Journal of Experimental Psychology: General
2002, Vol. 131, No. 2, 270–286

0096-3445/02/$5.00 DOI: 10.1037//0096-3445.131.2.270

Magnitude Comparisons Distort Mental Representations of Magnitude

Jessica M. Choplin and John E. Hummel
University of California, Los Angeles

Many cognitive processes rely on representations of magnitude, yet these representations are often malleable (H. Helson, 1964; J. Huttenlocher, L. V. Hedges, & J. L. Vevea, 2000; A. Parducci, 1965). It is likely that factors that affect these representations in turn affect the psychological processes that rely on them. The authors conducted 4 experiments to investigate whether language-expressible magnitude comparisons distort mental representations of compared magnitudes. Participants compared magnitudes and estimated those magnitudes in a variety of tasks. Experiments 1 through 3 demonstrated systematic comparison-induced distortions. Experiment 4 demonstrated that comparison-induced distortions might account for the asymmetric dominance effect discussed in the decision-making literature. Potential effects of comparison-induced distortions on other psychological processes (e.g., density effects, order effects, body-size estimation, pain estimation, and consumer decision making) are discussed.

Magnitude comparisons (*less*, *more*, etc.) often reflect our estimates of magnitude. Describing one brand of yogurt as more fattening than another suggests something different from describing the second as lighter than the first; saying, “I am fatter than she is” suggests something different than “She is thinner than I am.” The choice of different descriptions likely reflects a difference in the initial assessments of magnitude. More-than comparison words such as *more, fatter,* and so on are generally used to describe larger magnitudes than less-than comparison words such as *less, thinner,* and so on (Cruse, 1976; Rusiecki, 1985; see also Clark, 1969, and below for a discussion of how these two types of comparisons are related to the concept of markedness). Beyond reflecting initial assessments, however, the comparisons themselves might reify initial assessments of magnitude, distorting the reasoner’s representations of the values being compared. In this article we examine how explicitly comparing two magnitudes can affect the mental representations of those magnitudes.

Numerous researchers have investigated how magnitude values (see Banks, 1977, for a review) and reference points (Holyoak & Mah, 1982) influence magnitude comparisons. For example, under the well-known *semantic distance effect,* people are faster to verify that a horse is larger than an opossum than to verify that a dog is larger than an opossum. In this example, the represented sizes of the animals affect the process of comparing those sizes. (Even more obviously, the value of the magnitude determines the truth value of the relation: $3 > 2$ is true; $2 > 3$ is not.) Much less studied is the possibility that judgments about relations (e.g., “A dog is larger than an opossum”) might affect judgments of the related magnitudes (the sizes of dogs and opossums). Such an effect would be the converse of the semantic distance effect, in which articulating a magnitude comparison affects the representations of the compared magnitudes.

One reason to hypothesize such an effect is that analogous distortions have been observed in the context of visual illusions (although such effects may not require the perceiver to explicitly articulate the relation). For example, in the Ebbinghaus illusion, a test object appears smaller when surrounded by large objects than when surrounded by small objects (see Figure 1). The relation between the sizes of the test object and the surrounding objects produces the illusion.

Another reason to hypothesize such effects in higher cognition comes from a model of transitive inference recently proposed by Hummel and Holyoak (2001). This model reasons about transitive relations (e.g., inferring $a > c$, given $a > b$ and $b > c$) by mapping objects onto locations in a spatial array and “reading off” additional relations (as proposed by Huttenlocher, 1968, and others). Mapping objects to locations in this way assigns magnitudes to the objects based on their stated relations (e.g., assigning magnitudes to a and b based on the stated relation $a > b$) and in the model can distort the representation of the objects’ magnitudes even when they are known beforehand (e.g., making a seem greater and/or b less than they actually are). Specifically, mapping objects to specific values based on their categorical relations is underconstrained. For example, given only the statement $a > b$, the specific values of a and b are undetermined. For this reason, the Hummel and Holyoak model must use heuristics in order to assign objects to likely values based on their stated categorical relations (e.g., given $a > b$, assume that b takes a medium value and a takes a value above that of b). These heuristics operate whenever the model makes magnitude comparisons, even when the actual magnitudes are known. As a result, the model predicts that the act of making magnitude comparisons, which invokes the heuristic machinery for assigning objects to values based on their relations, can

Preparation of this article was supported by National Science Foundation Grant SBR-9709023, by a grant from Hughes Research Laboratories, and by grants from the University of California, Los Angeles (UCLA) Academic Senate. We are grateful to Keith Holyoak for his helpful comments on a draft and to Armanda Pruitt, Vinod Shankar, Mary Huang, Diana Arutyunyan, Shlomit Shlomof, Stanley Chiao, Alina Sgattoni, Athena Yoneda, Barbara Petrovic, Vera Yung, Seam Heng, Elysabeth Nguyen, and Tammy Caplan for their help running experiments.

Correspondence concerning this article should be addressed to Jessica M. Choplin or John E. Hummel, Department of Psychology, 405 Hilgard Avenue, UCLA, Los Angeles, California 90095-1563. E-mail: choplin@psych.ucla.edu or jhummel@psych.ucla.edu

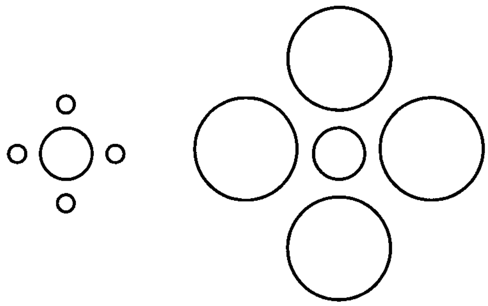

Figure 1. The Ebbinghaus illusion. The center circle appears smaller when surrounded by large objects than when surrounded by small objects.

distort the mental representation of object magnitudes, even when those magnitudes are known beforehand. The resulting distortions tend to be in the direction suggested by the comparison itself: If the actual difference between *a* and *b* is smaller than the difference suggested by the heuristics of the comparison $a > b$, then the model will tend to distort the magnitudes of *a* and *b* in such a way as to make the difference between them greater (e.g., distorting *a* upward and *b* downward); conversely, if the difference between *a* and *b* is larger than the difference suggested by the heuristics, then the model will tend to distort *a* downward and *b* upward. For this reason, we refer to comparison-induced distortions in the mental representation of magnitude as *comparison-induced* or *comparison-suggested* distortions and to the resulting magnitudes as *expected* magnitudes. However, it is important to emphasize that we do not assume that the "expectations" are in any way explicit; rather, we assume that they are a simple by-product of the heuristic algorithms that are invoked during magnitude comparison.

Magnitude comparisons are usually transitive—meaning that *relation* (*x y*) and *relation* (*y z*) implies *relation* (*x z*)—and so fall in the domain of the relations processed by the Hummel and Holyoak (2001) model. Although this model was originally proposed to explain response time and error rate data in transitive reasoning tasks, the more general suggestion that magnitude comparisons might affect represented magnitudes constitutes an important novel prediction with broader implications. One of our motivations for pursuing the research reported in this article was to test this prediction. Because the parameters of the Hummel and Holyoak model were set solely to explain the response times and error rates of participants engaged in transitive reasoning,[1] the experiments described here stand as strong tests of the model.

A third reason to suspect that magnitude comparisons may affect representations of magnitude is that memory for magnitude values is imperfect, and the information provided by a comparison may be useful for estimating the compared values (Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Vevea, 2000). If a person is ignorant of the true value of a variable but knows that it is larger than some other value, then estimating at the higher end of the general range of plausible values might be sensible.

How might comparisons affect mental representations of magnitude? We investigated two possible effects of comparisons on representations of magnitude: comparison-suggested differences and wording effects.

Comparison-Suggested Differences

Articulating a magnitude comparison emphasizes the difference between the arguments of that comparison. It is one thing to know that Albert is 6 ft (1.83 m) tall and Bill is 5 ft (1.52 m) tall; it is another to explicitly think about the fact that Albert is taller than Bill. As such, magnitude comparisons might influence mental representations of magnitudes simply by suggesting differences. Without knowing how tall either of them is, the statement "Albert is taller than Bill" would suggest to most English speakers that the difference between their heights is most likely greater than a millimeter and less than, say, a meter (Allan, 1986–1987; Rusiecki, 1985).[2] We hypothesize that when the comparison-suggested difference differs from the actual difference between the compared magnitudes, the represented magnitudes will be distorted in the direction of the comparison-suggested difference (Hummel & Holyoak, 2001). For example, if it happened that the actual difference between Albert and Bill was only a millimeter, then explicitly thinking about the fact that Albert is taller than Bill may exaggerate the represented difference between them, making Albert seem taller, Bill shorter, or both. Conversely, if it happened that the true difference between Albert and Bill was a meter, then explicitly comparing them might reduce the represented difference. Although these examples are based on extreme deviations from an intuitive comparison-suggested difference in height, we assume that more moderate deviations from comparison-suggested differences may produce analogous effects.

To investigate whether (and how) English speakers associate quantitative differences with comparison terms, Rusiecki (1985) gave participants sentences of the form "Sylvia's husband is shorter than her brother" and "Mary is older than Jane." Participants then filled in sentences such as "Sylvia's husband may be about __ tall" and "Mary is about __ years old" to indicate how tall or old they imagined these people might be. He found that participants generally agreed on the approximate range of the differences. For example, when the comparison terms *older* and *younger* were used, most participants imagined a 2- to 5-year difference, not a 1-month or 40-year difference.

No research to date has directly addressed the question of whether comparison-suggested differences, as discussed by Rusiecki (1985) and Allan (1986–1987), distort representations of known magnitudes. However, the work of Huttenlocher and her colleagues is suggestive in this context (Huttenlocher, Hedges, & Bradburn, 1990; Huttenlocher, Hedges, & Prohaska, 1988). They have found that temporal predicates biased recollection of the time at which autobiographical events occurred. For example, Huttenlocher et al. (1988) asked participants to recall when campus

[1] Specifically, the model accounts for the complex pattern of reaction time and error rate data reported by Sternberg (1980; see Hummel & Holyoak, 2001).

[2] The suggestion need not be in any way explicit. We refer to the difference as a *comparison-suggested difference* in part for lack of a better term and in part because it corresponds roughly with the difference that the comparison might suggest to people when more precise information is not available. However, even when precise information is available, people may (partially) code magnitude information in terms of magnitude comparisons. If so, the default comparison-suggested differences would likely affect representations of known magnitudes as well.

events (i.e., movies) occurred and found that recall was biased toward differences suggested by standard temporal predicates (e.g., toward exactly 7 days ago when the predicate was *week,* 30 days ago when the predicate was *month,* etc.). Huttenlocher et al. (1990) made follow-up calls to people who had participated in an unrelated interview and asked them to recall when the interview had occurred and observed similar effects. The hypothesis motivating the current experiments is that many (if not all) magnitude comparisons, even those that do not represent a specific numerical duration, suggest similar standard differences (Allan, 1986–1987; Rusiecki, 1985) and that these comparison-suggested differences can distort representations of magnitude.

