In: Psychology of Decision Making in Education
Editor: Jeanine A. Elsworth
ISBN 1-60021-933-7

Chapter-03

TOWARD A COMPARISON-INDUCED DISTORTION THEORY OF JUDGMENT AND DECISION MAKING

Jessica M. Choplin *

DePaul University

ABSTRACT

This chapter demonstrates how biases produced by verbal comparisons (Choplin & Hummel, 2002) might produce a variety of phenomena in the psychology of judgment and decision making. The biases produced by verbal comparisons cause people to overestimate the importance of small differences and underestimate the importance of large differences. Simulations will demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from a reference point (default value, status quo, etc.) would produce s-shaped evaluation functions. Simulations will also demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from other contextual stimuli would produce distribution-density effects (often called frequency effects). Because large differences are underestimated, when the variance in people's unbiased estimates is large as in anchoring effects, these biases will often look like assimilation effects. Biases produced by verbalized social comparisons would also overestimate the importance of small differences and underestimate the importance of large differences so that moderate downward comparisons will produce higher evaluations of the self than will extreme downward comparisons and moderate upward comparisons will produce lower evaluations of the self than will extreme upward comparisons. Comparison strategies might also help explain decoy effects such as asymmetric dominance effects and phantom decoy effects. Testable and falsifiable assumptions of this model are described thereby laying a foundation for future empirical research.

---

* Research supported by a grant from the National Science Foundation (NSF-0621664) awarded to Jessica M. Choplin. Correspondence to: Jessica M. Choplin; Department of Psychology; DePaul University; 2219 North Kenmore Avenue; Chicago, IL 60614-3504; jchoplin@depaul.edu

INTRODUCTION

Hedonic and other types of attribute evaluations have wide ranging implications for decision-making and the consequences of those decisions for well-being. Evaluations of price and other consumer-product attributes affect the decisions people make as consumers and, in turn, how satisfied they are with the products they buy (e.g., Choplin & Hummel, 2002; Choplin & Hummel, 2005; Huber, Payne, & Puto, 1982; Wedell & Pettibone, 1996). Evaluations of people's own personal attributes—often informed by social comparisons (e.g., Festinger, 1954; Lavine, Sweeney, & Wagner, 1999; Morse & Gergen, 1970; Mussweiler, 2003; Wedell & Parducci, 2000)—affect how people see themselves on any personal attribute one might imagine. Evaluations of food intake (taste, calories, portion size, Riskey, Parducci, & Beauchamp, 1979; Schifferstein & Frijters, 1992; Wansink, 2004) affect people's dietary decisions and all of the consequences of those decisions. People's subjective evaluations of their own well-being (i.e., how happy they see themselves, Kahneman, 1999; Parducci, 1995; Schwarz & Strack, 1999) often affect actual well-being.

In this chapter, I present an introduction to a model of attribute evaluation, comprehension, memory, and estimation wherein verbal, language-expressible comparisons of attribute values systematically bias how people evaluate, comprehend, remember, and estimate the compared attribute values (comparison-induced distortion theory, CID theory, Choplin & Hummel, 2002, 2005). Judging one product to be more expensive than a second product, for example, might bias people's judgments of how affordable those products are. I will present simulation data that suggests that comparison-induced biases might create a variety of phenomena in the psychology of judgment and decision making and provide key insights into yet other phenomena. My goal in this chapter is merely to introduce readers to the basic tenants of CID theory. I will use simulation data to demonstrate that biases produced by verbal comparisons are sufficient to explain these phenomena and to describe some key novel predictions that can be tested in future research. It is not my focus in this basic introduction to argue that it is necessary to appeal to comparison-induced biases to explain these phenomena nor is it my focus to argue that CID theory provides a better account of these phenomena than other theories provide. Arguments and empirical results demonstrating that CID theory provides a better account than other accounts are presented elsewhere (Choplin & Hummel, 2005, and in papers currently in process or preparation, and ongoing empirical research). Furthermore, the fact that CID theory can account for such a large variety of phenomena (a larger variety of phenomena than any of its competitors) is itself strong evidence in its favor.

COMPARISON-INDUCED DISTORTION THEORY

Comparison-induced distortion theory (CID theory, Choplin & Hummel, 2002) is a theory of attribute evaluation in which language-expressible magnitude comparisons (e.g., "I am older than Susan is") systematically bias how people evaluate, comprehend, remember, and estimate attribute values. The basic idea behind CID theory is that language-expressible magnitude comparisons suggest quantitative values. For example, to investigate the meanings of English age comparisons Rusiecki (1985) gave his participants sentences such as "Mary is older than Jane" and "Martin's wife is older than Ken's wife" and asked them to report the ages they imagined. Rusiecki found considerable agreement in the values imagined by his participants. In response to the comparison "Mary is older than Jane" participants imagined Mary to be 20.2 years old on average and Jane to be 17.9 years old on average. In response to the comparison "Martin's wife is older than Ken's wife" participants imagined Martin's wife to be 37.2 years old on average and Ken's wife to be 33.0 years old on average.

Of particular interest to the current discussion, the age differences imagined by Rusiecki's (1985) participants were remarkably similar. Regardless of the ages they imagined, participants imagined a difference between the ages of approximately 2 to 5 years (slightly larger for larger values; more on this topic below)—not 1 month or 30 years. Inspired by these results, Rusiecki argued that comparisons suggest quantitative differences between compared values. I will henceforth call these quantitative differences "comparison-suggested differences," because they are the differences suggested by comparisons. In the case of age comparisons, Rusiecki's results demonstrate that comparison-suggested differences are approximately 2 to 5 years. For ease of discussion I will operationally define the difference suggested by age comparisons to be 3.5 years in the discussion that follows. Please note, however, that the actual size of the difference likely depends upon many factors as I discuss below and that empirical measurements of the comparison-suggested difference will have to be cognizant of these factors.

Choplin and Hummel (2002) proposed that language-expressible magnitude comparisons (like those investigated by Rusiecki, 1985) might bias evaluations toward the quantitative values suggested by comparisons. For example, if the actual age difference between two people were 1.5 years (i.e., less than the comparison-suggested difference of 3.5 years), then a comparison might bias evaluations of their ages apart—toward a difference of 3.5 years. The younger person might be evaluated younger than she or he would have been evaluated otherwise and the older person might be evaluated older than she or he would have been evaluated otherwise. If the actual age difference between two people were 5.5 years (i.e., more than 3.5 years), then a comparison might bias evaluations of their ages together—again toward a difference of 3.5 years. The younger person might be evaluated older and the older person might be evaluated younger.

