Reference Price and Price Perceptions: A Comparison of Alternative Models

RONALD W. NIEDRICH
SUBHASH SHARMA
DOUGLAS H. WEDELL*

Reference price effects on consumer price perceptions are often explained by Helson's adaptation-level theory, in which the cognitive representation of reference price is the prototype of the relevant category. However, recent conceptualizations and empirical evidence suggest the possibility of an exemplar model, which may be specified using Volkmann's range theory or Parducci's range-frequency theory. In two experiments, these three contextual models of reference price effects are pitted against one another. Based on the MANOVA and model fitting, range-frequency theory accounted for reference price effects that the other theories could not, suggesting that consumers compare the target price against specific members of the category rather than the category prototype. A third experiment demonstrated that range and frequency effects are moderated by the stimulus presentation condition, suggesting that consumers place greater weight on extreme prices anchoring the range for internal reference prices than for external reference prices.

Price evaluations are important inputs to consumer decisions such as what, when, where, and how much to buy (Alba et al. 1994; Gupta 1988). One major approach to understanding how consumers arrive at such price evaluations involves the concept of a reference price (Blattberg, Briesch, and Fox 1995; Kalyanaram and Winer 1995), which can be defined as the price against which buyers compare the offered price of a product or service (Monroe 1990). Such comparisons presumably dictate whether a price is deemed too high or too low. Although the concept of reference price is well established, the nature of the reference price comparison process is less understood. In this article, we explore three different approaches to reference price comparison and evaluate the ability of each approach to predict the attractiveness ratings of prices in laboratory experiments.

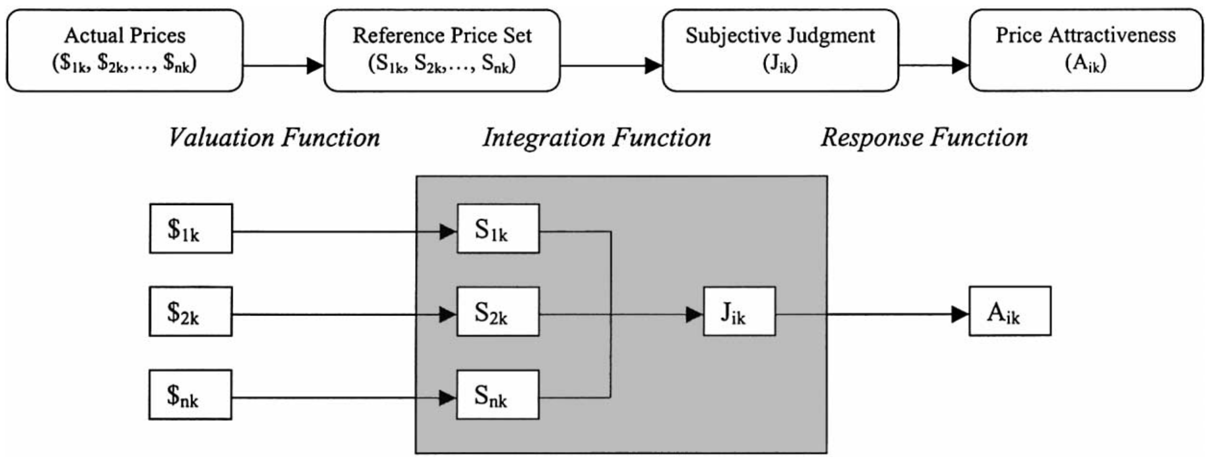
The most widely cited explanation of reference price and price judgment is based on Helson's (1947, 1964) adaptation-level theory, which asserts that judgments are proportional to deviations from a comparison standard. This standard, or adaptation level, is context sensitive, as it is conceived as the mean of the stimuli presented within a contextual set (Helson 1964; Wedell 1995). Following this view, the reference price that consumers use to evaluate a product is a weighted average of the prices from the relevant category (Monroe 1990). The adaptation-level model of reference price comparison is consistent with a prototype representation of categories in which a single prototypical value is abstracted from category instances and is used to represent the category (Medin, Altom, and Murphy 1984).

An alternative to a prototype representation of categories is provided by exemplar models, which posit that the category is represented by a distribution of its instances. In the psychological literature, exemplar models have consistently outperformed prototype models in empirical comparisons (Ashby and Maddox 1998). Exemplar models assume that judgments are based on comparisons to specific category members rather than a comparison to summary information about a typical member (Medin et al. 1984). The exemplar model of categorization is consistent with a conceptualization of reference price as a range of prices or the latitude of acceptable prices rather than a single summary price (e.g., Janiszewski and Lichtenstein 1999; Kalyanaram and Little 1994; Lichtenstein and Bearden 1989; Monroe 1971). However, with the exception of Janiszewski and Lichtenstein (1999), the prototype model has typically been used to explain price evaluations in such experiments.

*Ronald W. Niedrich is assistant professor of marketing, E. J. Ourso College of Business Administration, Louisiana State University, Baton Rouge, LA 70803 (niedrich@lsu.edu). Subhash Sharma is professor of marketing and distinguished foundation fellow, the Moore School of Business, University of South Carolina, Columbia, SC 29208 (sharma@sc.edu). Douglas H. Wedell is professor of psychology, University of South Carolina, Columbia, SC 29208 (wedell@sc.edu). This article is based on the first author's University of South Carolina dissertation. The authors would like to thank two additional dissertation committee members, Terence A. Shimp and William O. Bearden, as well as Richard G. Netemeyer, the editor, the associate editor, and three reviewers for their significant comments throughout the many stages of this article.

FIGURE 1

CONCEPTUAL FRAMEWORK



NOTE.—The price attractiveness rating of brand price *i* is measured given a set of *n* prices in context *k*.

In a series of four experiments, Janiszewski and Lichtenstein (1999) compared Volkmann's (1951) range theory to Helson's adaptation-level theory. Range theory may be classified as an exemplar model in which the cognitive representation is assumed to include only the highest and lowest values in the contextual set. In the main experiments, subjects were shown sets of prices that manipulated the range while holding the mean of the sets constant. Subjects were then asked to rate the attractiveness of the mean price. There were significant differences between the attractiveness ratings of the mean price in these skewed price sets. Range theory correctly predicted the direction of these differences. On the other hand, adaptation-level theory cannot account for differences in price judgments in sets with the same mean. Based on these data, analyzed in an ANOVA framework, Janiszewski and Lichtenstein concluded that range theory is more consistent with the data.

The experiments reported here extend the work of Janiszewski and Lichtenstein (1999) in a number of important ways. First, we attempt to overcome the limitations of testing contextual models using only the ANOVA framework. The contextual models make specific predictions of judgments given a particular contextual set. The ANOVA tests for differences in ratings across experimental conditions and not for differences between ratings and model predictions. Thus, the ANOVA may support a model that does not fit the data. In an attempt to overcome this limitation, data analyses include both the ANOVA and fits of specific models. Second, different models may fit the data better in different areas of the distribution. Rather than evaluate only the mean, we had participants evaluate several target stimuli spaced at different intervals along the range. Third, Janiszewski and Lichtenstein (1999) compared ratings in contextual sets in which the mean price was held constant and the range was varied. However, price judgments can be affected by frequency manipulations (Alba et al. 1994; Alba et al. 1999; Kalwani and Yim 1992). Therefore, we also investigated the complementary condition, comparing ratings across contextual sets in which the range was held constant and the mean was varied. Finally, in addition to adaptation-level theory and range theory, we also test Parducci's (1965) range-frequency theory. Range-frequency theory is an exemplar model in which the cognitive representation is assumed to include all prices in the contextual set. Work in psychological judgment has typically found range-frequency theory to provide better fits to psychophysical data than either range theory or adaptation-level theory (Birnbaum 1974; Parducci and Perrett 1971). Thus, we test the extent to which reference price effects could be predicted by range-frequency theory.

In sum, the purpose of this research is to provide a better understanding of the nature of reference prices and of how reference prices are used by consumers to evaluate the focal brand price. Consistent with the view of methodological pluralism, we combined the use of the ANOVA and model fitting to provide a more rigorous analysis of the data. Theoretical predictions were tested in a series of three experiments. Experiment 1 extended the work of Janiszewski and Lichtenstein (1999) by addressing the limitations previously discussed. Experiment 2 tested whether a prototype model or an exemplar model better accounted for the data. Finally, experiment 3 tested the moderating effect of stimulus presentation on the relative applicability of the models.

CONCEPTUAL FRAMEWORK

The conceptual framework, shown in figure 1, is based on Anderson's (1981) information integration model. In the model, an overt rating, A_{ik}, is measured in response to contextual stimuli. To evaluate brand price *i*, consumers utilize an appropriate evaluation context *k* (Herr 1989). The set of *n* actual prices of the brand in context *k* are the set of prices to which the consumer has been exposed, denoted $\$_{1k}$, $\$_{2k}$, . . . , $\$_{nk}$. The focal brand price, $\$_{ik}$, is the price of the brand currently under consideration and is a member of the set of actual prices. Three functional relationships are specified in Anderson's (1981) information integration model: the valuation function, the integration function, and the response function. Each of the three functions is discussed in turn.