Wording Effects

Comparison words often suggest values because they are used by the speakers of a language to describe relative values within specific ranges of a dimension. DeSoto, London, and Handel (1965, p. 3) presented a particularly striking anecdote to demonstrate this principle. One summer, there was an exhibition baseball game in which two baseball legends, Mickey Mantle and Willie Mays, were playing against one another. Both were having a bad day. Upset over the poor performance of the baseball greats, one fan yelled at Mantle, “I came to see which of you two guys was better, you or Mays. Instead, I’m seeing which is worse!” Apparently, comparisons of the form “*a* is better than *b*” are not used to describe items within the same range of values as comparisons of the form “*b* is worse than *a*,” despite the fact that in a truth conditional sense they are logically equivalent. (See Cruse, 1976, and Lehrer, 1985, for discussions on the semantics of the comparison words *better* and *worse.*)

Analogous examples of truth conditionally equivalent comparison words being used to describe items within different regions of a dimension are ubiquitous. For example, intuitively the price comparison “*c* is more expensive than *d*” suggests higher prices than the price comparison “*d* is cheaper than *c*.” Height comparisons using the word *taller* suggest taller values than those using *shorter.* Age comparisons using *older* suggest older values than those using *younger.* Rusiecki (1985) found empirical evidence to support these intuitions. As described above, he gave participants sentences that compared, for example, heights using the terms *taller* and *shorter* and ages using the terms *older* and *younger* and asked them to imagine the compared values. He found that the imagined values of both compared items were pulled in the direction suggested by the comparison word relative to its antonym, that is, downward for words like *shorter* and *younger* relative to words like *taller* and *older.*

In addition to the general fact that comparison words are used to describe values within specific ranges of a dimension, comparison words may also distort representations by virtue of the fact that the first argument of a magnitude comparison is assigned a value relative to the second (Allan, 1986–1987; Hummel & Holyoak, 2001). In the statement “Bill is shorter than Albert,” Bill’s height is specified relative to Albert’s height. More-than comparison terms imply that the first argument takes a higher value than the second argument; less-than terms imply that the first argument takes a lower value. This assignment of values relative to the second argument may also explain why less-than comparisons result in lower estimates than more-than comparisons.

These considerations suggest the possibility of a second effect of comparison words: If the “mental array” onto which magnitudes are mapped is finite, and if the referent of a less-than comparison is assumed to take a value below the middle of the array (consistent with the downward linguistic bias), then there may be less “room” between the referent and the bottom of the array than there is between the referent of a more-than comparison and the top of the array. As a result, the psychological distance between the arguments of a less-than comparison may be smaller than the psychological distance between the arguments of a more-than comparison (see Hummel & Holyoak, 2001).

Semantic Incongruity Hypothesis

The Hummel and Holyoak (2001) model of transitive inference suggests an important boundary condition on comparison-induced distortions. Specifically, the model predicts no distortions when the represented values are equivalent to the values suggested by comparisons. When the difference between the compared represented values is equivalent to the comparison-suggested difference assumed by the model, no comparison-suggested difference distortion effects would occur. We therefore hypothesize that semantic congruity serves as an important boundary condition on comparison-induced distortions, such that effects are predicted only when the values suggested by comparisons are incongruous with represented values. We call this proposal the *semantic incongruity hypothesis*, because magnitude distortions are hypothesized to be a function of the incongruity between the stated relation and the true difference between the arguments it relates.

The semantic incongruity hypothesis suggests that linguistic asymmetries in how comparisons are used by speakers of a language might produce asymmetries in comparison-induced distortions. For example, describing a tall basketball player as “shorter” than another tall basketball player is generally unacceptable in English (Clark, 1969; Cruse, 1976). The term *short* is semantically incongruous when used in reference to the heights of tall people. By contrast, it is acceptable in English to describe one short person as “shorter” than another. And because *taller* is the default English height comparison term, it is also acceptable to describe one short person as “taller” than another. This asymmetry in how the terms are used suggests that comparison words might distort some values (those for which comparison terms are semantically incongruous with actual values), but not other values (those for which either of two comparison terms is acceptable).

Psychologists have not always appreciated these asymmetries in comparison word use. Hunter (1957), for example, assumed that when people engaged in transitive reasoning they could simply replace a statement such as “*x* is greater than *y*” with a statement such as “*y* is less than *x*,” because the truth conditions of these comparisons are equivalent. Contrary to this assumption, DeSoto et al. (1965) found that participants had directions in which they preferred to reason, usually from the items with the highest values to the items with the lowest values (see also Huttenlocher, 1968) or from the ends toward the middle.

The reason for this preferred directionality, however, remained unclear until Clark (1969) pointed out a distinction between two types of magnitude comparison predicates—*unmarked* and *marked*—and noted a number of characteristics that distinguish the two. For example, unmarked predicates generally name the dimension on which the two are opposites. *Expensive,* for example, names the dimension of expense; *higher* names height, and so forth. (Of course, there are exceptions—e.g., *taller* is unmarked, but it compares values on the dimension of height, not "tallness.") Furthermore, unmarked terms are easier to remember (Clark & Card, 1969) and are learned earlier by children (Donaldson & Wales, 1970).

For the current discussion, the most important distinguishing characteristic between unmarked and marked magnitude comparisons is that unmarked comparisons have two possible meanings, whereas marked comparisons have only one. Unmarked comparisons might simply refer (in a neutral manner) to the difference between two compared items, or they might suggest particular values on the dimension along which they differ. By contrast, a marked predicate has only one meaning, which suggests a particular value on the dimension over which it is defined.[3] The dual meanings of unmarked terms result in a number of asymmetries in how the terms are used. For example, because the comparison word *shorter* is marked and the comparison word *taller* is unmarked, the question "How tall is the short man?" is acceptable, but the question "How short is the tall man?" is not (see Clark, 1969). The semantic incongruity hypothesis suggests that comparison words will be more likely to distort representations of magnitude in the latter case than in the former.

It is important to note that this semantic incongruity hypothesis is very different from the *semantic congruity effect* discussed in the symbolic comparison literature. Most important, the direction of effect is reversed: Under the semantic congruity effect, magnitude representations affect the processing of comparisons; but under the semantic incongruity hypothesis, comparisons affect magnitude representations.

In Experiments 1–3 we tested the effects of comparison-suggested differences and comparison words on estimates of magnitude. In Experiments 1a, 1b, and 2a we investigated the effects of comparison-suggested differences on estimates of magnitude. In Experiments 2a through 2c and 3 we investigated the effects of comparison words on estimates of magnitude. If magnitude comparisons affect representations of magnitude, then it is reasonable to expect that they will affect other psychological processes that depend on these representations. For example, if saying that one brand of yogurt is "lighter" than another makes the number of calories "seem" like less, then the revised estimate of the number of calories may affect other behaviors (i.e., choice, consumption, etc.). To test this hypothesis, in Experiment 4 we investigated how magnitude comparisons affect other psychological processes that depend on magnitude representation. In particular, we investigated whether magnitude comparisons can help explain the asymmetric dominance effect (ADE) observed in the judgment and decision-making literature (e.g., Huber, Payne, & Puto, 1982; Simonson, 1989; Wedell, 1991).

Experiment 1: Comparison-Suggested Differences

In Experiment 1 we investigated whether magnitude comparisons distort memory for specific magnitudes toward comparison-suggested differences. On each trial, a participant viewed either a circle or a triangle; after it was taken away, the participant recalled its width from memory. The triangles, as a category, were always slightly larger (both in height and width) than were the circles or vice versa. We used these arbitrary geometric shapes because they represented categories for which participants presumably did not hold base-rate size beliefs. To the extent that the actual difference between the circles and triangles differed from the comparison-suggested difference, we predicted that comparisons would bias recall of size to reflect expectations. Without knowing what the comparison-suggested differences are for any given relation or dimension, it is difficult to predict specifically what these bias effects will be. However, for sizes that are sufficiently similar, it is reasonable to assume that the comparison-suggested differences will be greater than the actual differences. We therefore kept the size difference between the circles and the triangles relatively small and predicted that comparisons would increase the difference between recalled size with successive comparisons. In Experiment 1a, participants periodically made judgments from memory comparing the relative sizes of the circles (as a group) to those of the triangles (as a group). Experiment 1b served as a control in which participants judged the widths of individual stimuli but made no explicit comparisons between them. We predicted a general increase in the difference between recalled sizes in Experiment 1a, but no such increase in Experiment 1b.

Experiment 1a

Participants viewed circles and triangles in four blocks of trials. On each trial, after the circle or triangle was removed from sight, the participant recalled its width from memory. For any given participant, the circles were larger than the triangles or vice versa. After each block, the participant made an explicit judgment as to which figures were larger (circles or triangles) and rated his or her confidence in this judgment.

Method

Participants. The experimenter approached individual people on the University of California, Los Angeles (UCLA) campus and asked them to participate in the experiment. One hundred forty people agreed to participate after being approached in this manner. All participants served voluntarily and successfully completed the experiment.

Materials and procedure. The stimuli consisted of four geometric figures of varying widths. The largest figure was 80 mm wide. This figure was scaled down to 80% of its original size to produce a second figure 64 mm wide. The second figure was scaled down to 64% of its original size (i.e., 80% $\times$ 80%) to produce a figure 41 mm wide, which was scaled down by 80% to produce a fourth figure 33 mm wide. For half of the participants, the two largest figures (80 and 64 mm) were circles and the two smallest figures (41 and 33 mm) were triangles. For the other half of the participants, the larger figures were triangles and the smaller figures were circles. Each figure was printed on its own 8.5 $\times$ 11 in. (21.6 $\times$ 27.9 cm) page of paper.

The four figures were presented one at a time in the following order: 33, 80, 41, and 64 mm. The experimenter presented each figure to the participant, let him or her look at it, then took it away. Participants recalled the width of each figure immediately after it was removed from sight. Participants reported the size of a figure by marking a line so that the distance

[3] The majority of marked terms suggest small values and unmarked terms suggest large values, but there are exceptions. Some marked terms suggest large values. Occasionally, there will be two marked terms, one suggesting small values and the other suggesting large values. See Allan (1986–1987).

between a stop mark on the left end of the line and the mark made by the participant was equal to the participant's estimate of the distance across the test object at its widest extent (i.e., the diameter of the circle or the base of the triangle; see Coren & Girgus, 1972). Following the presentation of the four figures, the participant was asked to circle either the sentence "Circles are bigger than triangles" or the sentence "Triangles are bigger than circles." Participants were asked to indicate how certain they were by circling "Not sure," "Sure," or "Very sure." To observe how magnitude comparison judgments affect recall of magnitude as participants have progressively more experience with the stimuli, this sequence was repeated four times, for a total of 16 presentations and four magnitude comparisons.

A booklet was prepared for each participant. Each page of the booklet was approximately 5 cm high and 21.6 cm wide. This booklet had 20 pages (4 pages to reproduce the widths of the objects followed by 1 for a magnitude comparison, with the sequence repeated four times). The pages used to reproduce the widths of objects had a horizontal line 14 cm long with a 1-cm vertical stop on the left end. The pages used to query participants had the two alternative magnitude comparisons (i.e., "Circles are bigger than triangles" or "Triangles are bigger than circles") at the top. Participants indicated the certainty of their judgment immediately below the magnitude comparison statements. Participants leafed through the booklet 1 page per trial, indicating each response on a separate page.