Formally, the comparison-suggested value of the smaller of two compared values (E_S; E for Expected) and the comparison-suggested value of the larger of two compared values (E_L) can be calculated from Equations 1a and 1b respectively:

$$E_S = S_L - D \qquad (1a) \qquad\qquad E_L = S_S + D \qquad (1b)$$

where S_L and S_S (S for Stimulus values) are the values of the larger and smaller values unbiased by comparisons respectively and D is a parameter representing the comparison-suggested difference (Choplin & Hummel, 2005). Participants probably do not calculate these comparison-suggested values on the fly. Rather, these values likely come from their memory of previous times in which they judged one value to be larger or smaller than another value (e.g., memory of the age of previous people they had judged to be older than 33.0 years old).

In the tradition of previous judgment models (Anderson, 1965; Huttenlocher, Hedges, & Vevea, 2000), CID theory is a weighted average model in that represented values are assumed to be a weighted mean of values unbiased by comparisons and comparison-suggested values:

$$R_S = wE_S + (1\text{-}w)S_S \qquad (2a) \qquad\qquad R_L = wE_L + (1\text{-}w)S_L \qquad (2b)$$

where w is the relative weights of the two values, is bound between 0 and 1, and is constrained so as to prevent impossible values (e.g., negative years or sizes of geometric figures) from being represented. For example, assuming a comparison-suggested difference, D, of 3.5 years (an oversimplification used here for demonstration purposes only; see discussion below on measuring comparison-suggested differences), a comparison between a 22-year old and a 28-year old would bias evaluations of their ages toward each other. If the weight given to comparison-suggested values were 0.2 (I use this value for demonstration purposes only; to model real data this value would be found by fitting the model to empirical data), then the represented age of the 22-year old would be 22.5 years and the represented age of the 28-year old would be 27.5 years. That is, the 22-year old would be evaluated, i.e., treated, half a year older and the 28-year old would be evaluated half a year younger.

Analogous to previous weighted-average judgment models that weight information from various sources (e.g., Huttenlocher et al., 2000), Parameter w represents the degree to which people rely upon comparisons to evaluate attribute values. If people have an accurate understanding of how large or small or good or bad attribute values are and they are completely certain of their own understanding, then they will have little need to use comparison information to evaluate attribute values and Parameter w would take a low value. By contrast, if they do not have an accurate understanding and need to evaluate attribute values relative to other contextual attribute values, then they will need to rely upon comparison information and Parameter w would take a high value. In cases where people need to recall attribute values from memory, if they remember exact values, they will have little need to rely upon comparison information to aid their recall and Parameter w would take a low value. If they do not remember exact values, they might use verbal comparison information to aid their recall and Parameter w would take a high value. Many cognitive processes could potentially produce this averaging and like other weighted-average models CID theory does not need to commit itself to any one specific cognitive process.

Sometimes people might hesitate to describe one value as larger (or smaller) than another value even if it is larger (or smaller). It might seem odd, for example, to describe a person who is 28 years and 4 months old as “older” than a person who is 28 years and 2 months old. People might prefer to describe these ages as “approximately the same” or “similar.” These comparisons suggest that there is little to no difference between the compared values and so biases in evaluation created by these comparisons can be modeled by setting parameter D in Equations 1a and 1b to zero. If Parameter w in Equations 2a and 2b were set to .2, the comparison “a 28 year and 4 month old person is approximately the same age as a 28 year and 2 month old person” would bias the evaluation of 28 years and 4 months to be about 12 days younger and the evaluation of 28 years and 2 months to be about 12 days older.

The decision to describe a given difference as “approximately the same” or “similar” can be modeled stochastically using Shepard’s (1987) law of generalization. That is, the probability of describing a difference between two stimuli (i.e., S_L and S_S) as approximately the same or similar can be modeled as:

$$p(\text{"}same\text{"} \mid S_L - S_S) = e^{-c(S_L - S_S)} \quad (3)$$

where c is a sensitivity parameter.

Equations 1a through 3 employ three parameters (D, w, and c). Note that although the theory uses these three parameters, Parameters D and c can usually be measured empirically so these two parameters will generally not be free. The ability to measure Parameters D and c empirically leaves only Parameter w as a free parameter in CID theory—the same number of free parameters as some of its competitors such as adaptation-level theory (AL theory, Helson, 1964) and range-frequency theory (RF theory, Parducci, 1965, 1995) and fewer free parameters than other competitors (Frederick & Loewenstein, 1999; Stevens, 1961). Parameter D in Equations 1a and 1b can be measured by asking a control group of participants about the types of differences they imagine in response to a comparison (as Rusiecki, 1985, did) or by looking at real-world differences under the assumption that participants have been exposed to these differences and that their experiences with these differences have shaped their understanding of the differences implied by these comparisons. Parameter c in Equation 3 which reflects the probability that people will judge values to be "approximately the same" can be measured by asking a control group of participants whether or not they would describe a given difference as "approximately the same." I discuss the ways one might measure these values in the next sections.

MEASURING COMPARISON-SUGGESTED DIFFERENCES

As described above, comparison-suggested differences (Parameter D in Equations 1a and 1b) and values (E_S and E_L in Equations 1a and 1b) can be measured empirically by asking control groups of participants to imagine that one value is more than, less than, larger than, smaller than, better than, worse than, etc. another value and then ask them about the differences they imagine. My students, colleagues, and I have measured comparison-suggested differences in this manner many times.

For example, in one study my students and I measured comparison-suggested differences and values in personal attributes (grade point average, height, weight, income, dates per month, and commute to campus) by asking a control group of participants about the differences they imagined. In particular, we described a fictional DePaul University undergraduate student (Jennifer for women; Brad for men; the personal attributes of these two people were collected from yet another control group of participants who imagined the average female or male undergraduate DePaul University student) and asked the control group to imagine the personal attributes of another DePaul University undergraduate student (Michelle for women; Michael for men). Michelle was better on every personal attribute than Jennifer (higher grade point average, taller, weighed less, earned more, went out on dates more often, and lived closer to campus); and Michael was better on every personal attribute than Brad (higher grade point average, taller, more muscular and so weighed more, earned more, went out on dates more often, and lived closer to campus). Jennifer and Brad's personal attributes as well as the median imagined personal attributes of Michelle and Michael are presented in Table 1. The median imagined personal attributes of Michelle and Michael represent comparison-suggested values. Comparison-suggested differences were calculated from these comparison-suggested values. Perhaps unsurprisingly the attributes of average female and male undergraduate DePaul University students imagined by these pretest participants fit gender stereotypes. The average female student imagined by women (Jennifer) earned a higher grade point average, was shorter, weighed less, earned much less money (65¢ on the dollar), went out on more dates, and lived closer to campus than did the average male student imagined by men (Brad). We will use these values later in the chapter when we discuss social comparisons.

