Title
Comparison-Induced Sequence Effects on Hedonic Evaluations

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN
1069-7977

Authors
Choplin, Jessica
Lombardi, Megan

Publication Date
2010

Peer reviewed

Comparison-Induced Sequence Effects on Hedonic Evaluations

Jessica M. Choplin (jchoplin@depaul.edu) & Megan M. Lombardi (mlombar7@depaul.edu)
DePaul University Department of Psychology
2219 North Kenmore Avenue
Chicago, IL 60614-3504

Abstract

Hedonic evaluations and emotional reactions to experiences depend not only upon the conditions being experienced, but also upon the sequences in which conditions are experienced. The authors propose a comparison-induced distortion (CID) model of sequence effects on evaluation in which to-be-evaluated exemplars are verbally compared (Choplin & Hummel, 2002; Choplin, 2007) to the most similar, recent exemplars. Predictions of this model were tested and pit against Helson's (1964) adaptation-level theory, Parducci's (1995) range-frequency theory, and Haubensak's (1992) consistency model using a paradigm in which sequences periodically improved (i.e., improved for *n* trials, returned to the original state on a single trial, and improved for *n* trials again) or periodically deteriorated by small or large amounts. The results were consistent with the predictions of the proposed CID model of sequence effects and inconsistent with adaptation-level theory, range-frequency theory, and the consistency model.

Sequence Effects

Most theories of the causes of emotions posit that people's emotional reactions to the conditions they experience (e.g., prices, salaries, pain, tastes, wait times, and so forth) depend in part upon their hedonic evaluations of those conditions—how good or bad they judge those conditions to be (e.g., Kahneman, 1999; Lazarus, 1991; Tesser & Martin, 1996). While some researchers conceptualize these evaluations as measurements on a single good/bad dimension (Kahneman, 1999) and other researchers conceptualize these evaluations as separate measurements of how good conditions are and how bad conditions are (Cacioppo, Gardner, & Berntson, 1999; Watson & Tellegen, 1985), the general idea underlying this common notion is that these evaluations play a central role in the emotions people experience. If people judge conditions to be good, then their emotional reactions will generally be positive. That is, the evaluation that a condition is good will evoke emotions such as happiness, delight, and relief. If people judge conditions to be bad, then their emotional reactions will generally be negative. That is, the evaluation that a condition is bad will evoke emotions such as anger, frustration, and worry.

A challenge for research on emotion and hedonic evaluation arises from the fact that evaluations are not a pure function of the objective conditions being evaluated. Rather, hedonic evaluations depend upon the context in which conditions are experienced (see, for example, Parducci, 1995). Predicting people's emotional reactions, therefore, requires an understanding of the contextual factors that affect evaluations. The research reported here investigated one type of context effect on evaluation, namely, the effect of the sequence in which conditions are experienced. Paying $2.85 per gallon of gasoline, for example, might seem more reasonable if recent prices have been over $2.85 than if recent prices have been under $2.85. Waiting 5 minutes for a bus might seem more reasonable if one has lately been waiting more than 5 minutes than if one has been waiting less. The purpose of the research reported here was to develop and test an account of how the sequences in which conditions are experienced affect evaluations. We will propose a comparison-induced distortion (CID) account of sequence effects on evaluation and pit this account against Helson's (1964) adaptation-level theory, Parducci's (1995) range-frequency theory, and Haubensak's (1992) consistency model.

Comparison-Induced Distortions

The basic idea underlying the CID model of sequence effects is that people will verbally compare to-be-evaluated items to the most similar items they have recently encountered. The model selects previous exemplars to be compared to the to-be-evaluated item using two criteria: 1) giving more weight to exemplars that are more similar to the to-be-evaluated item and 2) giving more weight to more recent items (the 1-back item is weighted higher than the 2-back item, the 2-back item higher than the 3-back item, etc.). Although the proposal that these two criteria determine recall and comparison processes has not previously been applied to sequence effects on hedonic evaluations, it has been applied to other domains in which recall of previously presented exemplars affects judgment (see, for example, Nosofsky & Palmeri, 1997; Smith & Zarate, 1992).