Results and Discussion

Participants' mean recalled widths of the four figures on successive blocks are shown in Figure 2. The data show a general spreading of recalled sizes, such that larger figures were recalled as progressively larger on successive blocks and smaller figures were recalled as progressively smaller. A 4 (block) $\times$ 4 (stimulus size) within-subjects analysis of variance (ANOVA) revealed that this interaction was significant, $F(9, 1251) = 29.16$, $MSE = 26.39$, $p < .01$. A Tukey post hoc analysis ($\alpha = .05$) revealed that the 33-mm figure was recalled as smaller on the last observation than on the first and the 80-mm figure was recalled as larger on the last observation than on the first, as was the 64-mm figure. The Tukey analysis failed to reveal any differences between the first and last observations of the 41-mm figure. Linear trend analyses revealed highly reliable linear trends for all four figures. Linearly increasing trends were observed for the 80-mm figure, $F(1, 139) = 90.64$, $MSE = 39.97$, $p < .01$, and the 64-mm figure, $F(1, 139) = 15.03$, $MSE = 20.9$, $p < .01$. Linearly decreasing trends were observed for the 33-mm figure, $F(1, 139) = 135.30$, $MSE = 11.03$, $p < .01$, and the 41-mm figure, $F(1, 139) = 8.95$, $MSE = 57.28$, $p < .01$.

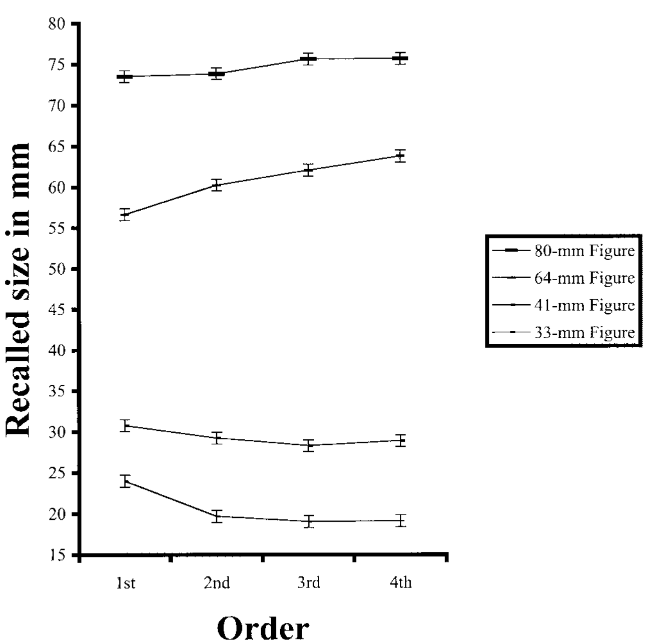

Figure 2. Results of Experiment 1a. When participants made magnitude comparisons, recalled sizes diverged in each of the four successive blocks. Error bars are 95% within-subjects confidence intervals described by Loftus and Masson (1994).

As predicted, magnitude comparisons gave rise to a spreading of recalled sizes. The difference between the recalled sizes of the larger figures and the recalled sizes of the smaller figures increased on each successive block as the participant articulated more magnitude comparisons.

Experiment 1b

A second group of participants was exposed to the same stimuli and asked to perform the same size judgment task as in Experiment 1a. However, participants in Experiment 1b were never asked to compare the sizes of the figures. If the explicit magnitude comparisons articulated in Experiment 1a were responsible for the progressive spreading of recalled sizes, then participants in Experiment 1b should not show the same progressive distortion in their recalled sizes. Instead, recalled sizes should remain roughly flat over successive presentations. By contrast, if simply viewing the stimuli repeatedly is sufficient to distort the mental representation of size—that is, if magnitude comparisons were not responsible for the distortions observed in Experiment 1a—then Experiment 1b should replicate the spreading of recalled sizes observed in Experiment 1a.

Method

Participants. The experimenter recruited individual participants in the UCLA dormitories. One hundred forty people agreed to participate in the experiment after the experimenter knocked on their dormitory doors and asked them to participate. All participants served voluntarily and successfully completed the experiment.

Materials and procedure. The stimuli and presentation order were identical to those of Experiment 1a, except participants in this experiment were never asked to compare the sizes of the circles and triangles. The sequence of figures was simply repeated four times, to produce a total of 16 presentations. The booklet prepared for this experiment contained 16 pages (1 page for each figure width reproduction).

Results and Discussion

The average recalled widths of the four figures in each of the four successive blocks are shown in Figure 3. A general reduction in the size of the difference between recalled widths of wide and narrow figures was observed, with wider figures recalled as progressively smaller on successive blocks and narrower figures as progressively larger. A 4 (block) $\times$ 4 (stimulus size) within-subjects ANOVA on recalled width revealed that this interaction was significant, $F(9, 1251) = 4.31$, $MSE = 36.9$, $p < .01$. A Tukey post hoc analysis ($\alpha = .05$) revealed that the 33-mm figure was recalled as wider on the last observation than on the first and the

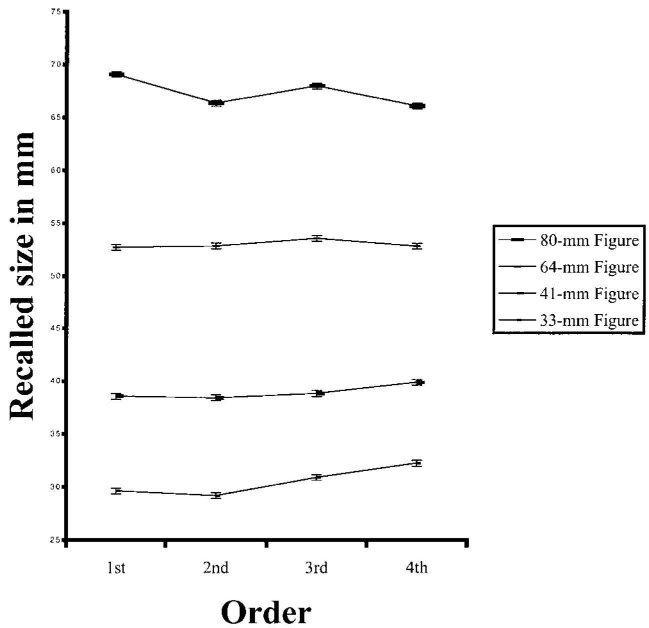

Figure 3. Results of Experiment 1b. When participants did not make magnitude comparisons, a general convergence of recalled sizes was observed. Error bars are 95% within-subjects confidence intervals described by Loftus and Masson (1994).

80-mm figure was recalled as narrower on the last observation than on the first. The Tukey analysis failed to reveal a difference between the recalled widths of the first and last observations of the 41-mm or 64-mm figures. A linearly decreasing trend was observed for successive recalled widths of the 80-mm figure, $F(1, 139) = 4.76$, $MSE = 79.53$, $p < .05$. A linearly increasing trend was observed for successive recalled widths of the 33-mm figure, $F(1, 139) = 11.99$, $MSE = 52.93$, $p < .01$. A marginally reliable linearly increasing trend was observed for successive recalled widths of the 41-mm figure, $F(1, 139) = 3.65$, $MSE = 35.78$, $p < .06$. Trend analysis failed to reveal a linear trend for successive recalled widths of the 64-mm figure ($F < 1$).

These results demonstrate that simply viewing the stimuli and recalling their widths is not sufficient to produce the increase in the difference between recalled widths observed in Experiment 1a. Indeed, the results of Experiment 1b revealed the opposite trend: When participants did not articulate magnitude comparisons, we observed a general decrease in the difference between recalled widths on successive blocks.

Between-experiments analyses were conducted contrasting recalled widths of the four figures in Experiments 1a and 1b. A 2 (experiment) $\times$ 4 (block) mixed-factors ANOVA revealed a significant interaction between the recalled widths of the 80-mm figure in the two experiments such that in Experiment 1a participants' recalled widths increased in each succeeding block, but in Experiment 1b recalled widths decreased in each succeeding block, $F(3, 834) = 5.62$, $MSE = 55.92$, $p < .01$. Analysis of the recalled widths of the 64-mm figure revealed a significant interaction such that in Experiment 1a recalled widths increased in each succeeding block, but they remained unchanged across succeeding blocks in Experiment 1b, $F(3, 834) = 18.81$, $MSE = 31.65$, $p < .01$. Analysis of the recalled widths of the 41-mm figure revealed a significant interaction such that in Experiment 1a recalled widths decreased in each succeeding block, but in Experiment 1b they increased in each succeeding block, $F(3, 834) = 6.25$, $MSE = 23.33$, $p < .01$. Finally, analysis of the 33-mm figure revealed a significant interaction such that in Experiment 1a recalled widths decreased in each succeeding block, but in Experiment 1b they increased in each succeeding block, $F(3, 834) = 33.43$, $MSE = 22.48$, $p < .01$.

The increase in the differences between recalled sizes when participants articulated magnitude comparisons (Experiment 1a) strongly contrasts with the decrease in the difference between recalled sizes in the absence of such comparisons (Experiment 1b). This difference between the experiments makes the effects of articulating magnitude comparisons especially salient.

Discussion, Experiments 1a and 1b

Experiments 1a and 1b demonstrated that magnitude comparisons can distort mental representations of magnitude. We interpret these distortions as projections toward comparison-suggested differences between the arguments of magnitude comparisons. The differences suggested by the magnitude comparisons apparently distorted participants' mental representations of the magnitudes.

As noted earlier, these distortions are analogous to bias in recall as studied by Huttenlocher and her colleagues, in that conceptual information is used to reestimate values after memory of exact values has faded (Huttenlocher et al., 1988, 1990, 1991, 2000). Experiments 1a and 1b extend the line of research started by Huttenlocher and her colleagues to cover magnitude comparisons. First, the information contained in the one-place attributional predicates used by Huttenlocher et al. (1991, 2000) cannot explain the results of Experiment 1a. If our participants had relied on their knowledge of the circle and triangle categories alone (i.e., knowledge of the central tendency of these categories), then biases in recall would have been toward the means of each category. The recalled sizes of the triangles would have moved toward each other as would the recalled sizes of the circles. The recalled widths of the middle-sized objects (64-mm and 41-mm figures) would have diverged (as we observed), but the recalled widths of the extreme-sized objects (80-mm and 33-mm figures) would have converged (contrary to our findings). Moreover, on the basis of one-place predicates, there is no reason to expect the observed differences between the results of Experiments 1a and 1b. Our results—specifically, the finding that articulating comparisons causes recalled widths to move progressively farther from the contrasting category and thus farther from the individual category means—suggest that participants relied on relational information to reestimate the figures' widths.