Table 1. Values and Differences Suggested by Comparisons to Michelle and Brad

	Jennifer's personal attributes (collected in another pretest)	Michelle's median imagined personal attributes (comparison-suggested values)	Comparison-suggested differences
g.p.a.	3.17	3.5	0.33
Height	64 inches	66 inches	2 inches
Weight	134 pounds	125 pounds	9 pounds
Income	$7,868.00	$9,000.00	$1,132.00
Dates	3.1	5	1.9
Commute	9.2 miles	5 miles	4.2 miles

	Brad's personal attributes (collected in another pretest)	Michael's median imagined personal attributes	
g.p.a.	3.1	3.38	0.28
Height	70 inches	72 inches	2 inches
Weight	173 pounds	190 pounds	17 pounds
Income	$12,084.00	$15,000	$2,916.00
Dates	2.45	4	1.55
Commute	18.9 miles	10 miles	8.9 miles

The sizes of comparison-suggested differences depend upon many factors. For one, the size of the comparison-suggested difference often depends upon the size of the base of the comparison. For example, Choplin and Hummel (2005) studied comparison-induced distortions in judgments of line length. To measure comparison-suggested values and differences in judgments of line length, John Hummel and I showed a control group of participants lines of various lengths (each control participant only saw one line) and then asked them to imagine a line that was longer or shorter than that line and draw it. Looking first at participants who drew longer lines than the lines they saw: Of those who viewed a line that was 10.0 mm long, the median redrawn longer line was 36.3mm (a difference of 26.3mm). Of those who viewed a 14.0 mm-line, the median redrawn line was 42.5mm (a difference of 28.5mm). Of those who viewed a 22.0 mm-line, the median redrawn line was 53.0 mm (a difference of 31.0 mm); and of those who viewed a 30.0 mm-line, the median redrawn line was 60.0 mm (a difference of 30.0 mm). Next looking at participants who drew shorter lines than the lines they saw: Of those who viewed a line that was 34.0 mm long, the median redrawn shorter line was 18.0 mm (a difference of 16.0 mm). Of those who viewed a 30.0 mm-line, the median redrawn line was 15.0 mm (a difference of 15.0 mm). Of those who viewed a 22.0 mm-line, the median redrawn line was 9.0 mm (a difference of 13.0 mm);

and of those who viewed a 14.0 mm-line, the median redrawn line was 6.0 mm (a difference of 8.0 mm). In this case, the sizes of the comparison-suggested differences depended upon the size of the base of the comparison such that as a general trend comparison-suggested differences were larger when the base of the comparison was larger.

Comparison-suggested differences might not always be larger when the base of the comparison is larger. Rather, the size of the comparison-suggested difference might depend upon the distribution of values out in the world. The comparison "a man taller than 5'8"," for example, might suggest a larger difference than would the comparison "a woman taller than 5'8"," because the distribution of men's heights is larger than the distribution of women's heights. Also if people have some idea of the amount of variance in a category, comparison-suggested differences might be larger when variance is larger. The variance in the high school g.p.a.s of all high school graduates, for example, might be larger than the variance in the high school g.p.a.s of DePaul University undergraduate students, because DePaul University undergraduate students are a self-selecting group and the self-selection process involves high school g.p.a. among other factors. Comparison-suggested differences would then likely be larger if one were comparing the high school g.p.a.s of two high school graduates than if one were comparing the high school g.p.a.s of two DePaul University undergraduate students.

Furthermore, not all ways of wording comparisons will suggest the same quantitative difference. Wording a comparison "slightly larger," for example, will suggest a smaller quantitative difference than will wording it "larger;" and wording a comparison "much larger" will suggest a larger quantitative difference than will wording it "larger." The fact that alternative ways of wording comparisons will suggest smaller or larger quantitative differences gives us a method of isolating the evaluation biases created by verbal comparisons from the evaluation biases produced by some other factors. Once a researcher has measured the quantitative differences suggested by alternative ways of wording verbal comparisons among control participants, an experimenter will be able to manipulate the quantitative differences suggested to experimental participants by manipulating how a comparison is worded while keeping all other factors constant. Once all other factors have been held constant, any observed effects would have to be due to the verbal comparisons that were used. Note that special care would be needed to hold emotional factors constant as emotional factors could also potentially be affected by different ways of wording comparisons. Emotional factors might be held constant by testing for the effects of alternative ways of wording comparisons on dimensions that have no hedonic consequences such as line length or size of geometric shapes.

MEASURING LIKELIHOOD OF JUDGING TWO VALUES THE SAME

The likelihood that people will describe two values as approximately the same can be modeled using Shepard's (1987) law of generalization (shown in Equation 3). The law of generalization has one parameter (Parameter c). A value for Parameter c can be measured empirically by asking control groups of participants how they would describe differences (e.g., as "same"—or perhaps "approximately the same" or "same ballpark"—versus "different," "larger than," or "smaller than"), fitting Equation 3 to these results using a root mean squared error (RMSE) criterion, and thereby finding the best-fit value for Parameter c. Empirically measuring the value of Parameter c for control participants allows CID theory to make a priori predictions about how experimental participants will evaluate attributes.

My students and I measured the likelihood that people will judge values to be the same and found the best-fitting value for Parameter c as described above in an experiment in which they judged whether prices for an all-you-can-eat lunch were the same as or different than \$6.53. In particular, participants imagined that they worked for a summer camp that had been charging \$6.53 for an all-you-can-eat lunch. For some of the participants, the camp was considering raising the price and they were asked whether they would consider each of the prices \$6.62, \$6.71, \$6.80, \$6.89, \$6.98, \$7.07, and \$7.16 to be the same as or different than \$6.53. Half of these participants simply judged whether the prices were "the same as" or "different than" \$6.53. The other half of these participants judged whether the prices were in "the same ballpark as" or "a completely different ballpark than" \$6.53. For other participants the camp was considering lowering the price and they were asked whether they would consider each of the prices \$6.44, \$6.35, \$6.26, \$6.17, \$6.08, \$5.99, and \$5.90 to be the same as or different than \$6.53. The proportions of participants describing each price as the same as \$6.53 are plotted in Figure 1. Shepard's (1987) law of generalization (Equation 3) was fit to the results using a root mean squared error criterion. Best-fits are also plotted in Figure 1. In the condition wherein participants simply judged whether the prices were "the same as" or "different than" \$6.53, the best-fitting value for Parameter c for raising prices was 3.33 (RMSE = 0.05) and the best-fitting value for lowering prices was 2.46 (RMSE = 0.12). In the condition wherein participants s judged whether the prices were in "the same ballpark as" or "a completely different ballpark than" \$6.53, the best-fitting value for Parameter c for raising prices was 1.77 (RMSE = 0.09) and the best-fitting value for lowering prices was 1.58 (RMSE = 0.09).