After a comparison item is selected, we hypothesize that the to-be-evaluated item is verbally (often sub-vocally, not out loud) compared to the comparison item. CID theory predicts that verbally comparing items causes people to exaggerate small differences and under appreciate the size of large differences. The reason for this pattern of evaluation is that evaluations are biased toward the central tendency (i.e., mean or median) of values associated with a comparison word (Huttenlocher, Hedges, & Vevea, 2000).

For example, a longer wait for a bus could be 1 extra minute or 30 extra minutes, but when we consulted city bus schedules we found that the most common wait time between buses was 10 minutes. That is, there is a distribution of differences in wait times between busses where the central tendency of the distribution of "longer wait times" is around 10 extra minutes. Huttenlocher et al. (2000) demonstrated that judgments of values are typically biased towards the central tendency of categories. If so, then a "longer wait time" of 5 extra minutes (i.e., less than the central tendency of the category of "longer wait times" which is 10 extra minutes) would be biased towards the evaluation of 10 extra minutes. That is, the difference would be exaggerated. Likewise, a "longer wait time" of 15 extra minutes (i.e., more than the central tendency of 10 extra minutes) would also be biased towards the evaluation of 10 extra minutes. This time, however, the bias would cause the size of the difference to be under appreciated.

The experiment described below was designed to test the predictions of this model. To do so, participants were asked to rate their aversion to several fictional wait times for a bus in the winter. Participants evaluated wait time sequences that improved or deteriorated in large or small increments, returned to a value near the original value, and then improved or deteriorated again. CID theory predicts that there will be an amount of change (small or large) by direction of change (improving or deteriorating) interaction effect. Specifically, wait times in the small-increment deteriorating sequence will be rated as worse than wait times within the small-increment improving sequence, because these small differences will be exaggerated. A little worse will seem like a lot worse and a little better will seem like a lot better. Conversely, wait times in the large-increment deteriorating sequence will be rated better than items in the large-increment improving sequence, because people will under appreciate the sizes of the differences. A lot worse will seem as if it is only a little worse and a lot better will seem as if it is only a little better.

Adaptation-Level Theory

Models of sequence effects on evaluation commonly start with Helson's (1964) proposal that evaluations are made relative the conditions to which people have adapted—that is, the conditions to which they have become accustomed, consider normal, and continue to expect (see Briesch, Krishnamurthi, Mazumdar, & Raj, 1997; Frederick & Loewenstein, 1999; Kalyanaram & Winer, 1995, for reviews). Helson (1964) modeled the conditions to which people adapted as the running average of all previously experienced conditions (see also Kalwani, Yim, Rinne, & Sugita, 1990; Rajendran & Tellis, 1994; Wedell, 1995). These running averages then serve as reference points against which all other conditions are evaluated. Models that appeal to this explanation of sequence effects on evaluation assume that sequence effects occur when the conditions people consider normal change as they experience more instances. If people experience additional favorable conditions, they will start to consider these favorable conditions to be normal. If people experience additional unfavorable conditions, they will start to consider unfavorable conditions normal. This change in what is considered normal, thereby, causes sequence effects wherein the same conditions might be evaluated as better (or worse) depending upon whether the previously experienced values were better or worse.

Contrary to CID theory, AL theory predicts that items in deteriorating sequences will always be judged worse than items in improving sequences regardless of the size of the difference between items. Since items are evaluated in comparison to the adaptation level—which is the average of all previous items—the same item (e.g., 36 minutes) will be evaluated differently based on the average of the items that precede it in the sequence. For a deteriorating sequence, the to-be-evaluated wait time will be worse than what people have gotten used to, the average that they consider normal. For an improving sequence, the to-be-evaluated wait times will be better than what they are used to and consider normal. This pattern would be true regardless of the sizes of the differences between wait times. That is, adaptation-level theory predicts that wait times will be evaluated as worse in deteriorating than in improving conditions.