Magnitude comparisons are similar to temporal relations (e.g., as studied by Huttenlocher et al., 1988, 1990) in that both are relational. They differ, however, in that magnitude comparisons can take a wider array of arguments (any quantitative attribute, not just time). They also differ in that Huttenlocher and colleagues' (1988, 1990) findings can be described as a bias toward a specific numerical duration, such as "1 month before," and so on. By contrast, our findings are based on judgments of categorical relations (e.g., *larger* or by analogy to temporal predicates *later* or *earlier*), which apply over an infinite range of values. In spite of the fact that they do not explicitly state any particular value difference (e.g., *larger* does not specify how much larger), our findings suggest that the mind nonetheless imposes an interpreted value on the relation (Rusiecki, 1985) and that the interpreted values bias memory of magnitudes (as predicted by Hummel & Holyoak, 2001). In this sense, Experiments 1a and 1b constitute an important extension of the research of Huttenlocher and her colleagues.

In Experiment 1a, the difference between recalled sizes increased after successive comparisons. It is important to note that our account, which is based on comparison-suggested differences, predicts the opposite trend when the actual differences between stimuli are larger than the comparison-suggested difference. That is, if the actual differences are sufficiently large, then explicit comparisons should cause the difference between estimates to decrease. Choplin and Zhang (2001) demonstrated this kind of convergence in a marketing context and showed that it affects consumer decision making.

Experiments 2 and 3: Comparison Word Effects

Experiment 1 demonstrated an effect of magnitude comparisons on memory of magnitudes. In Experiments 2 and 3 we investigated whether comparison words affect memory of magnitudes: Are Albert and Bill believed to be taller if their heights are compared by saying "Albert is taller than Bill" than if they are compared by saying "Bill is shorter than Albert?" Experiments 2 and 3 manipulated comparison words to see whether the arguments of less-than comparisons are judged to be lower in magnitude than the arguments of more-than comparisons. Experiment 2a also served to replicate the comparison-suggested difference effects observed in Experiment 1a, and Experiment 3 served to replicate the effects observed in Experiment 2a and 2b under different stimulus conditions. Experiment 3 also investigated whether less-than comparisons suggest a smaller difference between their arguments than do more-than comparisons.

Experiments 2a–2c

In Experiments 2a through 2c, participants viewed three stimuli—a small circle, a medium-sized triangle, and a large square. We then removed the stimuli and asked our participants to compare the size of the medium-sized triangle with the size of one of the other two shapes. This comparison used either the term *smaller* or the term *bigger.* We then asked our participants to recall the sizes of the three shapes from memory. In Experiment 2a we investigated how these comparisons affected the recalled size of the medium-sized triangle, in Experiment 2b we investigated the effects of the comparisons on the recalled size of the large square, and in Experiment 2c we investigated the effects of the comparisons on the recalled size of the small circle. We predicted that participants who compared sizes using *smaller* would recall the sizes as smaller than participants who compared them using *bigger.* We also predicted that in Experiment 2a distortions toward comparison-suggested differences would produce biases in recall of the size of the triangle. Assuming the actual differences in the shapes' sizes are smaller than comparison-suggested differences, we predicted that recalled size of the triangle would be biased away from the size of the shape with which it was compared. Participants who compared the medium-sized triangle with the small circle were predicted to recall the triangle as larger than would participants who compared the triangle with the large square.

Experiment 2a

In Experiment 2a, participants recalled the size of the triangle after comparing it with either the circle or the square using either the comparison word *smaller* or the comparison word *bigger.*

Method

Participants. The experimenter approached individual people either on the UCLA campus or at a West Los Angeles shopping mall. After being approached in this manner, 189 people agreed to participate in the experiment. The data from 29 participants were discarded (26 because they judged the true magnitude comparison statement as false, 2 because they left it blank, and 1 because the response was illegible), leaving 160 responses for analysis. Of these participants, 81 compared the triangle with the square (40 using the word *bigger* and 41 using the word *smaller*) and 79 compared the triangle with the circle (32 using the word *bigger* and 47 using the word *smaller*). All participants served voluntarily.

Materials and procedure. Three shapes (a circle with a 20-mm diameter, a triangle with a 40-mm-wide base, and a 55-mm-wide square) were printed on an 8.5 × 11 in. (21.6 × 27.9 cm) piece of paper. The circle was presented on the left side of the page, the triangle in the center, and the square on the right. Participants were instructed to memorize the entire page as they viewed these stimuli from a distance of approximately 1 m. After they indicated that they were ready, the page was removed and the participant was given a half-page (8.5 × 5.5 in.; 21.6 × 14.0 cm) questionnaire presenting a magnitude comparison sentence. This sentence compared the triangle with either the circle or the square using either the term *smaller* or the term *bigger.* As a means of ensuring that participants had completely processed the magnitude comparison, they were instructed to judge whether the sentence was true or false and circle their answer on the questionnaire. (The sentence was always true.)

At the bottom of the questionnaire, participants were instructed to recall the size of the triangle using a page depicting 14 triangles of increasing width. The 1st triangle (labeled *1*) had a base 30 mm wide. Successive triangles (labeled *2* through *14*) increased in width by 2 mm, giving the last triangle (labeled *14*) a width of 56 mm and the 6th (labeled *6*) a width of 40 mm (the true size of the original triangle; i.e., "6" was the correct response). Participants indicated the size they remembered the triangle to be by writing down the label of the triangle closest in size to their memory of the triangle.

Results and Discussion

A 2 (comparison item: circle or square) × 2 (comparison word: *smaller* or *bigger*) between-subjects ANOVA revealed a main effect of comparison words: When the triangle was compared with another object using *smaller*, participants picked smaller comparison triangles ($M = 40.34$ mm) than when it was compared using *bigger* ($M = 43.08$ mm), $F(1, 156) = 10.23$, $MSE = 7.21$, $p < .01$. This result supports the prediction that comparison words would affect recalled size. The ANOVA also revealed a main effect of the comparison object: When the triangle was compared with the circle, participants picked larger comparison triangles ($M = 43.14$ mm) than when the triangle was compared with the square ($M = 40.28$ mm), $F(1, 156) = 11.01$, $MSE = 7.21$, $p < .01$. If it is assumed that the actual differences between the stimuli are less than the differences suggested by the comparison, this effect supports the prediction that comparison-suggested differences affect estimates of magnitude. No interaction was revealed ($F < 1$), suggesting that the two effects are independent.

Experiment 2b

In Experiment 2b, participants recalled the size of the square. We were concerned that the procedure used in Experiment 2a might have resulted in demand characteristics. In particular, the sentence verification (although it was manipulated between subjects) could have placed demands on participants to respond in particular ways. As a means of avoiding these demand characteristics, the procedure used in Experiments 2b and 2c differed slightly from the procedure used in Experiment 2a. Rather than verifying a sentence, participants reported which shape was *smaller* (or *bigger*).

Method

Participants. The experimenter approached individual people on the UCLA campus and at a local pedestrian mall. After being approached in this manner, 203 people agreed to participate in the experiment. The responses from 4 participants were discarded (3 because they incorrectly thought the triangle was larger than the square and 1 whose response form was left blank), leaving 199 responses for analysis. Of these, 98 compared the square with the triangle using the word *smaller* and 101 made the comparison using the word *bigger.* All participants served voluntarily.

Materials and procedure. Participants viewed the circle, triangle, and square on the first page. After the participant indicated that he or she had viewed the shapes, the page was taken away, and the participant was given a distractor task that consisted of rating a "really bad analogy" on a scale ranging from 1 (*not so bad*) to 7 (*the worst*). After rating the analogy, participants were asked either "Which shape was smaller?" or "Which shape was bigger?" and circled either "square" or "triangle." The distractor task and the form of the question were intended to avoid demand characteristics associated with the comparative question by making the question appear to be a recall task. Finally, a page of comparison squares was presented. This page had 11 squares of increasing width. The 1st square (labeled *1*) was 44 mm wide. Successive squares (labeled *2* through *11*) increased in width (and height) by 2 mm, giving the last (labeled *11*) a width of 64 mm and the 4th a width of 50 mm (the true width of the square). Participants indicated what size they remembered the square to be by indicating the label of the square closest in size to their memory.

Results and Discussion

An independent-samples t test revealed an effect of comparison words such that participants who recalled which shape was *smaller* picked smaller comparison squares ($M = 50.54$ mm) than those who recalled which shape was *bigger* ($M = 54.08$ mm), $t(197) = 7.25$, $p < .01$. As in Experiment 2a, comparison words affected participants' estimates of magnitude.

Experiment 2c

Experiment 2c was identical to Experiment 2b except participants compared the circle with the triangle and recalled the size of the circle.

Method

Participants. The experimenter approached individual people on the UCLA campus; 206 people agreed to participate. The responses of 4 participants were discarded (all 4 because they incorrectly recalled that the circle was larger than the triangle), leaving 202 responses for analysis. Of these, 101 compared the circle with the triangle using the word *smaller* and 101 used the word *bigger.* All participants served voluntarily.

Materials and procedure. The materials and procedure were identical to those used in Experiment 2b except that participants compared the circle and the triangle. After rating the analogy, participants were asked "Which shape was smaller?" or "Which shape was bigger?" and circled either "triangle" or "circle." The page of comparison shapes had 11 circles of increasing width. The first (labeled *1*) had a diameter of 8 mm. Successive circles (labeled *2* through *11*) increased in diameter by 2 mm, giving the last circle (labeled *11*) a diameter of 28 mm and the 7th a diameter of 20 mm (the true width of the circle presented on the first page).

Results and Discussion

An independent-samples t test did not reveal an effect of comparison words: Participants who were asked to recall which shape was *smaller* picked comparison circles ($M = 18.64$ mm) that were not reliably different from those picked by participants who recalled which shape was *bigger* ($M = 19.06$ mm; $t < 1$). If the true comparison word effect for the circle had been as large as the observed comparison word effect for the triangle (2.74-mm difference), then we would have had sufficient power to detect a difference (power = .99).

Although it is risky to speculate about the causes of a null effect (especially because other factors such as shape of target object, source and number of participants tested, etc. varied across these three experiments), the failure to observe a comparison word effect for the circle is consistent with the semantic incongruity hypothesis. Note that the comparison term *smaller* is marked, and the comparison term *bigger* is unmarked, and therefore neither term creates incongruity when used in reference to the size of the circle. The marked comparison "The circle is smaller than the triangle" does not create an incongruity because the circle is already small, and the unmarked comparison "The triangle is bigger than the circle" does not create an incongruity because the term *bigger* is unmarked. By contrast, the marked comparison term smaller does create an incongruity when used to compare the size of the square because the square is not small. Consistent with this hypothesis, the effect sizes (i.e., the difference in means) across Experiments 2a through 2c scaled with object size. The effect is largest for the largest shape (the square; Experiment 2b), smallest for the smallest shape (the circle; Experiment 2c), and in the middle for the middle shape (the triangle; Experiment 2a). In future work this semantic incongruity hypothesis should be investigated with more precision.

Experiments 2a and 2b demonstrated that comparison words affect recall of magnitudes. In Experiment 3 we investigated whether comparison words would also affect more immediate judgments of magnitude, namely, magnitude ratings.