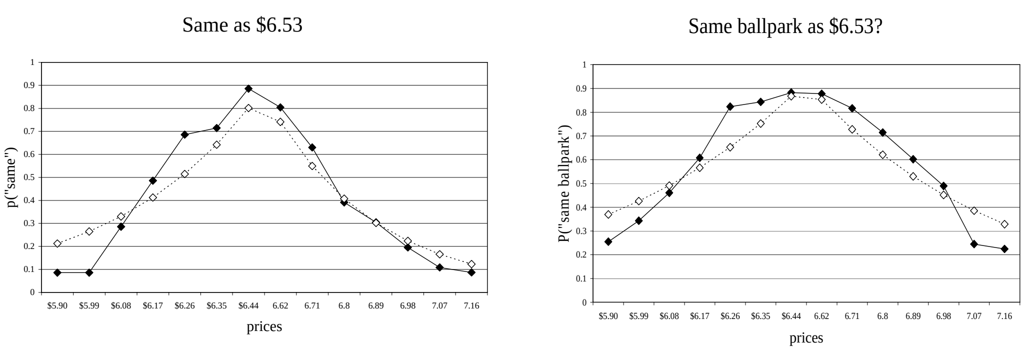

Figure 1. Probabililty of describing prices as the same as \$6.53. Black diamonds represent the proportion of participants who described each price as the same as \$6.53. White diamonds represent the best-fitting values when Equation 3 was fit to participants' responses.

Once comparison-suggested differences (Parameter D in Equations 1a and 1b) and likelihood of judging two values approximately the same (Parameter c in Equation 3) have been empirically measured, only the weighting of the information provided by comparisons (Parameter w in Equation 2a and 2b) will remain as a free parameter. Additionally, as described below CID theory will only be able to explain certain phenomena, if Parameters D and c take certain values. If CID theory is to explain these phenomena, then, it will have to predict that these parameters will take these values and these predictions can then be empirically verified (or else the CID theory account of these phenomena would be empirically falsified).

Quantitative modeling (discussed shortly) demonstrates that biases in evaluation created by verbal comparisons might play a key role in a number of well-known phenomena in the psychology of judgment and decision making. In particular, biases in evaluation created by verbal comparisons might play a key role in producing s-shaped evaluation functions, distribution-density effects, anchoring effects, social comparisons, and decoy effects. Although I will not be pitting CID theory against alternative theories here in this chapter, note that to compete with CID theory any alternative theory would have to provide a more parsimonious account of all of these phenomena combined, not just a more parsimonious account of one phenomenon. The fact that CID theory can account for such a large variety of phenomena is itself strong evidence in favor of CID theory. The remainder of this chapter will describe the role that biases produced by verbal comparisons might play in these phenomena. To lay the ground work for future empirical research, empirically testable predictions of the CID theory account of these phenomena will also be described.

S-SHAPED EVALUATION FUNCTIONS

Previous research has found considerable evidence that evaluations often follow s-shaped functions (see Figure 2, Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979). Following Helson's (1964) AL theory, these evaluation functions predict that evaluations are made relative to a single reference point (RP), i.e., the point at which people consider values "normal" or "average." This function differs from Helson's original formalization of AL theory, however, in that it has an s-shape. Evaluations are concave (downward) for positive changes from the RP and convex (concave upward) for negative changes from the RP. Helson assumed that evaluations are linearly transformed around the RP.

Figure 2. Evaluations made relative to reference points (RPs; 22 years and $50,000). Evaluation functions are generally s-shaped. They are formalized here using Stevens' (1961) Power Law (black s-shaped functions) and CID theory (speckled lines).

To see that evaluations are often s-shaped, consider how evaluations of age change over one's lifespan. As a child (e.g., the RP might be 6 years), all adults—and even teenagers—seem old. The portion of the curve representing adults and teenagers (i.e., the upper portion of the s-shape) is relatively flat. As a young adult of 22 years (i.e., the RP is 22 years; see Figure 2 Age Evaluations), children all seem young, mature adults all seem old, and the difference between 18-year olds and 25-year olds is huge. The portion of the curve representing children (i.e., the lower portion of the s-shape) and the portion of the curve representing mature adults (i.e., the upper portion of the s-shape) are both relatively flat, but the portion of the curve representing the difference between 18-year olds (just to the left of the RP) and 25-year olds (just to the right of the RP) is the steepest portion of the curve. As an adult of retirement age (e.g., the RP might be 65 years), the differences between children, teenagers, and young adults all seem insignificant, young parents appear to be kids having kids, and the lower portion of curve representing these ages is relatively flat.

Since this s-shape is generally thought to reflect the psychophysical law that sensitivity to differences decreases at greater magnitudes, it is often formalized as plotted in Figure 2 solid black lines using Stevens' (1961) Power Law (Kahneman & Tversky, 1979):

$$Evaluation(X) = \begin{cases} a_{more}(X - RP)^{b_{more}}, & if\ X > RP \\ -a_{less}(RP - X)^{b_{less}}, & otherwise \end{cases} \quad (4)$$

where X is the value being evaluated, RP is the reference point, a_{more} scales sensitivity to positive differences from adaptation, a_{less} scales sensitivity to negative differences from adaptation, b_{more} scales the curvature of the portion of the function representing positive differences, and b_{less} scales the curvature of the portion of the function representing negative differences. To produce s-shaped evaluation functions like those in Figure 2, the parameters b_{more} and b_{less} will take values greater than zero and less than one. They will be equal to one, if evaluations are linear (see Briesch, Krishnamurthi, Mazumdar, & Raj, 1997). They will be greater than 1 only in rare cases (e.g., perhaps evaluations of electric shocks, Stevens, 1962).