Range-Frequency Theory

One of the most important models of hedonic evaluation is Parducci's (1995) range-frequency theory. Range-frequency theory is based on the idea that judgments are made based on a compromise between range and frequency principles. According to the range principle, individuals evaluate items relative to the smallest and largest values that they have previously encountered. The individual's evaluation is based on a calculation of the range value for the to-be-evaluated item, which is the proportion of the range at which the to-be-evaluated item is located relative to the smallest and largest values. The midpoint between the highest and lowest values would be 50% of the way to the largest value from the smallest; half way between the smallest value and the midpoint would be 25%; and half way between the midpoint and the largest value would be 75%. According to the frequency principle, individuals evaluate to-be-evaluated items by calculating their percentile rank among all of the items that they have seen. Individuals compromise between these two principles when making evaluations.

Unlike comparison-induced distortion theory, range-frequency theory predicts that there will be no effect of the amount of change between wait times as long as there are no changes in the range or frequency values of the to-be-evaluated wait times. The experiment described below controls for this issue by keeping range and frequency values constant across the amount of change manipulation. Furthermore, the frequency values (percentile ranks) of the to-be-evaluated wait times would be larger (worse) in the deteriorating sequence than in the improving sequence, because better previous exemplars would be included in the context of judgment. Like adaptation-level theory, then, range-frequency theory predicts that wait times will be evaluated as worse in deteriorating than in improving conditions.

Consistency Model

Similar to comparison-induced distortion theory, Haubensak's (1992) consistency model of evaluation relies on the basic assumption that recalled exemplars affect judgment. According to the consistency model, people strive for internal consistency in their responses since judgments are subjective in that the mapping between the real-world dimension and the category-rating dimension is arbitrary. In attempting to maintain internal consistency in their responses, people often constrain their responses based on the first few exemplars they encounter. By making judgments about preceding stimuli in a sequence, for example, the person is confining subsequent judgments to a specific response scale. If the person evaluates the first few items in a sequence, doing so commits them to giving subsequent judgments that are consistent with the previous judgments. Evaluations that are consistent with those previous judgments can be calculated by linearly interpolating from previous evaluations.

To control for the effects of early judgments and the requirement that participants' evaluations be consistent with these judgments, an initial sequence was held constant across the direction of change manipulation for the current study. If participants linearly interpolate from the initial judgments they make, then evaluations should be the same for the deteriorating and improving conditions. To control for memory effects, a high and a low value were always present within the previous five trials (less than the seven represented in Haubensak's model) and these high and low values were constant across the direction of change conditions. Since this model predicts that the initial sequence will be the basis for wait time evaluations, this model predicts no effects of the amount of change between exemplars.

Experiment

To pit the predictions of the CID model of sequence effects against the predictions of the other models, we used a paradigm in which participants imagined that they had to wait for the bus in a rural town in northern Minnesota on each of 36 fictional winter days (manipulated within a single session). Wait times either periodically improved (i.e., times became successively shorter on each of *n* trials, returned to a value near the original state on a single trial, and then became successively shorter on each of *n* trials again) or periodically deteriorated (i.e., times became successively longer on each of *n* trials, returned to a value near the original state on a single trial, and then became successively longer on each of *n* trials again). Participants rated how aversive each wait time would be.

This sequence is effective in pitting the CID model against the other models. The CID model predicts that the size of the difference between consecutive exemplars will matter such that exemplars within periodically deteriorating series will be rated as worse than exemplars within periodically improving series when there is a small amount of change between consecutive exemplars. Additionally, exemplars within periodically deteriorating series will be rated as better than exemplars within periodically improving series when there is a large amount of difference between consecutive exemplars. Adaptation-level and range-frequency theories, by contrast, predict that exemplars within periodically deteriorating series will always be judged worse than exemplars within periodically improving series. The consistency model predicts no effects of the amount of change manipulation.