The Valuation Function

In information integration theory (Anderson 1981), the relationship between the physical stimulus ($\$_{ik}$) and the subjective value of that stimulus (S_{ik}) is called the valuation function. Parducci does not define the specific valuation function within the range-frequency model. In contrast, Helson's (1964) adaptation-level theory includes the psychophysical function as part of the model. Specifically, adaptation-level theory is based on Fechner's law (1860), in which the sensation scale is a logarithmic function of the stimulus values. However, requiring the valuation function to be logarithmic is overly constraining and likely to be false. For example, both Birnbaum (1974), using numerical stimuli, and Wedell (1996), with squares, provide convincing evidence that valuation is better conceived as a power function. Furthermore, this approach has often been taken in modeling the decreasing marginal value of monetary gains and losses (Kahneman and Tversky 1979). Thus, a power function is used to fit the data. Its form is given as

$$S_i = \$_i^r, \qquad \textbf{(1)}$$

where *r* is the value of the power exponent. It should be noted that all models were fit in two ways: estimating *r* and constraining *r* as equal to one. Allowing *r* to be a free parameter improves the model fit somewhat. However, as the improvement is essentially the same for all models, there is no difference in the conclusions using either procedure.

The Integration Function

The relationship between the subjective values (S_{ik}) and subjective judgments (J_{ik}) is defined as the integration function (Anderson 1981). Helson's adaptation-level model, Parducci's range-frequency model, and Volkmann's range model primarily differ in the integration function they specify.

Adaptation-Level Theory. In adaptation-level theory (Helson 1947, 1964), stimulus values are judged within a frame of reference. Equation 2 shows one version of Helson's model (Helson 1964; Marks and Algom 1998; Wedell 1995). This is a linear model, in which the constants *a* and *b* represent the intercept and slope, respectively. In this model, the adaptation level ($S_{al,\,k}$) is the arithmetic mean of all subjective values, and S_{ik} is the subjective value of stimulus *i* in context *k*.

$$J_{ik} = a + b(S_{ik} - S_{al,\,k}). \qquad \textbf{(2)}$$

It should be noted that the adaptation level is often described as the geometric mean rather than the arithmetic mean (Helson 1964; Marks and Algom 1998). This conceptualization is based on the idea that the value function is logarithmic, where the arithmetic mean of the subjective values (S_{ik}'s) is equivalent to the geometric mean of the physical values ($\$_{ik}$'s). To isolate the separate effects of the integration functions tested here, the fits of the adaptation-level model are based on the arithmetic mean. As previously discussed, the valuation function employed is provided in equation 1.

Range Theory. Volkmann's (1951) range theory is based on a range principle of judgment in which the judged value is based on the proportion of the contextual range lying below the stimulus value. This range principle asserts that equal segments of the psychological judgment scale are assigned to equal segments of the contextual range. That is, the judgment of a target within a set of stimuli is linearly related to the end points that anchor the subjective range. Thus, the subjective judgment (*J*) of stimulus *i* in context *k* is given by the proportion shown in equation 3:

$$J_{ik} = (S_{ik} - S_{\min,\,k})/(S_{\max,\,k} - S_{\min,\,k}), \qquad \textbf{(3)}$$

where S_{ik} is the subjective value of stimulus *i*, $S_{\min,\,k}$ is the minimum subjective value, and $S_{\max,\,k}$ is the maximum subjective value in context *k* (Wedell, Parducci, and Lane 1990). For example, if a stimulus lies three-fourths of the way toward the maximum stimulus, the range value of the stimulus is .75. Thus, in range theory the judgment of a stimulus is dependent only on its relationship to the minimum and maximum contextual values.

Range-Frequency Theory. Range-frequency theory (Parducci 1965, 1995) asserts that the judged value of a stimulus is determined by its location within the distribution of contextual stimuli that are brought to mind at the time of judgment. Two contextual principles determine the judged value of the stimulus on a single dimension. According to the "range principle," judgments reflect the location of the target stimulus relative to the most extreme values defining the relevant context. According to the "frequency principle," the location of the target stimulus is described by its rank within the contextual set of stimuli. As shown in equation 4, the subjective judgment (*J*) of stimulus *i* in context *k* is conceived as a compromise between the range and frequency principles:

$$J_{ik} = (w)R_{ik} + (1 - w)F_{ik}. \qquad \textbf{(4)}$$

This subjective value is represented by a weighted average of the range (*R*) and frequency (*F*) values, where the weighting parameter, w, may approach its theoretical limits of zero and one. Note that when $w = 1$, equation 4 reduces to equation 3. Thus, range theory is nested within range-frequency theory. Although a number of experimental conditions have been shown to affect w (Parducci and Marshall 1961; Parducci and Wedell 1986), a value closer to .5 is more typical, representing a roughly equal compromise between the range and frequency principles (Wedell et al. 1990).

The frequency principle has been shown to account for distributional effects, such as skewing (Parducci 1995). The frequency principle asserts that equal segments of the psychological scale are assigned to the same number of cognitive representations in the contextual set. That is, the judgment of a target within a set of stimuli is proportional to the number of representations falling below the target stimulus. Thus, the frequency value of stimulus i in context k, F_{ik}, is given by the proportion shown in equation 5:

$$F_{ik} = (\text{Rank}_{ik} - 1)/(N_k - 1), \quad \textbf{(5)}$$

where Rank_{ik} is the rank of stimulus i in context k, one is the minimum rank, and N_k is the total number of contextual stimuli (Wedell et al. 1990). In the special case in which $w = 0$ in equation 4, range-frequency theory reduces to the frequency principle. Because the contextual range is not included in equation 5, the frequency principle cannot explain effects resulting from the manipulation of the contextual range with ranks held constant.

The Response Function

Finally, the relationship between the subjective judgment (J_{ik}) and the overt rating (A_{ik}) is called the response function. As the range values (eq. 3) and the frequency values (eq. 5) are proportions, they are assigned a value between zero and one. Consequently, the subjective judgments (eq. 4) take on values between zero and one. Assuming a linear relationship between judgments and mean ratings, equation 6 can be used to transform subjective judgments into attractiveness ratings of stimulus i in context k, A_{ik}, where $A_{\min}$ is the lowest scale value and A_n is the number of scale points (Parducci 1995; Wedell et al. 1990). To evaluate the model fit, transformed judgments can be compared with category ratings. By holding the response function constant, we are better able to compare how predictions differ between the contextual models:

$$A_{ik} = (A_n - 1)J_{ik} + A_{\min}. \quad \textbf{(6)}$$

Hypotheses

The primary purpose of experiments 1 and 2 is to test the competing models defined in equations 2, 3, and 4. As each theory assumes a different cognitive representation of reference price, three competing conceptualizations are offered:

H1 (ALT): According to adaptation-level theory, consumers compare the target price against the mean of the contextual set of prices.

H1 (RT): According to range theory, consumers compare the target price against the two prices that define the range in the contextual set.

H1 (RFT): According to range-frequency theory, consumers compare the target price against all of the prices in the contextual set.

Given an exemplar representation of reference price, all prices may not be equally weighted. The weight given to prices in the reference price set is likely to be moderated by the processing environment. If consumers only use the prices available to them in the external environment, as with external reference prices, the price judgment is stimulus based. In contrast, if consumers must retrieve prices from memory, as with internal reference prices, the price judgment is memory based. Because extreme values are comparatively distinct, these values should be more salient than other stimuli in the context (Fiske and Taylor 1991). As the end points are expected to be more easily retrieved from memory (Tversky and Kahneman 1973), they are more likely to be used in a memory-based judgment task. In contrast, less salient intermediate values are less likely to be retrieved and used in a memory-based judgment task. As a result, consumers are expected to weigh intermediate prices more heavily when using external reference prices than when using internal reference prices. Thus, experiment 3 tests two related hypotheses:

H2: Frequency effects will be larger for stimulus-based price judgments than for memory-based price judgments.

H3: Range effects will be larger for memory-based price judgments than for stimulus-based price judgments.

EXPERIMENT 1

Experiment 1 contains two distinct tests of the three contextual theories. The first test extends the work of Janiszewski and Lichtenstein (1999) in the following ways: first, in addition to adaptation-level theory and range theory, we also test range-frequency theory. Second, in addition to an ANOVA, we use model fitting to test the contextual theories. Third, we study price perceptions for a product in an electronic shopping environment rather than studying grocery items with paper and pencil measures. Finally, by measuring attractiveness ratings of several prices in the contextual set, and not just the mean price, we evaluate the shape of the response function. The second test, not evaluated by Janiszewski and Lichtenstein (1999), provides a complementary test to the first. Whereas test 1 manipulates the range in distributions with the same mean price, test 2 manipulates the mean in distributions with the same range of prices. After a discussion of the experimental methods, each test will be addressed separately.

Methods

Subjects. One hundred and ninety-four students received extra credit to participate in the experiment. Of these subjects, 62% were female. The mean in age was 22 years, with a standard deviation of 3.3 years and a range of 19 to 47 years. Insofar as the purpose of this research is to test competing explanations of psychological processes, data collected from college students is appropriate (Mook 1983).

Design. The experimental design is a single between-subjects factor with three levels of price distribution (negative-skew low-mean, positive-skew low-mean, and negative-skew high-mean). Table 1 provides the number of prices at each discount level and the mean discount of the price set. Although price discounts are discussed in order to provide a more managerially relevant dialogue, the experimental stimuli are prices, and therefore the designation of distribution skew is with respect to prices. Subjects evaluated prices at 20%, 25%, 30%, 35%, and 40% discounts after they evaluated the 20 prices shown in table 1. The first 20 prices were used to establish the context. The last five price judgments were used in the MANOVA and model fitting. The presentation order of the first 20 contextual prices was randomized for each subject to mitigate order effects. Similarly, the presentation order of the last five target prices was randomized.