CID theory provides an alternative way of formalizing the s-shaped evaluation function than does Stevens' Law. The top-most lines in Figure 2 represent biases created by more-than comparisons simulated using Equations 1a through 2b. For example, the comparison "a 35-year old is older than a 22-year old" would produce the evaluation of the 35-year old shown in Figure 2 Age Evaluations. The middle lines represent biases created by approximately-the-same comparisons. The likelihood that people will describe values as approximately the same to the RP is simulated using Equation 3 and the biases produced by these approximately-the-same comparisons are simulated using Equations 1a through 2b with Parameter D set at 0 (since the comparison-suggested difference of describing two values as approximately the same would be zero difference). Unlike Stevens' Law, CID theory can capture indifference around the RP (Kalyanaram & Little, 1994), if people judge a wider range of values as "approximately the same as" the RP. The bottom lines represent biases created by less-than comparisons.

Continuous s-shaped evaluation functions similar to the Stevens' Law evaluation function would be produced by averaging across the three types of comparisons. When values are very close to the RP almost all participants will describe differences as the same, but the tendency to do so usually tapers off quickly for values that are farther from the RP (see Figure 1). It tapers off at a rate that usually produces a function that, once biases from all three types of comparisons are averaged together, is very similar to (but not identical to) the function produced by Stevens' Law.

To demonstrate what the evaluation function looks like when evaluations for less than, approximately the same, and more than comparisons are averaged together, I again simulated age evaluations from the perspective of a 22-year old who compares all ages to her or his own age (see Figure 3). Five hundred ages were randomly drawn from each of a normal, a positively skewed, a bimodal, and a uniform distribution. For demonstration purposes only, Parameter D in Equations 1a and 1b was set to 3.5 years, Parameter c in Equation 3 was set to 1.5 (before modeling empirical data, these two values would have to be empirically measured), and Parameter w in Equations 2a and 2b was set to .5 (to model empirical data, this value would be found by fitting this free parameter to data).

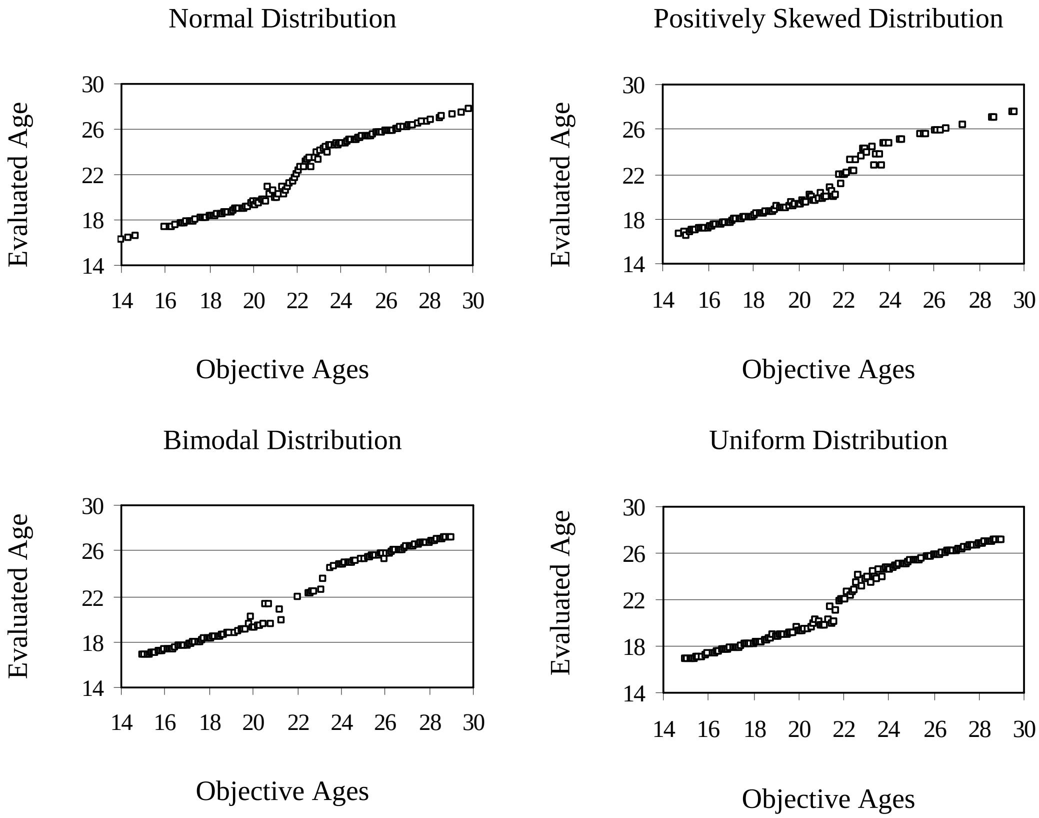

Figure 3. Simulated comparison-biased age evaluations—All ages are compared to 22 years.

At c set to 1.5, the age 22 years and 3 months was judged to be the same age as 22 years 69% of the time. The comparison "22 years and 3 months is approximately the same age as 22 years" biased the evaluation of 22 years and 3 months to be approximately 22 years and 1.5 months. The other 31% of the time 22 years and 3 months was judged older than 22 years. The comparison "22 years and 3 months is older than 22 years" biased the evaluation of 22 years and 3 months to be approximately 23 years and 10.5 months. Averaging across these two types of comparisons, 22 years and 3 months was evaluated as approximately 22 years and 4.3 months. An actual difference of 3 months was then evaluated as if it were (treated as if it were) 4.3 months. This extra sensitivity close to the RP can be seen in Figure 3 by the steeper slope around 22 years (the immediate region around the RP is not always steep, sometimes there is a region in which values are judged to be the same, see Kalyanaram & Little, 1994). Sensitivity quickly decreased for values that were farther away from the RP. An age of 24 years was evaluated as if it were 24 years and 10 months and an age of 26 years was evaluated as if it were 25 years and 10 months making a difference of 2 years seem like 1 year.

Notice also from Figure 3 that CID theory predicts that evaluations will be s-shaped and unaffected by the type of distribution from which values are drawn, if all values are verbally compared to a central RP and compared to no other contextual values. Simulation results presented in the next section demonstrate that evaluation functions will depend upon the type of distribution from which values are drawn, if values are compared to other contextual values.