Method

Participants. An experimenter, who was blind to the hypotheses, approached individual prospective participants on a university campus or in the surrounding community. Two hundred and five people volunteered after being approached in this manner. Approximately half of the participants (n = 101) experienced wait times that changed (improved or deteriorated) by small amounts (i.e., 5 minutes) on each trial, excluding periodic large changes. Of these, 50 participants were in the periodically improving condition and 51 were in the periodically deteriorating condition. The other half of the participants (n = 104) experienced wait times that changed (improved or deteriorated) by large amounts (i.e., 15 minutes) on each trial, excluding periodic larger changes. Of these, 54 participants were in the periodically improving condition and 50 were in the periodically deteriorating condition.

Materials and Procedure. Participants imagined that they were spending 36 days in northern Minnesota during the middle of the winter and had to rely upon an erratic bus for transportation. The amount of time they spent waiting for the bus each day was presented aloud and participants rated how aversive that wait time for each day would be on a scale from 0 ("not bad") to 10 ("extremely bad").

CID hypothesized that participants would overreact to differences smaller than 10 minutes (i.e., 5 minutes) and under-react to differences larger than 10 minutes (i.e., 15 minutes), because the median of values from the category of "longer wait times" for busses was 10 minutes in local bus schedules for the campus community. Furthermore, the results of the experiment (presented shortly) suggest that participants did overreact to 5-minute differences and under-react to 15-minute differences. Sequences of presented values were constructed by dividing the middle 26 days of the experiment into two 13-day periods. Within each 13-day period, periodically improving and deteriorating sequences like those in Table 1 were presented. The order of the three series shown in Table 1 (i.e., Series A, B, and C) was fully counterbalanced to produce six counterbalanced groups for each of the four—2 (amount of change: 5 minutes or 15 minutes) x 2 (direction of change: improving or deteriorating)—conditions. The sequence of wait times in the second 13-day period was identical to the sequence in the first 13-day period.

To control for primacy effects and introduce participants to the range of values they would see prior to the sequence manipulation, a sequence of 5 days was inserted at the beginning of the experiment. The sequence on these 5 days was 22, 35, 50, 35, and 22 minutes respectively in the small-difference condition and 2, 35, 70, 35, 2 minutes respectively in the large-difference condition. To make peak and end values equivalent across periodically improving and deteriorating sequences before asking participants to make a retrospective evaluation (Kahneman, Frederickson, Schreiber, & Redelmeier, 1993; Redelmeier & Kahneman, 1996), a sequence of 5 days was added to the end of the

Table 1. Wait Time Sequences (all values are in minutes)

Initial Sequences:			
For Participants Experiencing Small Changes	22, 35, 50, 35, 22		
For Participants Experiencing Large Changes	2, 35, 70, 35, 2		
Manipulated Sequences:	Series A	Series B	Series C
Deteriorating in Small Increments	26, 31, 36, 41, 46	27, 32, 37, 42	29, 34, 39, 44
Improving in Small Increments	46, 41, 36, 31, 26	42, 37, 32, 27	44, 39, 34, 29
Deteriorating in Large Increments	6, 21, 36, 51, 66	11, 26, 41, 56	16, 31, 46, 61
Improving in Large Increments	66, 51, 36, 21, 6	56, 41, 26, 11	61, 46, 31, 16
Final Sequence:			
For All Participants	35, 35, 35, 35, 35		

experiment. The wait time for each of these 5 days was 35 minutes.

After evaluating wait times for all 36 days, participants retrospectively evaluated all of the wait times they had seen in the experiment on a scale from -50 (not bad at all) to +50 (extremely bad).