Procedure. The experiment required a manipulation of internal reference price. This was accomplished by exposing subjects to brand prices over time. Consistent with the methods used by Wedell et al. (1990) and Kalwani and Yim (1992), subjects were exposed to a series of stimuli (brand prices) presented sequentially on a computer screen. These prices were intended to simulate price evaluations over a number of purchase occasions. Each screen was intended to represent one purchase occasion. After exposure to each brand price, subjects were required to rate the attractiveness of that price.

The discount sizes were established based on previous research in which discounts ranged from 10% to 40% (Kalwani and Yim 1992), and from 10% to 60% (Janiszewski and Lichtenstein 1999). To allow for enough price levels to manipulate discount distributions, the current research used 13 discount levels from 0% to 60% in increments of 5%. The 5% increment is consistent with the 6% "just noticeable difference" reported by Kalwani and Yim (1992, p. 96), and with the 5% promotion definition employed by Rao, Arjunji, and Murthi (1995). Twenty-five purchase occasions were selected in order to provide enough prices to manipulate the discount patterns, yet not so many as to cause subject fatigue.

TABLE 1

FREQUENCY DISTRIBUTIONS OF EXPERIMENTAL STIMULI

Discount (%)	NSLM	PSLM	NSHM	Normal	Bimodal
0		1	8	1	1
5		1	5	(1)	2 (1)
10		1	1	2	5
15		1	1	1 (1)	2 (1)
20	8 (1)	1 (1)	1 (1)	4	1
25	5 (1)	1 (1)	1 (1)	6 (1)	(1)
30	1 (1)	1 (1)	1 (1)	4	1
35	1 (1)	5 (1)	1 (1)	1 (1)	2 (1)
40	1 (1)	8 (1)	1 (1)	2	5
45	1			(1)	2 (1)
50	1			1	1
55	1				
60	1				
Mean discount (%)	30	30	10	25	25

NOTE.—Experiments 1 and 3 utilized the negative-skew low-mean (NSLM), positive-skew low-mean (PSLM), and negative-skew high-mean (NSHM) distributions, while experiment 2 utilized the normal and bimodal distributions. The stimuli for experiment 1 were airline ticket prices, and thus full price was $2,400. The stimuli for experiments 2 and 3 were prices for 2-liter bottles of cola, and thus full price was $1.80. After rating the contextual set of prices, subjects rated the five focal prices. The price frequency of the focal prices are shown in parentheses.

This number is more than the 11 prices used by Janiszewski and Lichtenstein (1999) but far less than the 60 choice occasions used by Kahn and Louie (1990) and Alba et al. (1994).

Experimental Materials. Airline ticket prices that are available to consumers in an electronic shopping environment was selected as the product category. As a large number of Web sites provide airline price information to consumers, the experimental paradigm would seem particularly relevant to price searches on the Internet. However, consumers may have existing internal reference prices for familiar airline tickets. To reduce the effect of previous price knowledge, a relatively complex airline ticket price was constructed. Specifically, subjects were provided airline ticket prices for a vacation that originated in the United States and made stops one week apart in London, Rome, Paris, Madrid, and Munich before returning to the United States. The range of prices selected for experiment 1 was intended to be representative of marketplace prices. To this end, a survey of airline ticket prices was conducted using a convenience sample of Web sites that were available on the Internet. Based on these data, the range of prices used in experiment 1 was from $2,400 (full price) to $960 (60% discount).

Measures. After exposure to ticket prices, subjects were asked to evaluate the price attractiveness on a nine-point scale that was anchored by "a very attractive price" (1) and "a very unattractive price" (9). This measure is similar to the price attractiveness rating used by Janiszewski and Lichtenstein (1999). The stimulus exposure time was self-paced to simulate the price elaboration time in electronic shopping environments. Nevertheless, response time was included as a potential covariate. The computer was used to measure unobtrusively the response time of each evaluation. After all prices were rated, a manipulation check was administered. Subjects were provided a list of nine prices ($2,400, $2,160, $2,040, $1,920, $1,800, $1,680, $1,560, $1,440, and $960) and were asked to identify the highest, lowest, and average prices in the experiment.

Preliminary Checks. The mean correlation between prices and ratings for all 194 subjects was .814. However, the nature of this scale appeared problematic for some subjects, as indicated by the near-zero or high-negative correlations. Apparently, some subjects reversed the measurement scale, while others did not seriously participate in the experiment. In an attempt to remove the systematic error resulting from incorrect scale usage, .50 was selected as a cutoff based on visual inspection of the data. With 19 subjects removed, the remaining 175 subjects had a mean correlation between prices and ratings of .908. Identical analyses were performed on the group consisting of all 194 subjects and on the more reliable group consisting of 175 subjects. The results are virtually identical in both groups. Consequently, only the results of the high-reliability subjects are reported. In addition, as response times did not significantly differ across distributions, the response time covariate is not included in the analysis reported here.

Manipulation Checks. The manipulation checks were evaluated in separate one-way ANOVAs. We expect large between-subject differences in price recognition when treatment conditions differ in high, low, or average prices. We also expect relatively small differences when treatment conditions do not differ in high, low, or average prices because of the effect of stimulus frequency on memory. In test 1, the range of the prices in the negative-skew low-mean and positive-skew low-mean conditions were manipulated. The price recognition of the range was significantly higher ($F(1, 131) = 218.1$, $p = .000$, $\eta^2 = .626$) in the positive-skew low-mean condition (M = $2,290, SD = $171.9) than in the negative-skew low-mean condition (M = $1,941, SD = $79.3). The range was significantly lower ($F(1, 131) = 4{,}219$, $p = .000$, $\eta^2 = .970$) in the negative-skew low-mean condition (M = $960, SD = $0.0) than in the positive-skew low-mean condition (M = $1,433, SD = $57.8). Finally, the average price was significantly higher ($F(1, 131) = 12.9$, $p = .000$, $\eta^2 = .090$) in the negative-skew low-mean condition (M = $1,724, SD = $118.6) than in the positive-skew low-mean condition (M = $1,645, SD = $131.9). As there are large differences in the range and small differences in the mean, the manipulation was successful.

In test 2, the mean in the positive-skew low-mean and negative-skew high-mean conditions were manipulated. The mean price was significantly higher ($F(1, 130) = 252.5$, $p = .000$, $\eta^2 = .662$) in the negative-skew high-mean condition (M = $2,079, SD = $178.9) than in the positive-skew low-mean condition (M = $1,645, SD = $131.9). The range was significantly higher ($F(1, 130) = 25.2$, $p = .000$, $\eta^2 = .163$) in the negative-skew high-mean condition (M = $2,400, SD = $0.0) than in the positive-skew low-mean condition (M = $2,290, SD = $171.9). The range was significantly lower ($F(1, 130) = 19.8$, $p = .000$, $\eta^2 = .133$) in the positive-skew low-mean condition (M = $1,433, SD = $57.8) than in the negative-skew high-mean condition (M = $1,494, SD = $96.4). Again, the large differences in the mean and the small differences in the range suggest that the manipulation was successful.

Test 1 Results

MANOVA. After exposure to a set of brand prices, each of the theories competes in their predictions of consumer price judgments. The first test of the three contextual models employs a 2 (distribution: positive-skew low-mean, negative-skew low-mean) by 5 (discount size: 20%, 25%, 30%, 35%, and 40%) MANOVA with repeated measures on discount size. Because the midpoint of the price range is higher in the positive-skew low-mean distribution than in the negative-skew low-mean distribution, range theory and range-frequency theory predict that ratings will be more attractive in the positive-skew low-mean distribution. In addition, because of the frequency effect, range-frequency theory predicts an interaction between price and distribution such that ratings in the negative-skew distribution will be convex, while ratings in the positive-skew distribution will be concave, with respect to attractiveness ratings. Because range theory does not account for frequency effects, range theory predicts no interaction. Finally, differences in discount distributions with the same mean cannot be explained by adaptation-level theory; thus, no differences were predicted between the ratings in the two distributions. A summary of the model predictions and the MANOVA findings are provided in table 2.

The mean attractiveness ratings, denoted by points, are plotted in figure 2. The critical tests are provided by the interaction between price and distribution, which is statistically significant ($F(4, 110) = 5.02$, $p = .001$, $\eta^2 = .043$), and the main effect of distribution, which is also significant ($F(1, 113) = 5.74$, $p = .018$, $\eta^2 = .048$). As predicted by both range theory and range-frequency theory, ratings are more attractive in the positive-skew low-mean distribution ($M = 5.543$, SD = 1.539) than in the negative-skew low-mean distribution ($M = 5.993$, SD = 1.514). Planned within-subject polynomial contrasts were evaluated to test the nature of the interaction. As predicted by range-frequency theory, the quadratic term for interaction between price and distribution is statistically significant ($F(1, 113) = 31.54$, $p = .000$, $\eta^2 = .218$). Thus, test 1 provides support for range-frequency theory, mixed support for range theory, and no support for adaptation-level theory.