A well-known phenomenon associated with s-shaped evaluation functions is that people's aversion to losses is often greater than their affinity for gains (Kahneman & Tversky, 1979). Consider evaluations of income (Figure 2 Salary Evaluations). For someone who earns \$50,000 a year, the subjective difference between \$40,000 and \$50,000 (a difference that represents a loss) is large, likely larger than the subjective difference between \$50,000 and \$60,000. Reflecting this fact, the evaluation function for values to the left of the RP in Figure 2 Salary Evaluations is steeper than the evaluation function for values to the right of the RP. This pattern would reverse, if large values represented loss (e.g., being charged more).

Previous research has shown that loss aversion is due to affective factors. These affective factors might affect how people use language to describe differences. For example, the comparison words they use to describe losses (e.g., loss, less, deficit, shortfall, or shortage) might suggest larger differences than do the comparison words they use to describe gains (e.g., gain, more, surplus, or excess). They also might be less likely to describe losses as approximately the same as RP values than to describe gains as approximately the same as RP values. If loss aversion causes people to use their language in either of these ways, then the biases produced by verbal comparisons would not only be s-shaped, but would also reflect the fact that people are loss averse. This model of the s-shaped evaluation function assumes that verbal comparisons mediate the effects of affective factors on evaluations. That is, this model assumes that affective factors play a role in evaluations by changing comparison words; verbalized comparisons in turn bias evaluations.

The speckled lines in Figure 2 Salary Evaluations demonstrate the point that biases produced by verbal comparisons would reflect loss aversion by plotting what the comparison-biased evaluation function would look like, if—because of loss aversion—comparison-suggested differences (parameter D in Equation 1) were larger for losses than for gains. As seen in Figure 2 Salary Evaluations, biases produced by verbal comparisons nicely capture the asymmetry in the evaluation function caused by loss aversion. To mathematically demonstrate this point, comparisons to an RP were simulated with Parameter D set higher for losses than for gains. Sensitivity to losses and gains was assessed by fitting Equation 4 to the simulation results using a minimized root mean squared error criterion. Demonstrating increased sensitivity to losses, the best-fitting value for Parameter a in Equation 4 was greater for losses than for gains.

Not plotted in Figure 2, a tendency to avoid describing losses as "approximately the same" (Parameter c) would also nicely capture the asymmetry in evaluations of losses and gains. To investigate the effects of asymmetries in Parameter c on sensitivity to losses and gains, comparisons to an RP were simulated with Parameter c set higher for losses than for gains. Again, sensitivity to losses and gains was assessed by fitting Equation 4 to the simulation results using a minimized root mean squared error criterion; and again the best-fitting value for Parameter a in Equation 4 was greater for losses than for gains.

Importantly, once Parameters D and c are measured empirically as described above, the CID theory account of the s-shaped evaluation function has fewer free parameters than does Steven's Law. The Steven's Law account has 4 free parameters (a_{more}, a_{less}, b_{more}, and b_{less}), while the CID theory account only has one (w, or two, if w takes different values for losses than for gains). Not only can Parameters D and c be measured empirically, but also to explain some phenomena CID theory will have to predict that Parameters D and c will take certain values. These predictions can then be tested empirically. These parameters, therefore, represent testable predictions of the model rather than free parameters. For example, the CID theory account of the s-shaped evaluation function predicts that either comparison-suggested differences will be larger for losses than for gains or people will be less likely to describe losses as “approximately the same” than to describe gains that way, or both. If people do neither of these things, then CID theory would be unable to capture loss aversion or explain the s-shaped evaluation function.

Of course, finding the predicted correlations between how people use language and biases in their judgments and choice behavior would not in itself establish that these biases are caused by verbal comparisons (as CID theory maintains) as both could be caused by third variables (i.e., affective factors, psychophysical factors, or both) or the causal direction could be reversed. To establish, that verbal comparisons cause biases in judgment, one can manipulate these parameters. Parameter D might be manipulated, for example, by tagging the adjectives “slightly” or “much” on to comparisons; Parameter c might be manipulated by describing sameness as “same,” “approximately the same,” or “in the same ballpark.” To control for affective factors, one might study evaluations of non-hedonic dimensions such as line lengths or sizes of geometric shapes.

Previous research has found that evaluation functions are not always s-shaped. Rather, evaluation functions will often depend upon the distribution of contextual attribute values. Typically, for example, (if the range of values is held constant) an attribute value that is drawn from a positively skewed distribution will be judged larger than the same attribute value that is drawn from a negatively skewed distribution (Birnbaum, 1974; Hagerty, 2000; Haubensak, 1992; Niedrich, Sharma, & Wedell, 2001; Parducci, 1965, 1995; Riskey et al., 1979; Stewart, Chater, & Brown, 2006). I will discuss these distribution-density effects and how CID theory might explain them in the next section.

DISTRIBUTION-DENSITY EFFECTS

Niedrich, Sharma, and Wedell (2001) described models of evaluation—like Helson’s (1964) adaptation-level theory and the model formalized in Equation 4 as prototype models. An alternative to these models is suggested by the view that values might be compared to other exemplars that are drawn from the distribution. That is, instead of being compared to a central prototypical example, a to-be-evaluated attribute value might instead be compared to large values sometimes, small values sometimes, and frequent values most commonly of all (i.e., most commonly compared to central values in a normally skewed distribution, large values in a negatively skewed distribution, and small values in a positively skewed distribution). The most successful exemplar model of evaluation in the literature is Parducci’s (1965; 1995) range-frequency theory (RF theory) in which evaluations are affected by two types of information: the range score of x in distribution k (R_{xk}) and the frequency score (percentile rank) of x in distribution k (F_{xk}). The range score is calculated as shown in Equation 5:

$$R_{xk} = \frac{X - Min_k}{Max_k - Min_k} \qquad (5)$$

where Max_k is the largest value in distribution k and Min_k is the smallest value in distribution k (see also Janiszewski & Lichtenstein, 1999; Volkmann, 1951). The frequency score represents the percentile rank of value x among all values in distribution k as shown in Equation 6:

$$F_{xk} = \frac{Rank_{xk} - 1}{N_k - 1} \tag{6}$$

where $Rank_{xk}$ is the rank of value x in distribution k and N_k is the number of exemplars in distrbution k. It is these frequency scores that create distribution-density effects. Category rating evaluations of x are assumed to be a linear function of the range-frequency compromise score of x ($RFscore_x$) calculated as a weighted average of R_{xk} and F_{xk} as shown in Equation 7:

$$RFscore_x = wR_{xk} + (1 - w)F_{xk} \tag{7}$$

where w is a weighting parameter.