Experiment 3

In Experiments 1 and 2 participants were required to recall magnitudes held in memory for several seconds. In Experiment 3 we investigated whether comparison words can affect magnitude judgments even when the stimuli remain in view. If comparisons can influence representations of magnitude in such cases, then their influence on other psychological processes might very well be ubiquitous. Experiment 3 also served to test a specific prediction of the Hummel and Holyoak (2001) model of transitive inference—namely, that a less-than comparison of two magnitudes suggests a smaller difference than a more-than comparison of the same arguments.

In Experiment 3, we gave participants a questionnaire presenting two models of VCR—a "new" model and an "old" model—with a list of their attributes. The new model was superior to the old model on every attribute except the length of the warranty (3 years for the old model vs. 2 for the new model). All attributes of the VCRs were held constant across conditions. All we manipulated was the comparison word of a statement comparing the lengths of the warranties: In one condition the old warranty was described as "longer than" the new warranty, and in the other condition the new warranty was described as "shorter than" the old warranty. The participants' task was to rate the length of the 2- and 3-year warranties as compared with the average length of a VCR warranty.

Method

Participants. The experimenter approached individual participants either on the UCLA campus or at a local pedestrian mall. After being approached in this manner, 94 people agreed to participate in the experiment. The data from 1 person were discarded because critical items were left blank, leaving 93 responses for analysis—49 in the more-than and 44 in the less-than condition. All participants served voluntarily.

Materials and procedure. Participants received one of two 1-page questionnaires. When a party of 2 or more volunteered, each member was given a separate questionnaire. They were told that each questionnaire would be different and were asked not to look at each other's questionnaires until everyone had completed the task. The questionnaires were identical in all respects except the wording of a description (as detailed below). There were two boxes at the top of each questionnaire. The box on the left was labeled *Our Old Model* and listed four attributes of the old model: three irrelevant attributes (two-speed fast-forward and reverse, 8 programmable recording schedules, two recording speeds) and the critical attribute (3-year warranty). The box on the right was labeled *Our New Model* and listed four analogous attributes of the new model: three irrelevant attributes (four-speed fast-forward and reverse, 16 programmable recording schedules, three recording speeds) and the critical attribute (2-year warranty). A short description of the differences between the models appeared below the boxes. On one questionnaire (the less-than comparison condition), this description began, "Although our new model has a shorter warranty than our old model . . . " and went on to list the virtues of the new model. On the other questionnaire (the more-than comparison condition), this description began, "Although our old model has a longer warranty than our new model . . . " and went on to list the virtues of the new model just as in the less-than comparison condition. The word used to compare the warranties (i.e., *longer* vs. *shorter*) and the necessary changes in word order were the only difference between the two questionnaires. Below the description of the differences between the two models, participants rated the length of the warranties (as compared with average VCR warranties) on 6-point scales (0 = *not long*, 5 = *very long*).

Results and Discussion

A 2 (comparison word) × 2 (the two warranties) mixed-factors ANOVA revealed a main effect of comparison word such that the lengths of the warranties (both 2-year and 3-year) were judged to be shorter when compared using the term *shorter* ($M = 2.65$) than when compared using *longer* ($M = 3.02$), $F(1, 91) = 4.23$, $MSE = 1.52$, $p < .05$. This effect replicates the comparison word effects of Experiments 2a and 2b. Because magnitude information was visible throughout the task, this result demonstrates that comparison-altered magnitude judgments generalize beyond delayed recall. However, because ratings are easily changed, sometimes without corresponding changes in representations, this result leaves open the possibility that only magnitude judgments, rather than representations, are immediately affected.

The ANOVA failed to reveal an interaction between comparison word and warranty policy on estimates of length ($F < 1$). The results of Experiment 3 therefore do not support the prediction that less-than magnitude comparisons suggest a smaller difference between their arguments than do more-than comparisons (Hummel & Holyoak, 2001).

Discussion, Experiments 2a through 2c and 3

Experiments 2a, 2b, and 3, like Experiment 1, demonstrated that magnitude comparisons can distort mental representations of magnitude. Specifically, they demonstrated that less-than comparison terms sometimes produce lower value estimates than do more-than comparison terms. Experiments 2a and 2b demonstrated comparison word effects on delayed recall of size. Experiment 3 demonstrated a comparison word effect on judgments of magnitude using a task in which magnitude information was continuously visible. Together, these findings suggest that magnitude comparisons may influence evaluations and representations of magnitude in a wide variety of tasks and contexts.

Asymmetric Dominance Effect

Experiments 1 through 3 investigated the effects of comparisons on representations of magnitude using explicit recall and ratings of magnitudes. However, representations of magnitude are most useful not as a basis for making explicit magnitude judgments but as a basis for other psychological processes, such as decision making, planning, and so forth. The remainder of this article is concerned with the question of whether distortions in magnitude representations—caused by magnitude comparisons—occur in these other tasks and processes. The basic idea is that people make comparisons in many situations and that these comparisons might distort their mental representations of the magnitudes involved. In turn, the magnitude distortions may affect subsequent operations performed on those magnitudes. For example, comparing the price of a generic brand to the price of a name brand might distort the difference between brands by evoking a comparison-suggested difference between them and distorting the difference toward the comparison-suggested difference (e.g., causing a difference of several pennies to be recalled as significantly more). Comparison words might also exert an influence, distorting the magnitude representation of both brands upward or downward. If so, then the resulting distortions could affect how people ultimately choose between the products.

We explored this possibility by concentrating on a single case, the ADE, which has been studied extensively in the decision-making literature (e.g., Huber et al., 1982). In studies of the ADE, the participant is given a choice between two options (labeled C_1 and C_2 in Figure 4). Each option is strong on one dimension and weak on another. For example, C_1 in Figure 4 is strong on

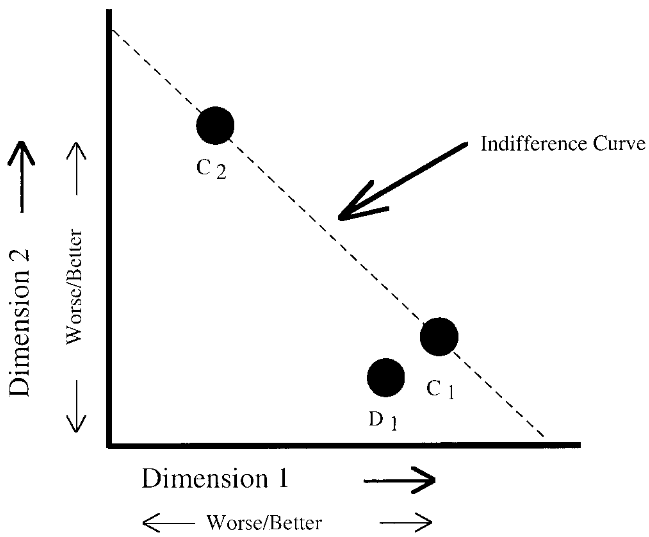

Figure 4. The asymmetric dominance effect. C_1 and C_2 represent choice items, and the dashed line represents the indifference curve. Introducing a decoy (D_1, in this figure) that is similar but inferior to a choice item (C_1, in this figure) biases people to pick that item.

Dimension 1 but weak on Dimension 2, whereas C_2 is strong on Dimension 2 but weak on Dimension 1. The values of C_1 and C_2 on the two dimensions are calibrated to make people equally likely to choose (i.e., indifferent with respect to) C_1 versus C_2. The dashed line in Figure 4 depicts the *indifference curve*—the curve along which people are indifferent to the options because the value gained on one dimension by moving to adjacent options is exactly equal to the value lost on the other dimension. Introducing a decoy item (such as D_1 in Figure 4) that is similar to, but clearly worse than (i.e., *asymmetrically dominated* by), one of the two alternatives biases people to choose the option that dominates the decoy (in this example, C_1). For example, Simonson (1989) gave his participants a choice of three supermarkets: A, B, and C. Supermarket A was more convenient than Supermarket C, but C had more variety than A. A and C were calibrated to lie on the same indifference curve. Supermarket B was similar to and slightly inferior to either A or C (for different participants) on both dimensions. When Supermarket B was similar to (dominated by) C, it biased participants to choose C over A, and vice versa.

Several accounts of this effect have been offered. According to Simonson (1989; see also Pettibone & Wedell, 2000; Wedell, 1991; Wedell & Pettibone, 1996), relations among the alternatives bias choice by providing a "justification" for choosing the dominating option. Because the decoy is clearly worse than one of the options, the decoy makes it possible for the participant to verbalize a reason for choosing that option. Verbalizing a reason reduces indecision and biases choice toward the option justified by the reason. In support of this account, Simonson found that the effect was stronger when participants were required to justify their choices. It is not clear, however, why these justifications ought to serve as justification: Options A and C, in themselves, remain unchanged by the characteristics of B.

Another account (the *value-shift* model; Pettibone & Wedell, 2000; Wedell & Pettibone, 1996) suggests that the perceived attractiveness of values becomes shifted because of changes in the range and ordinal rankings of options.[4] In Figure 4, for example, introducing D_1 makes C_1 appear better on Dimension 2 for two reasons. First, D_1 extends the range of values on Dimension 2. As a result, C_1's value on Dimension 2 is not at the very bottom of the range. Second, placing the decoy next to C_1 as in Figure 4 makes C_1 the second best value out of three on Dimension 2. If instead the decoy was placed next to C_2, then C_1 would be the third value out of three—the worst value—on Dimension 2. The value-shift model assumes that these changes increase the attractiveness of C_1. Because the critical factor in this value-shift account is the change in range and ordinal rankings, it predicts shifts in judgment or ratings of value but does not predict biases in recall of values from memory (D. H. Wedell, personal communication, April 16, 2001).

Consistent with Wedell and colleagues' (Pettibone & Wedell, 2000; Wedell & Pettibone, 1996) value-shift model, we hypothesize that the decoy alters the attractiveness of values. However, in contrast to Wedell and colleagues, we suggest an alternative mechanism that may be responsible for these alterations in attractiveness. We hypothesize that the proximity of the decoy to the dominating option causes the reasoner to compare them, and that the comparisons distort the reasoner's mental representations of the dominating and dominated options (see Figure 5). Specifically, when decision makers compare the decoy option (D_1 in Figure 5) to the option that dominates it (C_1 in Figure 5), the comparison changes the represented values of both options. If the difference is less than suggested by the comparison, then distortions caused by comparison-suggested differences would push the represented value of the choice item (C_1) outward, above the indifference curve, and those of the decoy D_1 downward. As a result, the distorted C_1 lies on a higher indifference curve than does C_2, making it the more attractive alternative. Analogous effects would be found when a D_2 decoy item is introduced instead of D_1. Comparison words would also likely make a difference. When larger values are desirable, more-than comparison terms will make the dominating option more attractive. When smaller values are desirable, less-than terms will make the dominating option more attractive.