Model Fitting. The models were fit to the data using least-squares iterative nonlinear regression. The overall mean attractiveness ratings were computed for the five focal prices in the positive-skew low-mean and negative-skew low-mean distributions. Thus, mean ratings for 10 prices

TABLE 2

SUMMARY OF MANOVA PREDICTIONS AND FINDINGS

Hypotheses	Experiment 1, test 1	Experiment 1, test 2	Experiment 2
H1 (ALT):			
Prediction	There are no differences between the price attractiveness ratings in the positive-skew low-mean and the negative-skew low-mean distributions	Price ratings will be more attractive in the negative-skew high-mean distribution than in the positive-skew low-mean distribution, and there is no price by distribution interaction	There are no differences between the price attractiveness ratings in the normal and the bimodal distributions
Finding	Not supported	Mixed support	Not supported
H1 (RT):			
Prediction	Price ratings will be more attractive in the positive-skew low-mean distribution than in the negative-skew low-mean distribution, and there is no price by distribution interaction	There are no differences between the price attractiveness ratings in the negative-skew high-mean and the positive-skew low-mean distributions	There are no differences between the price attractiveness ratings in the normal and the bimodal distributions
Finding	Mixed support	Not supported	Not supported
H1 (RFT):			
Prediction	Price ratings will be more attractive in the positive-skew low-mean distribution than in the negative-skew low-mean distribution, and there is a price by distribution interaction	Price ratings will be more attractive in the negative-skew high-mean distribution than in the positive-skew low-mean distribution, and there is a price by distribution interaction	Given the normal and bimodal distributions, price attractiveness ratings are affected by the interaction between price and distribution
Finding	Supported	Supported	Supported

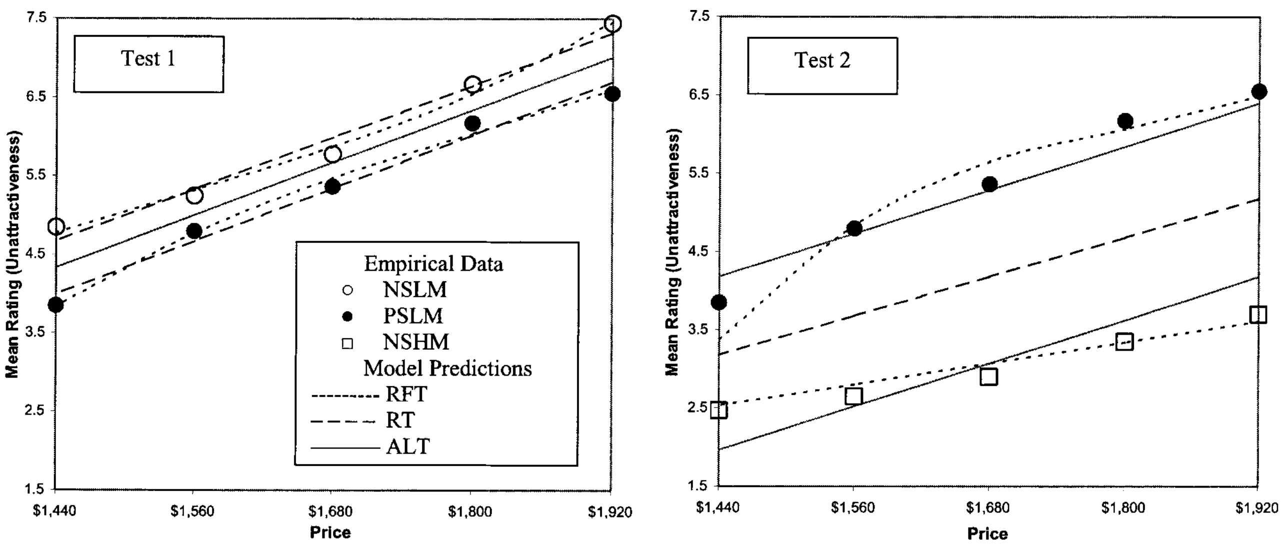
were used to fit the models in test 1. Figure 2 provides the mean attractiveness ratings (A_{ik}) at each price level ($\$_{ik}$) within each distribution condition, or context k, and the model predictions. Additional details of the model fitting procedures are provided in the appendix. As previously discussed, all models were fit with and without the valuation function parameter. Constraining r to one isolates the separate effects of the competing integration functions that are of primary concern in this experiment by holding the valuation function constant across all models. Thus, the results are provided where r is constrained to one.

The results of the model fitting for test 1 are provided in table 3. As seen in table 3, all parameter estimates are significant at $p < .01$. As indicated by the R^2 values, both range theory and range-frequency theory fit the data very well. However, as seen by the relatively low R^2 values and the model prediction shown in figure 2, adaptation-level theory systematically does not fit the data. Adaptation-level theory cannot account for the main effect of distribution. Further, since the range model is nested within the range-frequency model and the weighting parameter, w, is statistically significant, the range-frequency model provides a better fit to the data than does the range model. This is because range theory cannot account for the nonlinear interaction between price and distribution.

Discussion. Similar to Janiszewski and Lichtenstein (1999), test 1 compared contextual theories by manipulating the range in price distributions with the same mean. Using price judgments of five prices common to both distributions, both the MANOVA and model fitting results provide a clear interpretation of the data. Consistent with the conclusions of Janiszewski and Lichtenstein (1999), range theory provides a better account of the test 1 data than does adaptation-level theory. Thus, test 1 provides some support for the conclusion that the cognitive representation of reference price may include the two prices that define the range of the contextual set and not just the mean price. However, because Janiszewski and Lichtenstein (1999) measured only the mean price, the interaction between price and distribution could not be evaluated. Range theory cannot account for the nonlinear interaction in test 1. This interaction was correctly predicted by the frequency effect. Thus, test 1 provides support for range-frequency theory and the exemplar conceptualization of reference price.

FIGURE 2

EXPERIMENT 1: UNATTRACTIVENESS RATINGS BY PRICE AND DISTRIBUTION



NOTE.—Points denote empirical data, and lines represent model predictions. Attractiveness ratings are scaled from very attractive (1) to very unattractive (9). Predictions from the range-frequency theory (RFT) adhere closely to empirical points and reflect the contextual effects in both tests. Predictions from range theory (RT) capture basic effects in test 1 but not in test 2. Predictions from adaptation-level theory (ALT) capture basic effects in test 2 but not in test 1. Distributions were positive-skew low-mean (PSLM), negative-skew low-mean (NSLM), and negative-skew high-mean (NSHM).

Test 2 Results

MANOVA. The second test of the three contextual models employs a 2 (distribution: positive-skew low-mean, negative-skew high-mean) by 5 (discount size: 20%, 25%, 30%, 35%, and 40%) MANOVA with repeated measures on discount size. Because the mean price is higher in the negative-skew high-mean distribution than in the positive-skew low-mean distribution, adaptation-level theory and range-frequency theory predict that ratings will be more attractive in the negative-skew high-mean distribution. In addition, because of the frequency effect, range-frequency theory predicts an interaction between price and distribution such that the positive-skew distribution will be concave, while the negative-skew distribution will be linear, with respect to attractiveness ratings. Because adaptation-level theory is a linear model, it cannot account for nonlinear frequency effects. Finally, as differences in discount distributions with the same range cannot be explained by range theory, no differences were predicted between the ratings in the two distributions. A summary of the model predictions and the MANOVA findings are provided in table 2.

The mean attractiveness ratings, denoted by points, are plotted in figure 2. The critical test is provided by the interaction between price and distribution, which is statistically significant ($F(4, 113) = 25.99$, $p = .000$, $\eta^2 = .183$), and the main effect of distribution, which is also significant ($F(1, 116) = 68.10$, $p = .000$, $\eta^2 = .370$). As predicted by both adaptation-level theory and range-frequency theory, ratings are more attractive in the negative-skew high-mean distribution ($M = 3.013$, SD $= 1.665$) than in the positive-skew low-mean distribution ($M = 5.543$, SD $= 1.539$). Planned within-subject polynomial contrasts were evaluated to test the nature of the interaction. As predicted by range-frequency theory, the quadratic term for the interaction between price and distribution is statistically significant ($F(1, 116) = 11.65$, $p = .001$, $\eta^2 = .091$). Thus, test 2 provides support for range-frequency theory, mixed support for adaptation-level theory, and no support for range theory.

Model Fitting. The overall mean attractiveness ratings were computed for the five focal prices in the positive-skew low-mean and negative-skew high-mean distributions and were used to fit the models in test 2. Figure 2 provides the mean attractiveness ratings and the model predictions. The results of the model fitting for test 2 are provided in table 3. Except as noted, all parameter estimates are significant at $p < .01$. As indicated by the R^2 values, both adaptation-level theory and range-frequency theory fit the data very well. However, as seen by the low R^2 values and the model prediction in figure 2, range theory systematically does not fit the data. That is, range theory cannot account for the main effect of distribution. In addition, since the range model is nested within the range-frequency model and the weighting parameter, w, is statistically significant, the range-frequency model provides a better fit to the data than does the range model.

TABLE 3

MODEL FIT RESULTS

Model	Experiment 1, test 1		Experiment 1, test 2		Experiment 2		Experiment 3, simultaneous		Experiment 3, sequential	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Range:										
R^2	.982		.257		.994		.792		.959	
r					−.530	.349[a]				
$\$_{min,hi}$	912	63.84	914	545[a]			1.004	.062	1.015	.028
$\$_{min,lo}$	774	76.88			.930	.014	.868	.123	.805	.064
$\$_{min,hi}$	2,327	54.22	2,842	815	1.909	.049	1.794	.119	1.852	.062
$\$_{min,lo}$	2,226	47.66					1.640	.120	1.591	.048
Adaptation-level:										
R^2	.879		.955		.994		.904		.729	
r					−.530	.349[a]				
a	5.669	.124	5.288	.134	5.527	.093	4.892	.160	4.844	.174
b	.0056	.00073	.0046	.00035	−24.25	13.88[a]	8.461	.765	7.719	1.299
Range-frequency:										
R^2	.994		.966		.999		.992		.997	
r					−1.002	.270				
w	.807	.062	.335	.056	.790	.045	.447	.036	.755	.021
$\$_{min,hi}$	844	66.08	−1,174	3,403[a]			1.033	.029	1.027	.010
$\$_{min,lo}$	572	123.7			.926	.009	.559	.166	.714	.033
$\$_{min,hi}$	2,369	52.53	3,876	2,495[a]	2.005	.060	1.898	.075	1.902	.026
$\$_{max,lo}$	2,204	42.21					1.709	.100	1.582	.019

NOTE.—The parameter estimates are the power function exponent, r; the range-frequency theory weighting parameter, w; the adaptation-level theory intercept, a; and the slope, b; the range theory estimates for the minimum value in the high-price range, $\$_{min,\,hi}$; the maximum value in the high-price range, $\$_{max,\,hi}$; the minimum value in the low-price range, $\$_{min,\,lo}$; and the maximum value in the low-price range, $\$_{max,\,lo}$.