Results

We first analyzed the ratings participants gave for each of the 36 days as they went through the experiment. To reduce variance caused by idiosyncratic reactions to wait times, participants' judgments during the initial 5-day sequence were used as a baseline. Each participant's judgments on trials 6 through 31 were divided by the average of her or his average judgment on days 1 and 5, days 2 and 4, and her or his judgment on day 3. The results are presented in Figure 1.

As shown in Figure 1, of the participants in the small change condition, those who experienced periodically deteriorating sequences rated wait times more aversive than did those who experienced periodically improving sequences for 11 of the 13 wait times. This proportion (.85) was significantly greater than .50, $\chi^2(1, N = 13) = 4.92$, $p < .05$. This effect was very weak, however. In fact, a 2 (direction of change: improving or deteriorating) x 2 (portion of sequence: days 6-18 and days 19-31) x 13(wait times) Mixed-Factors Analysis of Variance (ANOVA) on the evaluations of participants in the small-difference condition failed to find a difference due to the direction of change, $F\ (1,99) = 0.04$, $MSE = 28.61$, $p > .05$.

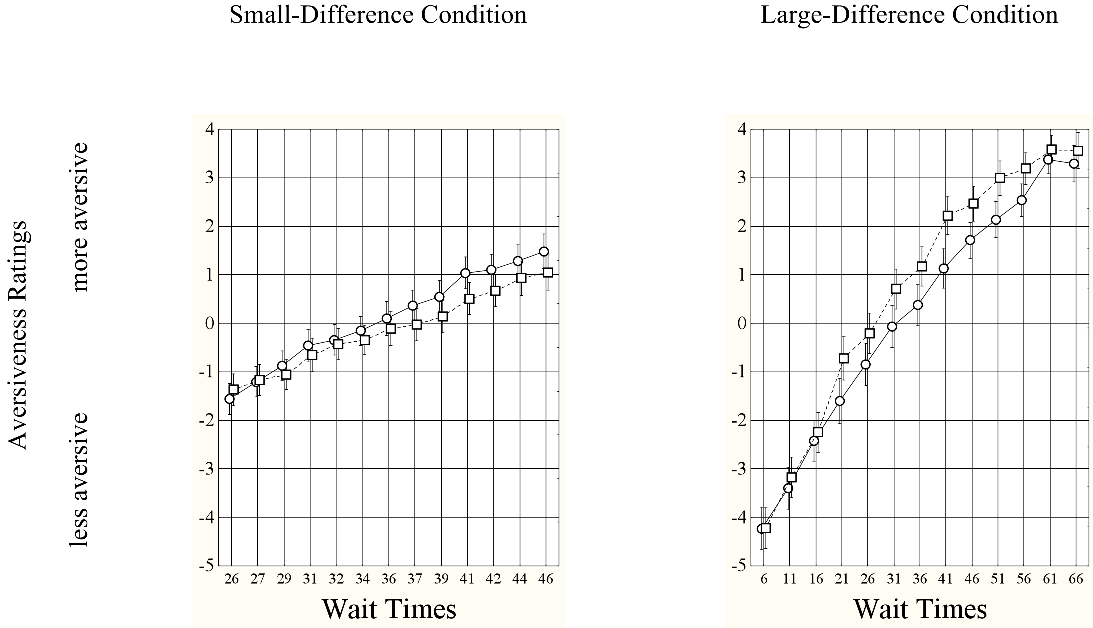
The participants in the large change condition showed a very different pattern of evaluations. Of these participants, those who experienced periodically improving sequences rated wait times more aversive than did those who experienced periodically deteriorating sequences for all 13 of the 13 wait times. This proportion (1.00) was significantly greater than .50, $\chi^2\ (1, N = 13) = 4.92$, $p < .01$. A 2 (direction of change: improving or deteriorating) x 2 (portion of sequence: days 6-18 and days 19-31) x 13 (wait times) Mixed-Factors ANOVA on the evaluations of participants in the large-difference condition also found a main effect of the direction of change, $F\ (1,102) = 8.92$, $MSE = 24.07$, $p < .05$.