Our comparison-induced distortion model makes a number of predictions, some of which are consistent with previous models, and others of which are unique to our model. In its emphasis on the role of explicit comparisons, our model is like that of Simonson (1989). It predicts (along with Simonson) that choice will become biased only when the choices are explicitly compared. However, in contrast to Simonson's model, our model predicts no special role for verbalizable justifications. Like the model of Wedell and colleagues (Pettibone & Wedell, 2000; Wedell & Pettibone, 1996), our model predicts that the dominating option will be shifted above the indifference curve, but the mechanisms responsible for the shift are very different in the two models. In Wedell's model, changes in range and ordinal ranking are responsible for the shift. Our model predicts shifts in the mental representation of magnitude in long-term memory. The resulting distortions would bias recall of those magnitudes in a manner analogous to bias toward the central

[4] The value shift is only one portion of the model proposed by Wedell and his colleagues, but it is the portion most relevant to our current purposes.

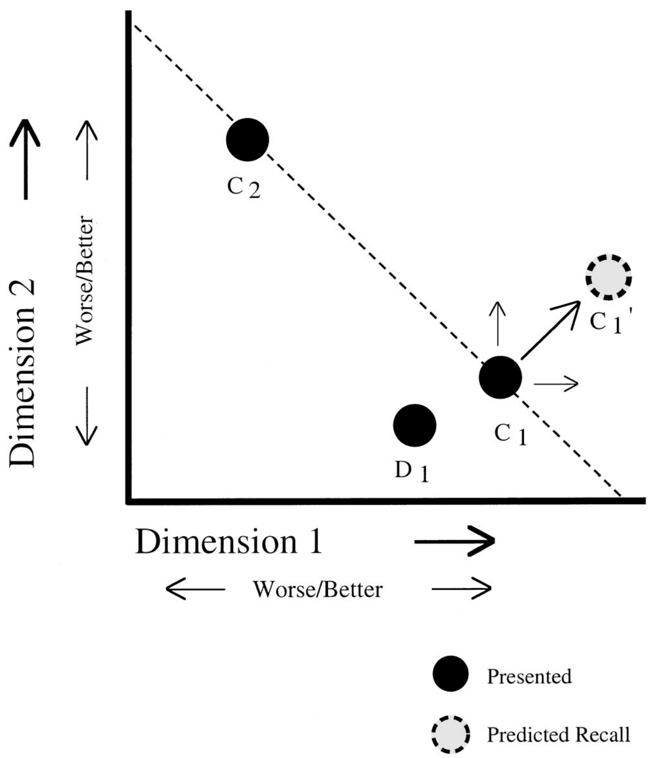

Figure 5. Our explanation of the asymmetric dominance effect. C_1 and C_2 represent choice items, D_1 represents a decoy item, and the dashed line represents the indifference curve. We propose that magnitude comparisons may distort participants' representations of the options' magnitudes—pushing the dominating option (C_1) above the indifference curve (recalled value of $C_1 = C_1'$). D_1 and C_2 would likely also become distorted (not depicted here).

tendency of a category as studied by Huttenlocher et al. (1991, 2000).[5] Finally, counter to the predictions of an attribute-only interpretation of Huttenlocher and colleagues' model (Huttenlocher et al., 1991, 2000), our model predicts that the dominating option should be recalled as more favorable to the consumer than it actually is (i.e., it should shift away from the central tendency of the category). Our model is unique in predicting that comparison words will affect estimates. To investigate whether such biases in recall obtain, in Experiment 4 we investigated the ADE using recall of values as our dependent measure.

Experiment 4

In Experiment 4, participants compared options by completing fill-in-the-blank sentences that compared the options' attributes using either less-than or more-than terms. After a delay, they recalled the values of the attributes they had compared. We predicted that the dominating choice option values would be recalled as more favorable to the consumer than the alternative choice option values. Because we wanted to avoid confounding comparison with choice, participants in this experiment did not choose between options. If the participants had chosen between options, then choice, rather than comparisons, could cause memory biases. This paradigm, of course, leaves open the question of whether memory biases are actually related to decision making. Our concern here was to investigate whether comparison-induced memory distortions would arise in choicelike situations. Investigations of whether and how comparison-induced distortions affect decision making are reported elsewhere (Choplin & Zhang, 2001). The Choplin and Zhang extension of the current theory is critical, because remembered values may not reflect the representations on which people act (see, e.g., Hastie & Park, 1986).

Method

Participants. The experimenters approached individual prospective participants on the UCLA campus or in the surrounding community and asked them to participate in the experiment. Seventy-nine people agreed to participate after being approached in this manner.

Design. This experiment used a 2 (comparison word) × 2 (value dominating the decoy: lowest or larger, as elaborated shortly) within-subjects design. There were four dimensions (price and layover for airplane tickets; rent and commute time for apartments), and conditions were assigned to dimensions. Manipulating which decoy was presented allowed us to counterbalance the dimension on which the decoy was dominated by the lowest value and the dimension on which the decoy was dominated by the larger value. For example, the participant might be given the airplane ticket option set with Ticket D_1 (costing $361 with a 181-min layover) as the decoy. D_1 was dominated by C_1, which had the lowest price in the option set ($338, i.e., the condition of decoy dominated by lowest value was assigned to price) and the longer layover time (153 min, i.e., decoy dominated by larger value was assigned to layover time). By contrast, for a participant who saw the airplane ticket option set with Ticket D_2 (costing $431 with a 95-min layover) as the decoy, D_2 was dominated by C_2, which had the shortest layover time in the option set (67 min, i.e., decoy dominated by lowest value was assigned to layover time) and the larger priced option ($408, i.e., decoy dominated by larger value was assigned to price). The two fill-in-the-blank comparison sentences used either less-than or more-than comparison terms (counterbalanced). The dependent measure was the recalled value of option attributes.

Materials and procedure. Participants were given a booklet. A cover page introduced them to the study and explained that they would be asked to compare values and recall the values they had compared. Two pages were prepared on which participants compared options (see Table 1). On the first page they compared three different airplane tickets that differed on price and layover time; on the second they compared three studio apartments that differed on rent and proximity to work (i.e., commute time). The options were arrayed left to right and labeled *A, B,* and *C* respectively. The decoy item (either D_1 or D_2) was always presented in the middle and labeled *B.* C_1 and C_2 were presented either on the left (and labeled *A*) or on the right (and labeled *C*). Assignment of C_1 and C_2 to A or C was counterbalanced. Because two comparisons were sufficient to describe the ordinal rankings of the three values, for each dimension (price and layover length for airplane tickets; rent and commute time for apartments), participants were given two identical fill-in-the-blank sentences that compared options using either the less-than or the more-than comparison term. If comparing airplane ticket prices, for example, the participant would fill in two identical sentences of the form "Ticket __ is less expensive than ticket __" or "Ticket __ is more expensive than ticket __." For the length of the layover and the length of the commute the less-than comparison term was *shorter* and the more-than term was *longer.* These two identical fill-in-the-blank sentences were presented for each of the two dimensions within each option set, producing a total of four sentences per page and eight sentences

[5] When the values are directly in front of the participant, we would not expect bias in recall. Nevertheless, choice could be biased even while the values are directly in front of the participant (see Experiment 3).

Table 1
Attribute Values of Airplane Ticket and Studio Apartment Options

Attribute	C_1	C_2	D_1	D_2
	Airplane ticket			
Cost ($)	338	408	361	431
Layover (min)	153	67	181	95
	Studio apartment			
Rent ($ per month)	632	733	665	766
Commute to work (min)	47	32	52	37

Note. C = choice item; D = decoy item.

in the entire experiment. To control for specific item effects, we counterbalanced assignment of decoy (D_1 vs. D_2) to option set (airplane tickets or studio apartments) as well as assignment of comparison words to dimensions (price vs. time). These counterbalances, along with the order counterbalance (C_1 on the left and C_2 on the right or vice versa), produced a total of eight counterbalanced conditions.

After turning a page, participants were not allowed to turn back. On the final page participants were asked to recall the values of all the airline ticket and apartment options. Written instructions on this page asked participants to estimate the value if they could not remember the exact value. Spaces were provided in which to recall every attribute value they had seen.

Results and Discussion

Recalled attribute values for ticket cost, layover time, rent, and commute time are presented in Figure 6 and Table 2. To make these diverse attributes comparable for analysis, in both Figure 6 and Table 2 we report the recalled values as a percentage of the original presented value (i.e., the values shown in Table 1, repeated in Table 2 in parentheses). Mean recalled values are shown in Figure 6. Bias in recall—that is, distortions of the dominating option above the indifference curve—can be seen for both the airplane tickets and the studio apartments with both decoys. Analyses of the differences between the mean scores shown in Table 2 revealed comparison word effects. Note that there are 24 comparisons on which we might find such effects: 4 dimensions (ticket cost, layover, apartment rent, and commute time) × 3 option values (lowest value, larger value, and decoy value) × 2 values

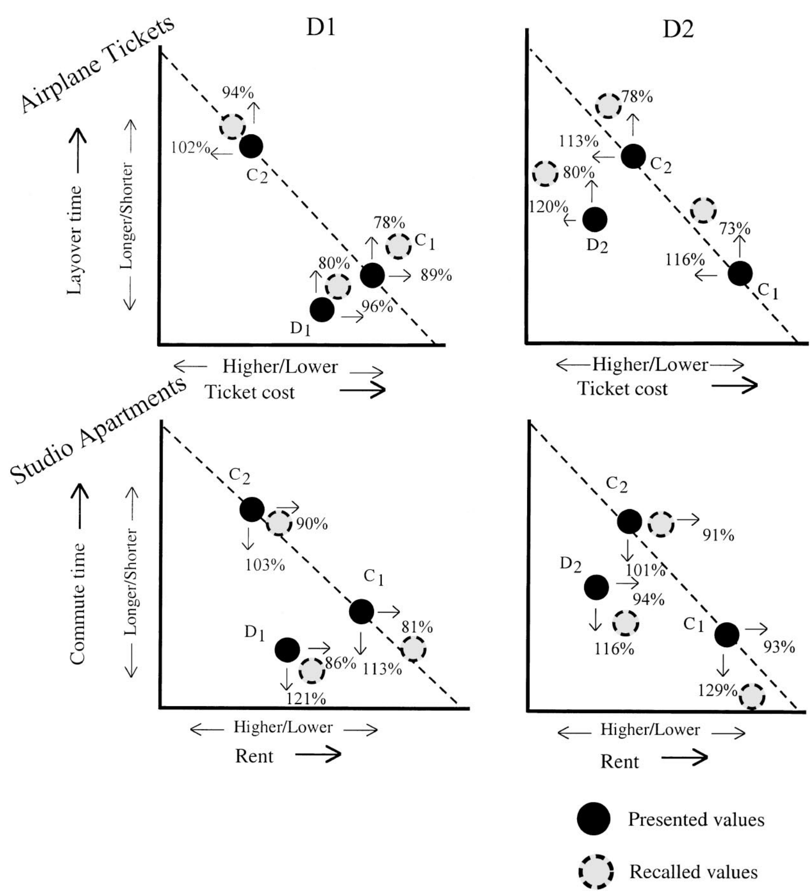

Figure 6. Results of Experiment 4: the effects of decoy position on recall of attribute values. C_1 and C_2 represent choice items, D_1 and D_2 represent decoy items, and the dashed lines represent the indifference curves. Results reported here are averaged across comparison words. Bias in recall appears to push the dominating option above the indifference curve. (Percentages are not exact.)