[a]Except as indicated by superscript, all parameter estimates are significant at $p < .01$.

Discussion. Test 2 provides the complementary condition to test 1 by manipulating the mean in price distributions while holding the range constant. As expected, adaptation-level theory provided a better account of the data than did range theory because the distribution main effect cannot be explained by differences in the distribution range. The MANOVA results provided evidence that the interaction and main effects were significant and displayed the pattern predicted by range-frequency theory. Also as expected, adaptation-level theory was partially supported through the predicted main effect, but it could not account for the observed interaction. Further analysis of the models indicated that both the adaptation-level and range-frequency models provide good fits to the data. As expected, the range model fit was poor. Thus, test 2 provides additional support for range-frequency theory and the exemplar conceptualization of reference price.

EXPERIMENT 2

Test 1 evaluated price judgments where the distribution mean was held constant, while test 2 evaluated price judgments where the distribution range was held constant. Experiment 2 extends experiment 1 by holding both the range and mean constant. This test is similar to the procedure used by Birnbaum (1974) in testing contextual theories with numbers rather than prices. Although range theory was not tested, Birnbaum (1974) found that range-frequency theory provided a better account of judgments of numbers than did adaptation-level theory. Experiment 2 also extends experiment 1 by testing whether a prototype or an exemplar model of reference prices provides a better account of consumer price judgments. Finally, experiment 2 generalizes the previous experimental findings to a different product category.

Methods

Subjects. One hundred and fifty-three students received extra credit for participating in the experiment. Of these subjects, 51% were female. The mean age of subjects was 22 years, with a standard deviation of 3.5 years and a range of 18 to 43 years.

Materials and Procedure. A 2-liter bottle of carbonated beverage was selected as the product category for the following reasons: first, price distributions of many grocery items in high-low promotional stores (i.e., not everday low-price stores) are bimodal (Raju, Srinivasan, and Lal 1990; Rao et al. 1995; Villas-Boas 1995). Second, grocery items were used in the experiment by Janiszewski and Lichtenstein (1999). Finally, carbonated beverages are purchased frequently and have considerable variability in price over time and across brands and stores.

The procedure was similar to the first experiment with the following exceptions: first, because the range was not manipulated between subjects, less discount depth was required. Thus, price discounts ranged from 0% to 50% rather than from 0% to 60%. Second, additional contextual prices were required to create the modality manipulation. Thus, 22 prices rather than 20 prices were used in the contextual set. Again, the range of prices were intended to be representative of marketplace prices. To this end, a survey of grocery store prices of 2-liter bottles of carbonated beverages was conducted using a convenience sample of local grocery stores. Based on these data, the range of prices used in experiment 2 was from $1.80 (full price) to $.90 (50% discount).

Design and Measures. The prototype representation can be tested by manipulating the modality in price sets with the same range and mean. Therefore, we used a one-factor, between-subjects design with two levels of price distribution (normal and bimodal). Judgments of prices at 5%, 15%, 25%, 35%, and 45% discounts were taken after the 22 prices shown in table 1. To mitigate order effects, the presentation order of the first 22 prices was randomized for each subject. Similarly, the presentation order of the last five target prices was randomized. Because the mean and range were the same in these two conditions, the manipulation measures from experiment 1 were not used.

Preliminary Checks. The overall correlation between prices and ratings for the 153 subjects in experiment 2 was .836. Consistent with experiment 1, the lowest acceptable correlation was set at .50. Consequently, 12 subjects were removed from the sample. The remaining 141 subjects had a mean correlation between prices and ratings of .887. Identical analyses were performed on two groups, one consisting of all 153 subjects and the other consisting of the 141 subjects in the high-reliability group. The conclusions of both groups are the same. Thus, only the results of the high-reliability subjects are reported here. Further, since there were no effects of response latencies in experiment 2, the results reported here do not include the response time covariate.

Results

MANOVA. The analysis of the three contextual models employs a 2 (distribution: normal, bimodal) by 5 (discount size: 5%, 15%, 25%, 35%, and 45%) MANOVA with repeated measures on discount size. Given a bimodal and a normal distribution of prices with the same range and mean, adaptation-level theory and range theory predict no differences between these two distributions for price judgments of any stimulus value. In contrast, range-frequency theory predicts an interaction between distribution and price and no main effect resulting from distribution. More specifically, compared with the bimodal distribution, ratings in the normal distribution are predicted to be more attractive for prices between the mean price and the lowest price in the contextual set, and less attractive for prices between the mean price and the highest price in the contextual set.

The mean attractiveness ratings, denoted by points, are plotted in figure 3. As predicted by all three theories, the main effect of distribution was not significant ($F(1, 139) = 0.06$, $p = .804$, $\eta^2 = .000$). Only range-frequency theory predicted the interaction between price and distribution, which was statistically significant ($F(4, 136) = 3.67$, $p = .007$, $\eta^2 = .097$). Planned within-subject polynomial contrasts tested the nature of the interaction. As predicted by range-frequency theory, the cubic term for the interaction between price and distribution was statistically significant ($F(1, 139) = 10.47$, $p = .002$, $\eta^2 = .070$). These results provide support for range-frequency theory and no support for the competing theories. A summary of the model predictions and the MANOVA findings are provided in table 2.

Model Fitting. The overall mean attractiveness ratings were computed for the five focal prices in the normal and bimodal distributions and were used to fit the models in experiment 2. Figure 3 provides the mean attractiveness ratings and the model predictions. All procedures are identical to those used in experiment 1. However, the results reported here are for the set of models that allows for the estimation of the valuation function parameter. As the range of the focal prices extend from 5% to 45%, versus from 20% to 40% in experiment 1, the effect of the valuation function becomes more pronounced. Thus, the fully specified models are reported because the better fitting models are more easily interpreted in figure 3. As before, allowing r to be a free parameter does not change the conclusions.

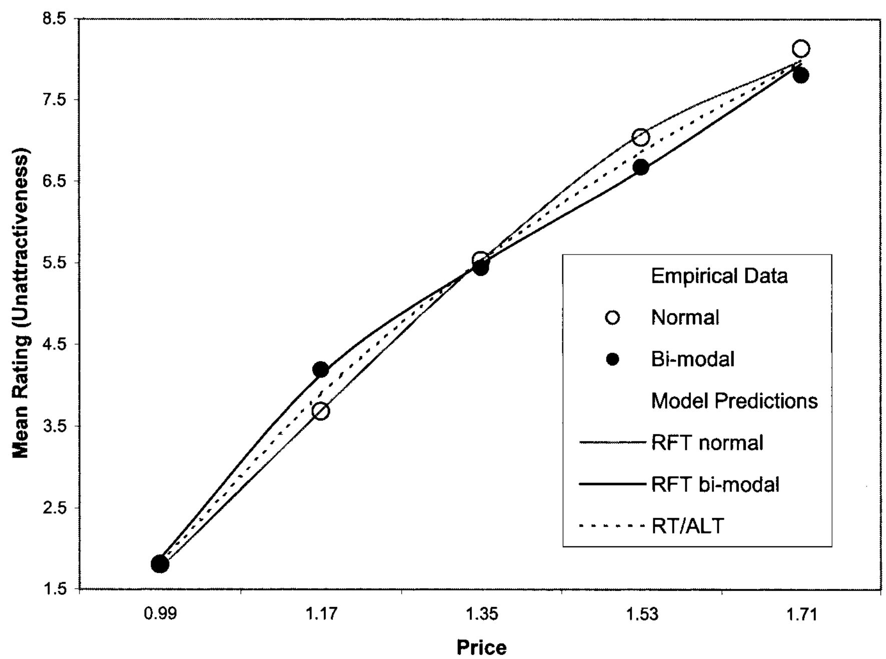
The results of the model fitting for experiment 2 are provided in table 3. Except as noted, all parameter estimates are significant at $p < .01$. In the models, $-r$ can be interpreted as $S = -\$^{-r}$, where the positive relationship between ratings and stimuli is maintained by the sign of the linear parameter estimates. As indicated by the R^2 values, all of the models fit the data very well. Since the midpoint of the range and the mean are in the same location in these two price sets, the adaptation-level and range models provide identical fits to the data. However, as seen in figure 3, range theory and adaptation-level theory systematically misfit the data because these two theories cannot account for the interaction between price and distribution. As the range model is nested within the range-frequency model and the weighting parameter is statistically significant, the range-frequency theory provides a better fit to the data than does range theory.

Discussion

Manipulating the modality of the two price sets was used to test the representation of reference price. As the two distributions share the same range and mean, range theory and adaptation-level theory predict no differences in the attractiveness ratings between these two price sets. However, range-frequency theory predicts an interaction between price and distribution. The hypotheses were evaluated using two separate methods. In the MANOVA framework, the interaction between price and distribution was both significant and in the direction predicted by range-frequency theory. Model fitting showed that range-frequency theory provides a significantly better fit to the data than range theory and is the only model that can account for the crossover effect.