RF theory predicts s-shaped evaluation functions for any distribution that is dense at the center of the distribution and becomes sparse toward the tails on each side (e.g., normal distributions; see Figure 4, top left). It predicts s-shaped evaluation functions in these cases, however, not because people are less sensitive to differences at greater distances from the RP (i.e., not because of Stevens', 1961, Law), but rather because the density or frequency of exemplars in the distribution produces frequency scores (i.e., F_{xk}; Equation 6) that are s-shaped. The middle portion of the curve is steep, because values are dense at the center of the distribution. The slopes of the lower and upper portions of the s-shaped curve are shallow, because values are sparse at the tails of the distribution.

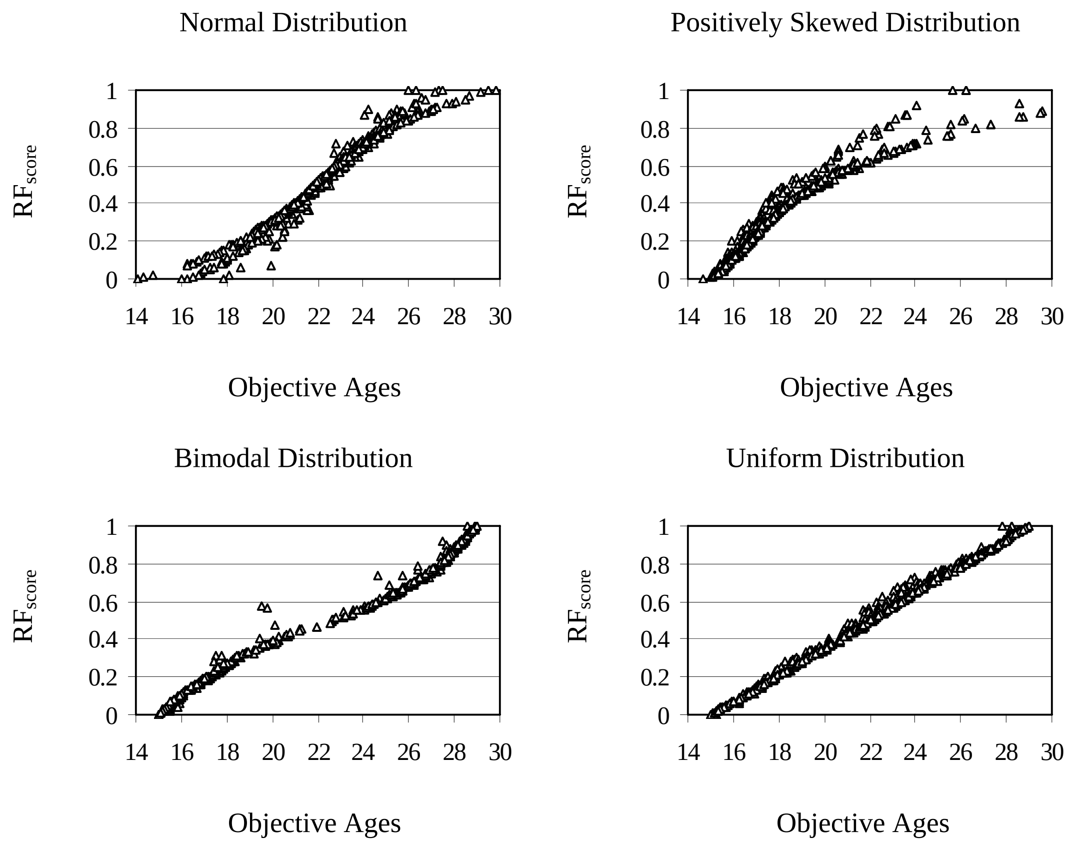

Figure 4: Age evaluation functions predicted by range-frequency (RF) theory. Evaluations follow range-frequency compromise scores (R-F Scores) and produce s-shaped evaluation functions for normal distributions, but not for skewed, bimodal or uniform distributions.

Unlike prototype models of evaluation, RF theory does not predict s-shaped evaluation functions for skewed (Figure 4; upper right), bimodal (Figure 4; lower left; the bimodal distributions in the simulations and proposed experiments below are normal distributions with the tails flipped in giving the end points the greatest density), or uniform (Figure 4; lower right) distributions. Rather, RF theory predicts that evaluation functions will be concave (downward) for positively skewed distributions, because the dense region at the lower end of positively skewed distributions produces larger frequency scores than would have been produced otherwise. It predicts that evaluation functions will be convex (concave upward) for negatively skewed distributions, because the sparse region at the lower end of negatively skewed distributions produces smaller frequency scores than would have been produced otherwise. Evaluation functions ought to flatten somewhat at the center of bimodal distributions because frequency at the center of bimodal distributions is sparse. Evaluations ought to be linear for uniform distributions because frequency is uniformly distributed.

Consistent with both prototype models and RF theory, Niedrich et al. (2001) found that evaluations of prices were s-shaped for normal distributions. But contrary to the predictions of prototype models[†], evaluations of prices in skewed and bimodal distributions were exactly as RF theory predicted they ought to have been. That is, evaluation functions were concave (concave downward) for positively skewed distributions and convex (concave upward) for negatively skewed distributions. Unlike normal distributions, the evaluation functions for bimodal distributions flattened in the sparse middle region. These results are consistent with previous research that has found distribution-density effects on evaluations (Birnbaum, 1974; Mellers & Birnbaum, 1982; Parducci, 1965, 1995; Riskey et al., 1979; Sokolov, Pavlova, & Ehrenstein, 2000; Wedell, Parducci, & Roman, 1989).

Although these results support RF theory, this theory is not able to account for several important findings. Most notably, consistent with CID theory and inconsistent with RF theory, in some cases people seem to have explicit RPs in mind (Holyoak & Mah, 1982). When evaluating body size, for example, people sometimes have a very particular body-size ideal in mind (Irving, 1990; Phelps et al., 1993; Richins, 1991). Unlike CID theory, RF theory also has no explanation for why evaluation functions are often steeper for changes that represent losses than for changes that represent gains (Kahneman & Tversky, 1979; but see Stewart et al., 2006).

Like prototype models of evaluation, CID theory predicts that whenever people verbally compare values to prototypical RPs, comparison-induced biases will generally produce s-shaped evaluation functions regardless of the type of distribution from which values are drawn (see Figure 3). Like RF theory, CID theory predicts that whenever people compare values to other exemplars drawn from the distribution evaluation functions will depend upon distribution density. Like RF theory, CID theory predicts s-shaped evaluation functions for normal distributions, but not skewed, bimodal, or uniform distributions when values are compared to other exemplars drawn from the distribution. Unlike RF theory, CID theory provides a natural account of how greater aversion to losses than affinity for gains could produce asymmetries in sensitivity to losses and gains.