The interaction between amount of change and direction of change was also significant as revealed by an omnibus 2 (amount of change: small or large) x 2 (direction of change: improving or deteriorating) x 2 (portion of sequence: days 6-18 and days 19-31) x 13 (wait times) Mixed-Factors ANOVA, $F\ (1,201) = 8.25$, $MSE = 26.3$, $p < .05$. This finding is consistent with the predictions of the CID model presented above and inconsistent with the predictions of adaptation-level theory (Helson, 1964) and range-frequency theory (Parducci, 1995) and not predicted by the consistency model (Haubensak, 1992). Post hoc least significant difference analyses found that participants who experienced periodically improving large differences rated wait times more aversive than did the other three groups. Evaluations on days 19-31 did not significantly differ from evaluations on days 6-18, $F\ (1,102) = 0.11$, $MSE = 2.32$, $p > .05$.

A 2 (amount of change: small or large) x 2 (direction of change: improving or deteriorating) Between-Subject ANOVA on participants' retrospective evaluations showed no main effect of the amount of change [$F\ (1,201) = 3.18$, $MSE = 621.1$, $p > .05$], no main effect of the direction of change [$F\ (1,201) = 0.46$, $MSE = 621.1$, $p > .05$], and no interaction between them [$F\ (1,201) = 0.87$, $MSE = 621.1$, $p > .05$] suggesting that the algorithms responsible for retrospective evaluations might be different from the algorithms responsible for online evaluations. This result also suggests that the finding that people prefer improving to deteriorating sequences (Hsee & Abelson, 1991; Hsee et al., 1991; Schifferstein & Frijters, 1992; Varey & Kahneman, 1992) might not generalize to online evaluations of periodically improving and deteriorating sequences such as the sequences investigated here.

Discussion

Theories of sequence effects on hedonic evaluation were assessed using a paradigm in which values periodically



Note: Error bars represent one standard error about the mean.
Periodically Improving Periodically Deteriorating

Figure 1. Results

improved or deteriorated by small or large amounts. When values changed by small amounts, participants evaluated periodically improving sequences more positively than periodically deteriorating sequences; but when values changed by large amounts, participants evaluated periodically deteriorating sequences more positively than periodically improving sequences.

Helson's (1964) adaptation-level theory cannot explain the finding that evaluations differed based on the size of the difference between exemplars since AL theory makes the prediction that items in deteriorating sequences will always be judged worse than items in improving sequences regardless of the size of the difference between items. Since wait times in the large change deteriorating sequence were preferred over wait times in the large change improving sequence it seems as though participants' evaluations were not based on a comparison of the to-be-evaluated wait time and the average of the preceding wait times, which would result in the opposite finding.

Since the range and frequency values of each wait time in the small change condition mapped onto the range and frequency values of each wait time in the large change condition, range-frequency theory (Niedrich, Sharma, & Wedell, 2001; Parducci, 1995) cannot explain the finding that the amount of change between items influenced wait time evaluations. Furthermore, range-frequency theory cannot explain the finding that wait times were evaluated as better in the large change deteriorating condition than in the large change improving condition; the frequency principle predicts that wait times will be worse in the deteriorating condition than in the improving condition since there would be more positive exemplars in the context of judgment. Based on the current findings, it seems as though range and frequency principles were not used to make evaluations of wait time since these principles would produce the opposite pattern of evaluations.

Haubensak's (1992) consistency model of judgment also cannot explain the finding that the amount of change between wait times influenced evaluations. The consistency model predicts that evaluations are based on linear interpolations from initial items in a sequence. Because linearly interpolating between the evaluations made during the initial sequence would have made evaluations in the periodically improving and deteriorating sequences identical, the finding that there were differences in the evaluated wait times in these sequences suggests that participants did not base their evaluations off of the initial sequence of wait times that was presented.