FIGURE 3

EXPERIMENT 2: UNATTRACTIVENESS RATINGS BY PRICE AND DISTRIBUTION



NOTE.—Points denote empirical data, and lines represent model predictions. Attractiveness ratings are scaled from very attractive (1) to very unattractive (9). Neither range theory (RT) nor adaptation-level theory (ALT) predicts any differences in ratings for the two distributions. Range-frequency theory (RFT) predicts the contextual effect.

Thus, range-frequency theory was supported, while adaptation-level theory and range theory were not.

EXPERIMENT 3

The results of the first two experiments strongly support the idea that consumers may use an exemplar representation of reference price when price judgments are based on internal reference prices. One possible criticism of the procedure used in the first two experiments is that subjects were forced overtly to rate each price in the contextual set, one price at a time. Although consumers may evaluate a number of prices in relatively short duration in an electronic shopping environment, there are a number of situations in which this is clearly not the case. The forced judgment task may map prices to overt categories, resulting in an increased frequency effect. To mitigate this issue in experiment 3, attractiveness ratings were not collected for the prices in the contextual set.

A second issue concerns situations in which consumers use external reference prices in a price judgment task. The presentation conditions in experiment 3 were intended to represent price judgments using either external or internal reference prices. As stated in hypotheses 2 and 3, we predicted that subjects will give more weight to the range when price judgments are memory based than when price judgments are stimulus based. As a result, the affect of distribution on price judgments, as well as the model fits, are expected to be moderated by the stimulus presentation condition.

Although experiment 3 provides an additional test of competing contextual theories, the primary purpose is to test hypotheses 2 and 3. The hypotheses were tested in a number of ways. First, we evaluated the attractiveness ratings and response times in a MANOVA framework. Second, we evaluated the price recognition data using cross-tabulations. Finally, we fit the three contextual models to the data using the same procedure as in the previous experiments.

Methods

Subjects and Materials. Two hundred and eighty-seven students received extra credit for participating in the experiment. Of these subjects, 56% were female. The mean age of subjects was 22.5 years, with a standard deviation of 4.5 years and a range of 19 to 49 years. As in experiment 2, carbonated beverages was selected as the product category.

Design and Procedure. Subjects were randomly assigned in a 3 (price distribution: negative-skew low-mean, positive-skew low-mean, and negative-skew high-mean) by

2 (stimulus presentation: sequential and simultaneous) between-subjects design. The discount distributions were the same as in experiment 1. Thus, the range of cola prices used in experiment 3 was from $1.80 (full price) to $.72 (60% discount). Subjects were exposed to a total of 25 prices, as shown in table 1. The presentation order of the 20 contextual prices was randomized for each subject to mitigate order effects. Similarly, the presentation order of the five focal prices, evaluated at 20%, 25%, 30%, 35%, and 40% discounts, was also randomized.

In contrast to the previous experiments, attractiveness ratings were not measured for the 20 contextual prices. The sequential presentation procedure was the same as the one used in the first two experiments, with subjects exposed to the 25 prices, one at a time. However, in the simultaneous presentation condition, subjects were provided with all the prices at the same time. Specifically, they were presented with five computer screens, one for each focal price. Each screen provided subjects with the 20 contextual prices and one randomly selected focal price. The measures were the same as in experiment 1.

Preliminary Checks. The overall correlation between prices and ratings for the 287 subjects in experiment 3 was .689. As before, the lowest acceptable correlation was set at .50. Consequently, 44 subjects were removed from the sample. The remaining 243 subjects had a mean correlation between prices and ratings of .887. Identical analyses were performed on two groups, one consisting of all 287 subjects and the other consisting of the 243 subjects in the high-reliability group. The conclusions of both groups are the same. Thus, only the results of the high-reliability subjects are reported here.

Manipulation Checks. As in experiment 1, the range of the prices in the negative-skew low-mean and positive-skew low-mean conditions were manipulated. The price recognition of the range was significantly higher ($F(1, 210) = 130.6$, $p = .000$, $\eta^2 = .385$) in the positive-skew low-mean condition (M = $1.672, SD = $0.178) than in the negative-skew low-mean condition (M = $1.462, SD = $0.009). The range was significantly lower ($F(1, 210) = 199.1$, $p = .000$, $\eta^2 = .488$) in the negative-skew low-mean condition (M = $0.815, SD = $0.166) than in the positive-skew low-mean condition (M = $1.080, SD = $0.005). Finally, there was no significant difference between the average price in the two conditions ($F(1, 210) = 0.0$, $p = .969$, $\eta^2 = .000$). As there are large differences in the range and no difference in the mean, the manipulation was successful.

Similarly, as in experiment 1, the mean of the prices in the positive-skew low-mean and negative-skew high-mean conditions were manipulated. The average price was significantly higher ($F(1, 158) = 88.7$, $p = .000$, $\eta^2 = .361$) in the negative-skew high-mean condition (M = $1.466, SD = $0.145) than in the positive-skew low-mean condition (M = $1.280, SD = $0.102). The range was significantly higher ($F(1, 158) = 12.9$, $p = .000$, $\eta^2 = .075$) in the negative-skew high-mean condition (M = $1.760, SD = $0.122) than in the positive-skew low-mean condition (M = $1.672, SD = $0.178). Finally, there was no significant difference between the two groups in the low price recognition ($F(1, 158) = 2.4$, $p = .120$, $\eta^2 = .015$). As there are large differences in the mean and small differences in the range, the manipulation was successful.

Results

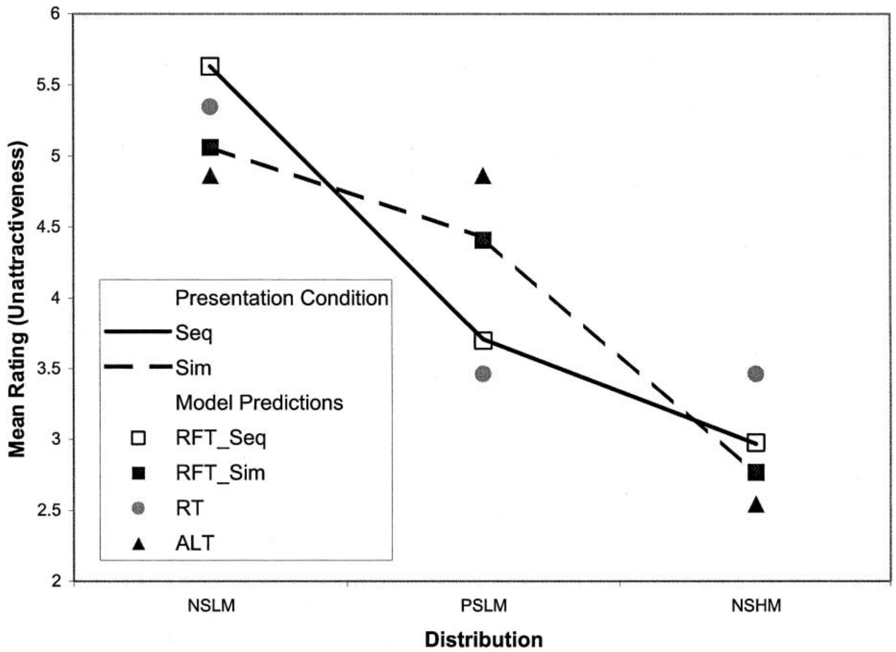
MANOVA. The hypotheses were tested in a 3 (distribution: negative-skew low-mean, positive-skew low-mean, negative-skew high-mean) by 2 (stimulus presentation: sequential, simultaneous) by 5 (discount size: 20%, 25%, 30%, 35%, and 40%) MANOVA with repeated measures on discount size. The dependent variables were response times and attractiveness ratings for the five focal prices. If subject evaluations consider only the high and low prices, ratings should not differ between the negative-skew high-mean and positive-skew low-mean conditions. Conversely, if subject evaluations consider only the mean price, ratings should not differ between the negative-skew low-mean and positive-skew low-mean conditions. As shown in figure 4, there is a significant interaction between distribution and presentation condition on attractiveness ratings ($F(2, 237) = 5.64$, $p = .004$, $\eta^2 = .045$). Mean ratings, denoted by lines in figure 4, in the positive-skew low-mean and negative-skew high-mean conditions are more similar in the sequential condition than in the simultaneous condition. On the other hand, mean ratings in the positive-skew low-mean and negative-skew low-mean conditions are more similar in the simultaneous condition than in the sequential condition. Thus, relative to each other, subjects in the simultaneous presentation condition place more weight on the mean, while subjects in the sequential presentation condition place more weight on the range of reference prices.

The previous finding suggests that subjects are comparing the focal price against more contextual prices in the simultaneous presentation condition than in the sequential presentation condition. If this were the case, we would expect price judgments to take longer in the simultaneous condition. Confirmation of this process is provided by the significant main effect of presentation condition on response times in the repeated measures MANOVA ($F(1, 237) = 74.56$, $p = .000$, $\eta^2 = .239$). Specifically, response times were significantly longer for subjects in the simultaneous presentation condition ($M = 11.048$ secs., SD = 4.606) than for subjects in the sequential presentation condition ($M = 6.570$ secs., SD = 2.624). Thus, both the attractiveness ratings and response times provide support for hypotheses 2 and 3.