Table 2
Mean Recalled Values in Experiment 4 and These Values as a Percentage of the Presented Values (With Presented Values in Parentheses)

	Larger dominates				Lowest dominates			
	Less-than term		More-than term		Less-than term		More-than term	
Attribute	$ or min	%	$ or min	%	$ or min	%	$ or min	%
Airplane ticket								
Cost ($)								
Lowest (338)	367.42	110	416.18	123	297.46	88	304.43	90
Larger (408)	446.47	112	469.59	115	414.63	102	420.00	103
Decoy	501.05 (431)	118	525.05 (431)	122	345.46 (361)	96	347.30 (361)	96
Layover (min)								
Lowest (67)	58	86	68	101	50	75	55	82
Larger (153)	118	77	121	79	106	79	114	81
Decoy	144 (181)	80	144 (181)	80	75 (95)	79	76 (95)	81
Studio apartment								
Rent ($)								
Lowest (632)	598.32	95	576.37	92	497.74	79	528.13	84
Larger (733)	665.55	91	672.21	91	649.35	89	667.25	91
Decoy	719.41 (766)	94	718.32 (766)	94	566.30 (665)	85	580.42 (665)	87
Commute (min)								
Lowest (32)	32	99	34	107	30	94	35	108
Larger (47)	52	110	54	116	58	132	60	127
Decoy	62 (52)	114	66 (52)	127	40 (37)	110	45 (37)	121

dominating the decoy (lowest and larger). In 20 out of these 24 comparisons the mean recalled values compared with less-than terms were lower than the mean recalled values compared with more-than terms. The proportion of comparisons showing comparison word effects (83.3%) was reliably greater than chance (50%), $\chi^2(1, N = 24) = 9.38, p < .01$.

Analyses of the differences between these mean scores also reveal the effects of decoy position. We predicted that the dominating option would be recalled as smaller (because smaller is better, i.e., more favorable to the consumer) as a percentage of the presented value than would be the nondominating option. This prediction means that, in the condition in which the larger value dominates, the larger value was predicted to be recalled as smaller as a percentage of the presented value than the lowest value and vice versa when the lowest dominates. There are 16 comparisons on which effects of decoy position might be found: 4 dimensions (airplane ticket cost and layover, apartment rent and commute time) × 2 comparison words (less-than and more-than) × 2 values dominating the decoy (lowest and larger). When the larger value dominated, the larger value was recalled as smaller than the lowest value five times, and the lowest value was recalled as smaller three times. By contrast, when the decoy was dominated by the lowest value, the lowest value was recalled as smaller than the larger value seven times, and the lowest value was recalled as smaller one time. This difference is reliable, $\chi^2(1, N = 16) = 4.27, p < .05$, and supports the hypothesis that comparison-induced distortions can account for the ADE.

We next analyzed the recalled attribute values of the two choice options. A 2 (comparison word) × 2 (value dominating the decoy: lowest and larger value) × 2 (option values: lowest and larger value) within-subjects ANOVA on the values as a percentage of the presented values revealed an interaction between the option and decoy position, $F(1, 87) = 16.00$, $MSE = 0.05$, $p < .01$. Post hoc Tukey analyses ($\alpha = .05$) revealed that when the decoy was dominated by the lowest value, the lowest value was recalled as smaller as a percentage of the presented value ($M = 87.3\%$, $SD = 32.2\%$) than the larger value ($M = 98.3\%$, $SD = 51.1\%$; smaller is more favorable to the consumer for these stimuli). By contrast, when the larger value dominated, the lowest value was not recalled as smaller as a percentage of the presented value ($M = 101.6\%$, $SD = 37.4\%$) than the larger value ($M = 98.7\%$, $SD = 41.5\%$). This interaction supports the view that comparisons distort representations toward standard comparison-suggested differences. These distorted representations may very well be responsible for bias in choice observed in the ADE. However, the ANOVA did not reveal a main effect of comparison word; in this analysis, the values compared with more-than terms ($M = 99.4\%$, $SD = 39.4\%$) were not recalled significantly differently from values compared with less-than terms ($M = 93.6\%$, $SD = 38.0\%$), $F(1, 87) = 2.02$, $MSE = 0.29$, $p > .05$. This result suggests that comparison words may play a minor role in the ADE.

Finally, we analyzed participants' recall of the decoy attribute values by running a 2 (comparison word) × 2 (value dominating the decoy) within-subjects ANOVA on the recalled decoy attribute values. In this experiment the option that dominated the decoy was necessarily confounded with changes in decoy attribute values themselves. Specifically, because the decoys had different values, the values being recalled were different. Nevertheless, analysis of recalled values as a percentage of presented values reveals an interesting pattern. Decoys dominated by the larger valued options were recalled as larger as a percentage of the presented values (M = 103.3%, SD = 45.6%) than decoys dominated by the lowest valued options (M = 94.2%, SD = 32.2%), $F(1, 87) = 5.61$, $MSE = 0.13$, $p < .05$. This result supports the prediction that recall of decoy attribute values would be biased farther away from the dominating item than presented. The ANOVA did not reveal an effect of comparison word; in this analysis the values compared with more-than terms (M = 101.1%, SD = 39.8%) were not significantly different from the values compared with less-than terms (M = 96.4%, SD = 39.5%; $F < 1$).

In summary, Experiment 4 showed that dominating options are recalled more favorably to the consumer than nondominating options. These results are consistent with the view that comparison-induced distortions can account for the ADE. However, caution is warranted in interpreting these results. First, remembered values may not reflect the representations on which people act (see, e.g., Hastie & Park, 1986; see Choplin & Zhang, 2001, for a careful investigation into this crucial link between comparison-induced distortions and decision making). Furthermore, to explain choice biases in the ADE, comparisons would have to affect the evaluation of choice attributes immediately. One might question, however, whether biases in recall after a delay measure a person's immediate evaluation of attribute values. This last concern is allayed by the results of Experiment 3, which suggest that even if comparison-induced biases exclusively affect judgments of magnitude at short time intervals, and representations of magnitude in long-term memory only after a delay, comparison-induced biases in judgment could nevertheless account for the ADE. In addition, it is clear that these results demonstrate a unique prediction of the comparison-induced distortion account of the ADE. Indeed, we used recall rather than magnitude ratings as our dependent measure, because our account made unique predictions about recall. Even in the absence of prior ratings and/or choice, the option that dominated the decoy was recalled as having more favorable attribute values than the option that did not dominate the decoy. Neither the value-shift model of Wedell and colleagues (Pettibone & Wedell, 2000; Wedell & Pettibone, 1996) nor the justification-based model of Simonson (1989) predicted the observed biases in recall. These results also cannot be explained as recall bias toward the mean of the category (Huttenlocher et al., 1991, 2000), because recalled values shifted away from the central tendency of the presented categories. The range of recalled values was larger than the range of presented values for 11 of the 16 cases (4 [dimensions: airplane ticket cost and layover, studio apartment rent and commute time] × 2 [comparison word: less-than and more-than] × 2 [value dominating the decoy: lowest and larger]). These effects suggest that the relational information provided by magnitude comparisons served as an aid in recall.

Experiment 4 demonstrated comparison-induced distortions consistent with the hypothesis that distortions may be responsible for the choice biases observed in the ADE. Evaluating the relative magnitudes of choice and decoy attribute values entails explicitly comparing these alternatives with one another. On our account, the particular distortions we observed, which tended to increase the psychological distance between the choice items and their decoys, are especially likely to the extent that the actual differences between attribute values are smaller than the difference participants expect to exist between them.

General Discussion

Four experiments demonstrated effects of magnitude comparisons on reports of magnitude. Experiments 1a, 1b, and 2a showed that magnitude comparisons distort the reported difference between items toward comparison-suggested differences. Experiment 1a showed that participants who compared sizes exaggerated differences when they recalled these sizes; Experiment 1b showed that participants who did not compare sizes did not exaggerate differences. Experiment 2a showed that the size of a medium-sized triangle was reported as smaller by participants who compared it with a larger square than by participants who compared it with a smaller circle in spite of the fact that all three shapes were observed by all participants. Experiments 2a, 2b, and 3 showed distortions toward lower magnitudes when less-than comparison terms were used as compared with when more-than comparison terms were used. Consistent with the semantic incongruity hypothesis, Experiment 2a showed this effect for recalled size of a medium-sized triangle and Experiment 2b showed it for recalled size of a larger square, but Experiment 2c did not show it for recalled size of a small circle. In Experiment 3, comparison words affected ratings of warranty length while the warranty lengths were still in view, suggesting that comparison word effects generalize beyond delayed recall. Experiment 4 was designed to investigate whether comparison-induced distortions could account for the bias in choice observed in the ADE and revealed a pattern of comparison-induced misestimation consistent with the view that they can. Experiment 4 demonstrated biases in recall of option attributes in which the dominating options were recalled as more favorable to the consumer than the nondominating options. Taken together, these four experiments provide considerable evidence that magnitude comparisons distort mental representations of magnitude.

Potential Insights Into Other Psychological Processes

Magnitude comparisons are ubiquitous. This fact suggests that comparison-induced distortions are likely to affect a considerable number of cognitive processes. For example, the model inspiring these studies (Hummel & Holyoak, 2001) was designed to account for patterns of behavior in tasks requiring transitive inference (it is not a model of magnitude ratings, magnitude recall, or choice behavior). According to this model, response times and error rates in tasks that require people to reason about magnitudes are likely to be affected by comparison-induced distortions. Similarity, mental arithmetic, understanding of proportions and probabilities, the ability to budget time, money, calories, and so forth are all likely to be affected by comparison-induced distortions. Understanding the role of such distortions in these tasks may therefore be important for understanding behavior in these tasks.

One particularly salient example is consumer decision making. To the extent that magnitude comparisons distort estimates of product attribute values, such as price, quality, and so on, the distorted values are likely to affect the attractiveness of products. If the representation of a price, for example, were distorted downward, then the product would appear less expensive and, consequently, more attractive. The results of Experiment 4 are consistent with this idea and suggestive. However, because recalled values may not be the representations on which people act (Hastie & Park, 1986), investigations into the link between comparison-induced distortions and decision making are critical (see Choplin & Zhang, 2001).

Density effects. Density effects are a collection of well-known psychological phenomena in which the concentration of comparison items affects assessments of particular items. For example, two items will generally be thought more similar if they are adjacent and no other values fall in between them (sparse region) than if a third value is placed in between them (dense region). Density effects manifest themselves in a wide variety of domains, including perceptions of fair grading (Wedell, Parducci, & Roman, 1989), sweetness judgments (Riskey, Parducci, & Beauchamp, 1979), and assessments of physical attractiveness (Wedell, Parducci, & Geiselman, 1987), among others. Comparison-induced distortions may be responsible for at least some of these density effects, and, as seen in the context of Experiment 4, a model of density effects based on the role of comparison-induced distortions would produce unique and testable hypotheses about the origins and nature of density effects.