Of the four evaluation models presented in this paper, only the CID model proposed above was able to predict and explain the observed results. The finding that participants preferred periodically improving to periodically deteriorating sequences when wait times changed by small amounts is consistent with the CID prediction that differences will be exaggerated toward the central tendency of values that have been associated with a comparison word (Huttenlocher et al., 2000). Similarly, the finding that participants preferred periodically deteriorating to periodically improving sequences when wait times changed by large amounts is consistent with the CID prediction that differences will be under appreciated when the central tendency of values associated with a comparison word is smaller than the amount of change.

The results reported here have implications for how managers, price strategists, administrators, politicians, and other bearers of good and bad news ought to present news to others. Consistent with research on hedonic editing (Thaler & Johnson, 1990), if circumstances (e.g., prices, salaries, service quality, and so forth) are going to become better, perhaps it is best to present the good news a little bit at a time. People will appreciate the good news; and they will be likely to exhibit positive emotions such as happiness, delight, and relief each time that good news is presented. If circumstances were going to become worse, however, perhaps it would be best to present all of the bad news at once. People might not realize how bad circumstances have actually gotten; and while they will likely exhibit negative emotions such as anger, frustration, and worry, the sum total of these negative emotions might be less than if the bad news were presented a little bit at a time. The CID model builds on this previous research by offering guidance on the size of the changes that will be underappreciated or exaggerated. Changes that are larger than the central tendency of the distribution of previously observed changes will be underappreciated; and changes that are smaller than this central tendency will be exaggerated.

References

Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research, 24*, 202-214.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*, 839-855.

Choplin, J. M., & Hummel, J. E. (2002). Magnitude comparisons distort mental representations of magnitude. *Journal of Experimental Psychology: General, 131*(2), 270-286.

Choplin, J.M. (2007). Toward a comparison-induced distortion theory of judgment and decision making. In J.A. Elsworth (Ed.), *Psychology of decision making in education, behavior and high risk situations* (pp. 11-40). Hauppauge, NY: Nova Science.

Frederick, S., & Loewenstein, G. (1999). Hedonic Adaptation. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 302-329). New York: NY: Russell Sage.

Haubensak, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 303-309.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Hsee, C. K., & Abelson, R. P. (1991). Velocity relation: Satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology, 60*, 341-347.

Hsee, C. K., Abelson, R. P., & Salovey, P. (1991). The relative weighting of position and velocity in satisfaction. *Psychological Science, 2*, 263-266.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage.

Kahneman, D., Frederickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science, 4*, 401-405.

Kalwani, M. U., Yim, C. K., Rinne, H. J., & Sugita, Y. (1990). A price expectations model of customer brand choice. *Journal of Marketing Research, 27*, 251-262.

Kalyanaram, G., & Winer, R. S. (1995). Empirical Generalizations from reference price research. *Marketing Science, 14*, G161-G169.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Niedrich, R. W., Sharma, S., & Wedell, D. H. (2001). Reference price and price perceptions: A comparison of alternative models. *Journal of Consumer Research, 28*(3), 339-354.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*, 266-300.

Parducci, A. (1995). *Happiness, pleasure and judgment: The contextual theory and its applications*. Mahwah, NJ: Lawrence Erlbaum.

Rajendran, K. N., & Tellis, G. J. (1994). Contextual and temporal components of reference price. *Journal of Marketing, 58*, 22-34.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66*, 3-8.

Schifferstein, H. N. J., & Frijters, J. E. R. (1992). Contextual and sequential effects on judgments of sweetness intensity. *Perception & Psychophysics, 52*, 243-255.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3-21.

Tesser, A., & Martin, L. (1996). The psychology of evaluation. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 400-432). New York: Guilford.

Thaler, R.H. and Johnson, E.J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science, 36*, 643-660.

Varey, C., & Kahneman, D. (1992). Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making, 5*, 169-185.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Wedell, D. 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*, 1158-1173.