Price Recognition. We have argued that, relative to subjects in the simultaneous presentation condition, subjects in the sequential presentation condition place more weight on high and low prices because these extreme values are more salient than other less-available reference prices. If this were true, then price recognition for high and low prices

FIGURE 4

EXPERIMENT 3: UNATTRACTIVENESS RATINGS BY DISTRIBUTION AND PRESENTATION CONDITION



NOTE.—Lines denote empirical data, and points represent model predictions. Attractiveness ratings are scaled from very attractive (1) to very unattractive (9). The significant interaction between distribution and presentation condition indicates that, relative to each other, range effects are stronger in the sequential presentation condition (Seq) and frequency effects are stronger in the simultaneous presentation condition (Sim). Predictions from the range-frequency theory (RFT) adhere closely to empirical lines. Predictions from adaptation-level theory (ALT) and range theory (RT) do not fit the data. Distributions were negative-skew low-mean (NSLM), positive-skew low-mean (PSLM), and negative-skew high-mean (NSHM).

should be better for subjects in the sequential rather than in the simultaneous condition. To test this proposition, responses for high and low price recognition used for the manipulation check were recoded into correct and incorrect categories. As expected, subjects were more likely to recognize the high price correctly in the sequential presentation condition (86.2%) than in the simultaneous presentation condition (72.5%), and this difference is statistically significant ($\chi^2 = 6.95$, $p = .008$). Also as expected, subjects were more likely to recognize the low price correctly in the sequential presentation condition (91.9%) than in the simultaneous presentation condition (80.0%), and this difference is statistically significant ($\chi^2 = 7.11$, $p = .008$). Thus, the price recognition data provide additional support for hypotheses 2 and 3.

Model Fitting. The model fitting procedure used in experiment 3 was the same one that was used in the previous experiments, except that ratings from three distributions were used rather than two. The overall mean attractiveness ratings were computed for the five focal prices in the negative-skew low-mean, positive-skew low-mean, and negative-skew high-mean distributions. Thus, mean ratings for 15 prices were used to fit the models in each presentation condition. Consistent with experiment 1, the valuation parameter was constrained to one. As before, allowing r to be a free parameter does not change the conclusions. The results of the model fitting for experiment 3 are provided in table 3. All parameter estimates are significant at $p < .01$.

Each contextual model differs in its operationalization of reference prices. If hypotheses 2 and 3 are correct, each model will be affected by the presentation condition. Because adaptation-level theory is heavily influenced by frequency effects, hypothesis 2 suggests that adaptation-level theory should fit better in the simultaneous presentation condition than in the sequential presentation condition. As shown in table 3, the adaptation-level model R^2 value is higher in the simultaneous presentation condition (.904) than in the sequential presentation condition (.729). It is also clear from figure 4 that adaptation-level theory fits the data better in the simultaneous presentation condition, providing additional support for hypothesis 2.

In contrast, range theory considers only the high and low prices. Consequently, hypothesis 3 predicts that range theory should fit better in the sequential presentation condition than in the simultaneous presentation condition. As shown in table 3, the range model R^2 value is higher in the sequential presentation condition (.959) than in the simultaneous presentation condition (.792). It is clear from figure 4 that range theory fits the data better in the sequential presentation condition, providing further support for hypothesis 3.

In range-frequency theory, *w* describes the relative weight given to the range and frequency principle. As the weighting parameter approaches the theoretical limit of one, range-frequency theory approaches range theory. While, the model fit is not expected to be affected by the presentation condition, hypotheses 2 and 3 suggest that the weighting parameter will be moderated by the presentation condition. As shown in table 3 and figure 4, range-frequency theory provides excellent fits in both presentation conditions. As expected, the weighting parameter is significantly higher ($t = 10.85$, $p = .000$) in the sequential presentation condition ($w = .755$, SE $= .0209$) than in the simultaneous presentation condition ($w = .447$, SE $= .0359$). Thus, the presentation condition moderates the weighting parameter in the range-frequency model, providing additional support for hypotheses 2 and 3.

Discussion

Hypotheses 2 and 3 were tested using measures of price attractiveness, response time, and price recognition. The analysis provides empirical evidence that subjects in the sequential presentation condition took less time to evaluate the focal prices and were more likely to identify the highest and lowest prices correctly than were subjects in the simultaneous condition. These data are consistent with the ideas that subjects in the sequential presentation condition compared the focal price to fewer prices and that the most salient prices were the highest and lowest prices in the contextual set. In addition, the MANOVA provides evidence that attractiveness ratings are affected by the interaction between distribution and stimulus presentation condition. Consequently, the fit of the contextual models was moderated by the presentation condition. Taken together, these data provide strong evidence that price contexts, and thus judgments, are moderated by the processing environment. Specifically, subjects in the sequential presentation condition place more weight on extreme values in the contextual set. As a result, range effects are stronger in the sequential presentation condition, and frequency effects are stronger in the simultaneous presentation condition.

GENERAL DISCUSSION

The purpose of this research was to provide a better understanding of the cognitive representation of reference prices and of how these prices are used by consumers in a price judgment task. To address these questions, Volkmann's (1951) range theory, Parducci's (1965) range-frequency theory, and Helson's (1947, 1964) adaptation-level theory were pitted against each other under a number of experimental conditions. The results of these experiments provide a number of theoretical contributions.

Reference Prices

As previously discussed, each contextual model conceptualizes reference price differently. Thus, support for one model provides evidence in favor of the assumed reference price representation. The prototype (adaptation-level theory) and range (range theory) conceptualizations were inconsistent with the data in experiment 1 and could not explain the effects of price distribution in experiment 2. Over a number of experimental conditions, range-frequency theory provided the best account of the data, which supported the idea that consumers store, retrieve, and use a rich array of price information in the process of generating price judgments. Consumers not only have a sense of the range but also of the relative frequencies of prices they have encountered. The distributional effects of price on judgments are consistent with an exemplar model of reference price.

While the cognitive representation of reference price appears to be exemplar based, not all exemplars in the relevant price set appear to be equally important in price judgment. The observation of high values of the range weighting in some conditions suggests that the two prices defining the range may at times explain the majority of reference price effects. Furthermore, the weight given to the reference prices appears to depend on how these are experienced. We have argued here that range values are salient and thus more easily retrieved from memory (Fiske and Taylor 1991; Tversky and Kahneman 1973). Intermediate values, on the other hand, while readily available in the simultaneous condition, may not be as easily retrieved in the sequential condition. Thus, prices defining the range receive greater weight in the sequential presentation condition than in the simultaneous condition. These data suggest that, relative to each other, the range values of internal reference prices are more heavily weighted and the intermediate values of external reference prices are more heavily weighted.

Support for hypotheses 2 and 3 may provide a number of implications to reference price researchers. For example, Mazumdar and Papatla (2000) find that the weight given to internal and external reference prices may vary by product category. This finding underscores the need to consider range and frequency effects in the operationalization of reference prices in choice models. One implication for brand choice modelers is that internal reference price may need to include a range component in addition to the exponentially weighted model that is often suggested (Briesch et al. 1997; Jedidi, Mela, and Gupta 1999). In addition, hypotheses 2 and 3 may suggest an additional mechanism for some of the effects reported by Alba and his colleagues. Alba et al. (1994) found that discount frequency had a stronger effect on price perceptions than did discount depth in a number of experiments when the stimulus presentation mode was predominantly simultaneous. On the other hand, Alba et al. (1999) found the reverse was typically true when prices were presented sequentially. The findings here suggest that the stimulus complexity explanation provided by Alba et al. (1999) may be moderated by the stimulus presentation condition.

Applicability of Contextual Models

While range-frequency theory provided the most robust explanation for the experiments conducted here, there were conditions in which adaptation-level theory and range theory provided very good accounts of the data. In reference price sets containing only one or two values, all three models will provide the same fit. Further, all three models will provide similar fits in larger reference price sets when the mean and the midpoint of the contextual range are the same. One exception to this statement was demonstrated in experiment 2, in which adaptation-level theory and range theory could not account for differences in distribution modality. In addition, adaptation-level theory may provide a relatively good explanation of price perceptions that are dominated by frequency effects, such as was found with external reference prices. In contrast, range theory may provide a relatively good explanation of price perceptions that are dominated by range effects, such as was found for internal reference prices.

Future Research

This research has investigated reference price effects in a laboratory setting. However, it is not clear how many reference price categories consumers use in evaluating price (Kalyanaram and Winer 1995). Winer (1988) has argued that there may be as many as eight types of reference prices. Consumers may use multiple reference prices in a single judgment. Multiple reference price categories suggest that consumers may make multiple evaluations, which are eventually integrated into a single price judgment of the brand. Alternately, multiple reference price categories could be integrated into a single category against which a single evaluation is made. Given the use of multiple reference price types, it is not clear when integration of this information occurs. Providing an understanding of the timing of integration is an interesting avenue for future research.

In addition, a better understanding of the cognitive representation of reference price is needed. This experiment suggests that reference price may be more accurately represented as a set of exemplars rather than how it is represented in the current prototype view. However, it may be possible to explain these data with other reference price conceptualizations, such as an exemplar-prototype hybrid. More generally, the values in the context at the time of the evaluation are important since they will determine the price judgment. For example, assume that a consumer has purchased a product over a 10-year period. Although the consumer may encode all 10 years of prices, it unlikely that all 10 years of prices are retrieved at the time of the evaluation. It seems reasonable to expect that some prices are more salient than others. Perhaps the context is set by the high and low prices, which are more easily retrieved, along with some subset of the more available prices.