Consider a set of stimuli that are clustered together in psychological space (see Figure 7). The area within the cluster will be relatively dense, whereas surrounding areas will be relatively sparse. Comparing items within dense regions of space will tend to push them apart, making dense regions sparser and making adjacent (previously sparse) regions denser. That is, because interstimulus distances will become biased to reflect comparison-suggested differences, the intervals between adjacent stimuli should tend to become more homogeneous. This prediction could be tested using paradigms similar to those used in this article. Furthermore, under a comparison-induced distortion account of density effects, distributions of stimuli will change (or not) depending on whether task constraints promote or discourage explicit comparisons (see Wedell, 1996, for an account of how memory for stimulus distributions influences the effects of density on similarity). A long line of research has demonstrated an intriguing pattern of circumstances under which density effects on similarity judgments are observed (Corter, 1987; Mellers & Birnbaum, 1982; Wedell, 1996). The extent to which comparison-induced density effects can account for this pattern remains to be explored. Such an account based on the Hummel and Holyoak (2001) model would have the benefit, unique among accounts of density effects, of being a process model—that is, a detailed algorithmic account of the mental operations underlying distortions—rather than simply a verbal or mathematical theory.

Order effects. Comparison-induced distortions may also produce order effects in which successive weights (Fernberger, 1931; Guilford & Park, 1931), prices (Choplin & Zhang, 2001), pains (Jones & Gwynn, 1984; Varey & Kahneman, 1992), and so forth are judged to have higher or lower magnitudes depending on the order in which they are presented. For example, consumers might judge \$1.11 to be more expensive when presented in the series starting at \$1.06, then going to \$1.11 and finally \$1.16 than when presented in the series starting at \$1.16, then going to \$1.11 and finally \$1.06. When people observe a series of stimuli (e.g., price changes over time), they may compare the current stimulus to previous stimuli, and these comparisons would likely lead to distortions. First, order would constrain which comparison word would be thought appropriate. In an ascending series more-than comparisons of the form "The current stimulus is more than the last" are appropriate; conversely, in a descending series less-than comparisons of the form "The current stimulus is less than the last" are appropriate. These comparisons would, in turn, likely distort representations of magnitude values. Second, the difference between the current stimulus and the previous stimulus would also likely have an effect. When the difference between the current stimulus and the previous stimulus is less than participants expect, then the representations would likely diverge—moving the current stimulus away from the previous stimulus. Conversely, when the difference is more than participants expect, then representations would likely converge—moving the current stimulus toward the previous stimulus. Using a simulated shopping paradigm, Choplin and Zhang reported results consistent with these predictions.

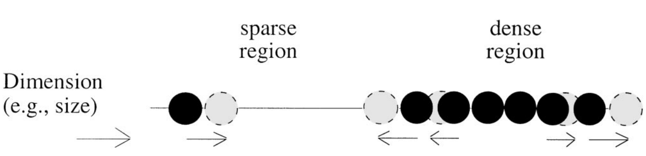

Figure 7. Density effects. Comparing items within dense regions of space may push items apart, making dense regions sparser, and adjacent (previously sparse) regions denser.

Analogous order effects have also been observed for hedonic judgments. For example, Varey and Kahneman (1992) exposed their participants to hypothetical series of aversive experiences (e.g., lifting heavy weights and carrying them for long distances) and had their participants judge how aversive the experience would be. These hypothetical experiences were identical in every respect except that they were presented either in an improving or a deteriorating order. Participants judged the series as less aversive in the improving than in the deteriorating order. Similarly, Jones and Gwynn (1984) administered pairs of shocks to participants' forearms, and participants judged how aversive the pair of shocks was as a whole. Jones and Gwynn observed a recency effect in which ascending pairs (the second stimulus was more intense than the first) were judged to be more aversive than descending pairs. Perhaps, going through the series, participants compare the event to the previous event. In ascending series the comparison words *worse* or *more painful* would be appropriate, whereas in descending series the comparison words *better* or *less painful* would be appropriate. The differences between successive events would also have an effect such that when the difference between successive events was less than the comparison-suggested difference, the differences between events would become distorted—once again, making an event presented in an ascending series appear worse than an event presented in a descending series. (See Varey & Kahneman, 1992, and Jones & Gwynn, 1984, for alternative explanations of these effects.) Of importance, as can be seen from this discussion, testable predictions regarding order effects (different from the predictions of previous pain perception theories) can be derived from the interaction of these two factors.

Price appraisal and pain perception are two domains in which comparison-induced order effects might be important. Undoubtedly, there are others.

Language. Languages differ in which comparison term—a more-than or a less-than term—serves as the marked (value-specified) comparison term and which serves as the unmarked (default) comparison term, and sometimes both are value specified (see Allan, 1986–1987; Cruse, 1976). This linguistic fact is important because the semantic incongruity hypothesis predicts different patterns of comparison-induced distortions as a function of which terms are value specified.

There are interesting differences, for example, between English and Vietnamese body-size comparison terms. To investigate these differences, Choplin, Hovannisian, and Nguyen (2001) measured semantic incongruity in English by asking English speakers to rate how "natural" it is to make various comparisons. They found that when English speakers compared two overweight people, they rated comparisons that used the word *fatter* as more natural than comparisons that used the word *thinner*. By contrast, when they compared two underweight people, they rated comparisons that used the word *thinner* as more natural than comparisons that used the word *fatter*. This interaction suggests that both *thinner* and *fatter* are value specified in English. Using the word *fatter* to describe the relationship between two underweight people or the word *thinner* to describe the relationship between two overweight people produces incongruity, and, consequently, the semantic incongruity hypothesis predicts large effects of English body-size comparison words on estimates of other people's body sizes. However, when English speakers compared themselves with others, they rated both *thinner* and *fatter* as equally natural and both were given very high naturalness ratings. Consequently, the semantic incongruity hypothesis predicts no effects of English body-size comparison words on estimates of one's own body size. By contrast, the Vietnamese body-size comparison term for *fatter* (*map tot tuong hon,* which back-translates as *more prosperous looking*) is used to compare body size almost to the exclusion of the body-size comparison term for *thinner* (*gay om hon,* which back-translates as *more ghastly–sickly*). That is, *more ghastly–sickly* is value specified and *more prosperous looking* is the default. (See Boucher & Osgood, 1969, for an account of how positive and negative evaluations of terms and social norms might produce these linguistic differences.) The semantic incongruity hypothesis, therefore, predicts that Vietnamese comparisons will affect healthy (i.e., not ghastly or sickly) people's estimates of their own body size, because the term *more ghastly–sickly* produces an incongruity when used to describe their body sizes. All of these predictions received support. The effects of English body-size comparisons (i.e., *thinner* vs. *fatter*) on estimates of other people's body sizes were quite pronounced, but no effects of English body-size comparisons on estimates of one's own body size were observed. By contrast, Vietnamese body-size comparisons did affect estimates of one's own body size. If languages systematically differ in their use of comparison terms, the language one speaks might, in a Whorfian manner, affect how one estimates magnitudes.

Conclusion

Many cognitive processes rely on representations of magnitude, yet as we have seen, these representations may be altered by comparing them with one another. Because comparisons are ubiquitous in daily life, their influence on the cognitive processes that rely on representations of magnitude is likely to be considerable.

References

Allan, K. (1986–1987). Interpreting English comparatives. *Journal of Semantics, 5,* 1–50.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 101–159). San Francisco: Academic Press.

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior, 8,* 1–8.

Choplin, J. M., Hovannisian, G., & Nguyen, E. (2001). *She is fatter than I am: Comparison-induced body size misestimation.* Unpublished manuscript.

Choplin, J. M., & Zhang, S. (2001). *This product is bigger, better, faster, and cheaper than that product: Comparison-induced distortions of attribute values and decisions.* Unpublished manuscript.

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76,* 387–404.

Clark, H. H., & Card, S. K. (1969). Role of semantics in remembering comparative sentences. *Journal of Experimental Psychology, 82,* 545–553.

Coren, S., & Girgus, J. S. (1972). A comparison of five methods of illusion measurement. *Behavior Research Methods & Instrumentation, 4,* 240–244.

Corter, J. E. (1987). Similarity, confusability, and the density hypothesis. *Journal of Experimental Psychology: General, 116,* 238–249.

Cruse, D. A. (1976). Three classes of antonyms in English. *Lingua, 38,* 281–292.

DeSoto, C., London, M., & Handel, S. (1965). Social reasoning and spatial paralogic. *Journal of Personality and Social Psychology, 2,* 3–71.

Donaldson, M., & Wales, R. J. (1970). On the acquisition of some relational terms. In J. R. Hayes (Ed.), *Cognition and the development of language.* New York: Wiley.

Fernberger, S. W. (1931). On absolute and relative judgments in lifted weight experiments. *American Journal of Psychology, 43,* 560–578.

Guilford, J. P., & Park, D. G. (1931). The effect of interpolated weights upon comparative judgments. *American Journal of Psychology, 43,* 589–599.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258–268.

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper.

Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology, 14,* 328–352.

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

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 279–305). Cambridge, MA: MIT Press.

Hunter, I. M. L. (1957). The solving of three-term series problems. *British Journal of Psychology, 48,* 286–298.

Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review, 75,* 550–560.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 196–213.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98,* 352–376.

Huttenlocher, J., Hedges, L., & Prohaska, V. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review, 95,* 471–484.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgments? *Journal of Experimental Psychology: General, 129,* 220–241.

Jones, B., & Gwynn, M. (1984). Functional measurement scales of painful electric shocks. *Perception & Psychophysics, 35,* 193–200.

Lehrer, A. (1985). Markedness and antonymy. *Journal of Linguistics, 21,* 397–429.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 582–601.

Parducci, A. (1965). Category judgment: A range–frequency model. *Psychological Review, 72,* 407–418.

Pettibone, J. C., & Wedell, D. H. (2000). Examining models of nondominated decoy effects across judgment and choice. *Organizational Behavior & Human Decision Processes, 81,* 300–328.

Riskey, D. R., Parducci, A., & Beauchamp, G. K. (1979). Effects of context in judgments of sweetness and pleasantness. *Perception & Psychophysics, 26,* 171–176.

Rusiecki, J. (1985). *Adjectives and comparison in English.* New York: Longman.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174.

Sternberg, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109,* 119–159.

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

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 767–778.

Wedell, D. H. (1996). A constructive–associative model of contextual dependence of unidimensional similarity. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 634–661.

Wedell, D. H., Parducci, A., & Geiselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology, 23,* 230–249.

Wedell, D. H., Parducci, A., & Roman, D. (1989). Student perceptions of fair grading: A range–frequency analysis. *American Journal of Psychology, 102,* 233–248.

Wedell, D. H., & Pettibone, J. C. (1996). Using judgments to understand decoy effects in choice. *Organizational Behavior & Human Decision Processes, 67,* 326–344.

Received August 28, 2000
Revision received January 8, 2002
Accepted January 9, 2002 ■