<td>5</td>
</tr>
<tr>
<td>s. Textra</td>
<td>9.3</td>
<td>4.3</td>
<td>5.3</td>
<td>5.3</td>
<td>5.2</td>
<td>4.0</td>
<td>20</td>
</tr>
<tr>
<td>t. Superwriter</td>
<td>4.3</td>
<td>2.3</td>
<td>4.9</td>
<td>5.2</td>
<td>6.1</td>
<td>9.0</td>
<td>1</td>
</tr>
</tbody>
</table>

1. The importance weights used in this example are the across-group, grand mean. The rank order of brands for the three groups, generated from each groups weights, were virtually identical.

2. The RATER rank order column was added here for comparison purposes.