APPENDIX

MODEL FITTING

The adaptation-level model is shown in equation A1. The adaptation-level model has three parameters to be estimated (a, b, and r) and two variables ($\$_{ik}$ and $\$_{al,\,k}$).

$$A_{ik} = a + b(\$^r_{ik} - \$^r_{al,k}). \qquad \textbf{(A1)}$$

The range-frequency model (eq. A2), has four parameters to be estimated ($\$_{\max,\,k}$, $\$_{\min,\,k}$, r, and w) and three variables ($\$_{ik}$, Rank$_{ik}$, and N_k). Since the price attractiveness scale varied from 1 to 9, $A_{\min}$ and $A_n - 1$ from equation 6 are 1 and 8, respectively. However, as shown in table 1, the distributions do not share the same range in all cases. To account for the range effect across these distributions, two additional parameters are required. For the purpose of nomenclature, let k represent the distribution in either the "hi" or "lo" range of prices and let "max" and "min" represent the extreme values that define the range in distribution k. Thus, the range-frequency model tested here has six parameters estimates (w, r, $\$_{\max,\,\text{hi}}$, $\$_{\max,\,\text{lo}}$, $\$_{\min,\,\text{hi}}$, and $\$_{\min,\,\text{lo}}$). Three dummy variables were used to account for the different range values across the distributions.

$$A_{ik} = 1 + 8[w(\$^r_{ik} - \$^r_{\min,k})/(\$^r_{\max,k} - \$^r_{\min,k}) + (1 - w)(\text{Rank}_{ik} - 1)/(N_k - 1)]. \qquad \textbf{(A2)}$$

The range model (eq. A3), is equivalent to equation A2 when $w = 1$. The range model has one variable ($\$_{ik}$) and three parameters to be estimated (r, $\$_{\max,\,k}$, and $\$_{\min,\,k}$). Two additional parameters are required to account for the range effect across these distributions. Thus, the range model tested here requires five parameter estimates (r, $\$_{\max,\,\text{hi}}$, $\$_{\max,\,\text{lo}}$, $\$_{\min,\,\text{hi}}$, and $\$_{\min,\,\text{lo}}$). Again, dummy variables were used to account for the difference in range across the distributions.

$$A_{ik} = 1 + 8(\$^r_{ik} - \$^r_{\min,k})/(\$^r_{\max,k} - \$^r_{\min,k}). \qquad \textbf{(A3)}$$

[*Received January 2000. Revised February 2001. David Glen Mick served as editor, and Joel Huber served as associate editor for this article.*]

REFERENCES

Alba, Joseph W., Susan M. Broniarczyk, Terence A. Shimp, and Joel E. Urbany (1994), "The Influence of Prior Beliefs, Frequency Cues, and Magnitude Cues on Consumers' Perceptions of Comparative Price Data," *Journal of Consumer Research*, 21 (September), 219–235.

Alba, Joseph W., Carl F. Mela, Terence A. Shimp, and Joel E. Urbany (1999), "The Effect of Discount Frequency and Depth on Consumer Price Judgments," *Journal of Consumer Research*, 26 (September), 99–114.

Anderson, Norman H. (1981), *Foundations of Integration Theory*, New York: Academic Press.

Ashby, F. Gregory and W. Todd Maddox (1998), "Stimulus Categorization," in *Measurement, Judgment, and Decision Making*, ed. Michael H. Birnbaum, San Diego, CA: Academic Press.

Birnbaum, Michael H. (1974), "Using Contextual Effects to Derive Psychophysical Scales," *Perception and Psychophysics*, 15 (October), 89–96.

Blattberg, Robert C., Richard Briesch, and Edward Fox (1995), "How Promotions Work," *Marketing Science*, 14, Part 2 (Summer), G122–G132.

Briesch, Richard A., Ladshman Krishnamurthi, Tridib Mazumdar, and Sevilimedu P. Raj (1997), "A Comparative Analysis of Reference Price Models," *Journal of Consumer Research*, 24 (September), 202–214.

Fechner, Gustav (1860), *Elemente der psychophysik*, Leipzig: Breitkopf & Härtel.

Fiske, Susan T. and Shelley E. Taylor (1991), *Social Cognition*, 2d ed., New York: McGraw-Hill.

Gupta, Sunil (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342–355.

Helson, Harry (1947), "Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data," *American Journal of Psychology*, 60 (January), 1–29.

——— (1964), *Adaptation-Level Theory*, New York: Harper & Row.

Herr, Paul M. (1989), "Priming Price: Prior Knowledge and Context Effects," *Journal of Consumer Research*, 16 (June), 67–75.

Janiszewski, Chris and Donald R. Lichtenstein (1999), "A Range Theory Account of Price Perception," *Journal of Consumer Research*, 25 (March), 353–368.

Jedidi, Kamel, Carl F. Mela, and Sunil Gupta (1999), "Managing Advertising and Promotion for Long-Run Profitability," *Marketing Science*, 18 (Winter), 1–22.

Kahn, Barbara E. and Therese A. Louie (1990), "Effects of Retraction of Price Promotions on Brand Choice Behavior for Variety-Seeking and Last-Purchase-Loyal Consumers," *Journal of Marketing Research*, 27 (August), 279–289.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decisions under Risk," *Econometrica*, 47 (March), 263–291.

Kalwani, Manohar U. and Chi Kin Yim (1992), "Consumer Price and Promotion Expectations: An Experimental Study," *Journal of Marketing Research*, 29 (February), 90–100.

Kalyanaram, Gurumurthy and John D. C. Little (1994), "An Empirical Analysis of Latitude of Price Acceptance in Consumer Package Goods," *Journal of Consumer Research*, 25 (December), 408–418.

Kalyanaram, Gurumurthy and Russell S. Winer (1995), "Empirical Generalizations for Reference Price Research," *Marketing Science*, 14, Part 2 (Summer), G161–G169.

Lichtenstein, Donald R. and William O. Bearden (1989), "Contextual Influences on Perceptions of Merchant-Supplied Reference Prices," *Journal of Consumer Research*, 16 (June), 55–66.

Marks, Lawrence E. and Daniel Algom (1998), "Psychophysical Scaling," in *Measurement, Judgment, and Decision Making*, ed. Michael H. Birnbaum, San Diego, CA: Academic Press.

Mazumdar, Tridib and Purushottam Papatla (2000), "An Investigation of Reference Price Segments," *Journal of Marketing Research*, 37 (May), 246–258.

Medin, Douglas L., Mark W. Altom, and Timothy D. Murphy (1984), "Given versus Induced Category Representations: Use of Prototype and Exemplar Information in Classification," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10 (July), 333–352.

Monroe, Kent B. (1971), "Measuring Price Thresholds by Psychophysics and Latitudes of Acceptance," *Journal of Marketing Research*, 8 (November), 460–464.

——— (1990), *Pricing: Making Profitable Decision*, New York: McGraw-Hill.

Mook, Douglas G. (1983), "In Defense of External Validity," *American Psychologist*, 38 (April), 379–387.

Parducci, Allen (1965), "Category Judgment: A Range-Frequency Model," *Psychological Review*, 72 (November), 407–418.

——— (1995), *Happiness, Pleasure, and Judgment: The Contextual Theory and Its Applications*, Mahwah, NJ: Erlbaum.

Parducci, Allen and Louise M. Marshall (1961), "Context Effects in Judgments of Length," *American Journal of Psychology*, 74 (December), 576–583.

Parducci, Allen and Linda F. Perrett (1971), "Category Rating Scales: Effects of Relative Spacing and Frequency," *Journal of Experimental Psychology*, 89 (August), 427–452.

Parducci, Allen and Douglas H. Wedell (1986), "The Category Effect with Rating Scales: Number of Categories, Number of Stimuli, and Method of Presentation," *Journal of Experimental Psychology: Human Perception and Performance*, 12 (November), 496–516.

Raju, Jagmohan S., V. Srinivasan, and Rajiv Lal (1990), "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science*, 36 (March), 276–304.

Rao, Ram C., Ramesh V. Arjunji, and B. P. S. Murthi (1995), "Game Theory and Empirical Generalizations concerning Competitive Promotions," *Marketing Science*, 14, Part 2 (Summer), G89–G100.

Tversky, Amos and Daniel Kahneman (1973), "Availability: A Heuristic for Judging Frequency and Probable Cause," *Cognitive Psychology*, 5 (September), 207–232.

Villas-Boas, J. Miguel (1995), "Models of Competitive Price Promotions: Some Empirical Evidence from the Coffee and Saltine Crackers Market," *Journal of Economics and Management Strategy*, 4 (Spring), 85–107.

Volkmann, John (1951), "Scales of Judgment and Their Implications for Social Psychology," in *Social Psychology at the Crossroads*, ed. John H. Rohrer and Muzafer Sherif, New York: Harper & Row, 273–294.

Wedell, Douglas H. (1995), "Contrast Effects in Paired Comparisons: Evidence for Both Stimulus-Based and Response-Based Processes," *Journal of Experimental Psychology: Human Perception and Performance*, 21 (October), 1158–1173.

——— (1996), "A Constructive-Associative Model of the Contextual Dependence of Unidimensional Similarity," *Journal of Experimental Psychology: Human Perception and Performance*, 22 (June), 634–661.

Wedell, Douglas H., Allen Parducci, and Michael Lane (1990), "Reducing the Dependence of Clinical Judgment of the Immediate Context: Effects of Number of Categories and Type of Anchors," *Journal of Personality and Social Psychology*, 58 (February), 319–329.

Winer, Russell S. (1988), "Behavioral Perspectives on Pricing: Buyer's Subjective Perceptions of Price Revisited," in *Issues in Pricing: Theory and Research*, ed. Timothy Devinney, Lexington, MA: Lexington.