To understand how comparisons to other exemplars in the distribution could create distribution-density effects, consider the positively skewed distribution of ages presented in Figure 5. Filled-in arrows represent biases created by comparisons between values that are closer together than the comparison-suggested difference and that are, therefore, biased apart. Outlined arrows represent biases created by comparisons between values that are farther apart than the comparison-suggested difference and that are, therefore, biased together. Values in dense regions (i.e., 18 – 22 years in Figure 5) are more likely to be closer than the comparison-suggested difference biasing evaluations apart. Values in sparse regions (i.e., 23 – 28 years) are more likely to be farther apart than the comparison-suggested difference biasing evaluations together. This difference in the effects of comparisons within dense versus sparse regions allows comparisons to create density effects.

[†] A model utilizing Equation 4 in which the RP is always updated to be the value presented on the most recent trial or trials (Frederick & Loewenstein, 1999) would also produce distribution-density effects as described here.

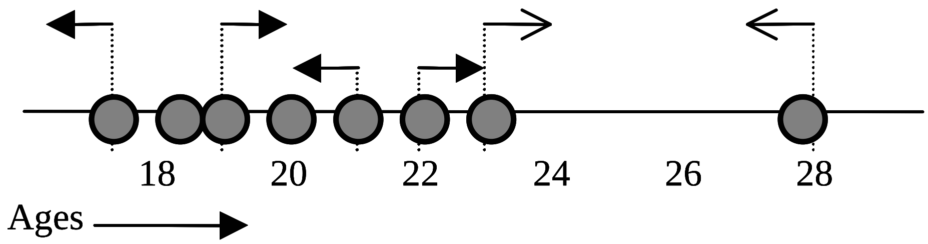

Figure 5: Comparison-induced biases in a positively skewed distribution of ages.

To demonstrate how biases produced by verbal comparisons could produce distribution density-effects, five hundred ages were randomly drawn from each of a normal, a positively skewed, a bimodal, and a uniform distribution. As in the simulation presented in Figure 3, Parameter D in Equations 1a and 1b was set to 3.5 years, Parameter c in Equation 3 was set to 1.5 (these values were chosen for demonstration purposes only; before modeling empirical data, these two values would have to be empirically measured), and Parameter w in Equations 2a and 2b was set to .5 (this value was also chosen for demonstration purposes only; to model empirical data, this value would be found by fitting this free parameter to data). Each of the 500 values was compared to one other randomly drawn value. The results are presented in Figure 6 and they demonstrate that, like RF theory, CID theory predicts s-shaped evaluation functions for normal distributions (upper left), but not skewed (upper right), bimodal (lower left), or uniform (lower right) distributions when values are compared to other exemplar values in the distribution.

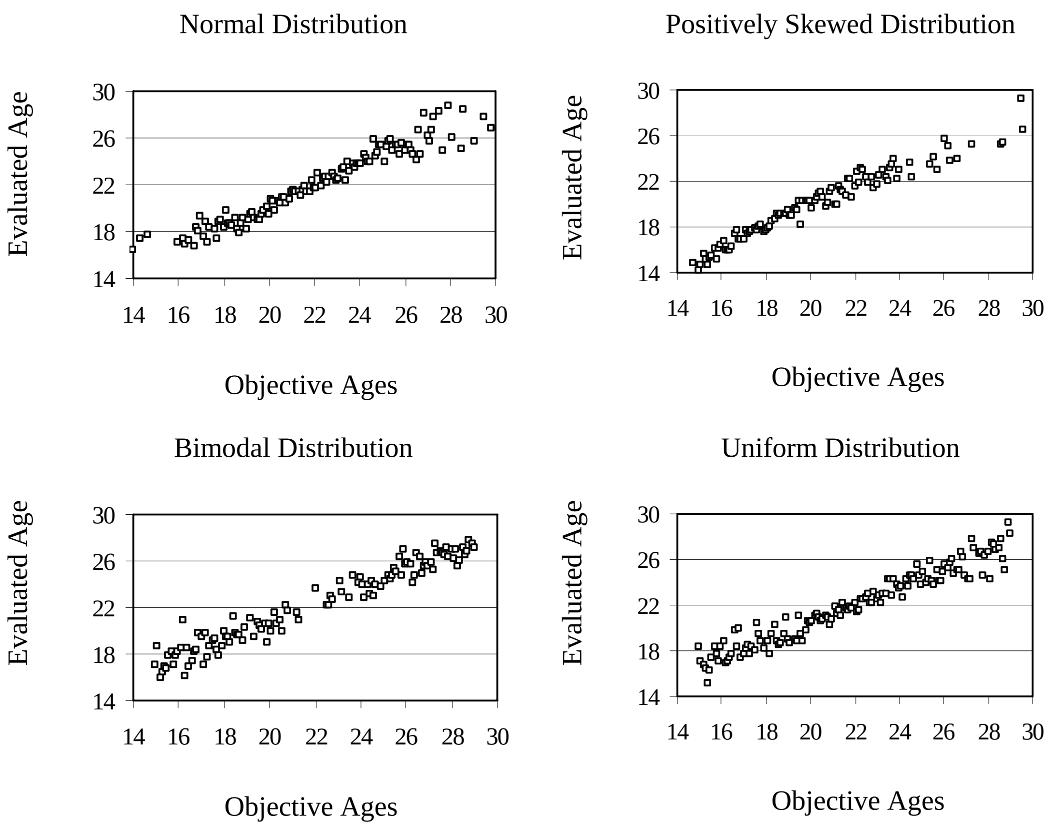

Figure 6. Simulated comparison-biased age evaluations. All values compared to contextual values.

Before I conclude the discussion of distribution-density effects, please note that CID theory does not predict that evaluations will always be affected by distribution density. As demonstrated above, CID theory predicts that evaluation functions will be s-shaped whenever people compare all values to one RP and compared them to no other values. In addition, CID theory predicts no biases, if people do not compare values or Parameter w in Equation 2 is very small.

ANCHORING EFFECTS

In the anchoring effect paradigm, participants compare an unknown value to an arbitrary value called the anchor and then estimate the unknown value (but see Wilson, Houston, Brekke, & Etling, 1996, for a slightly different anchoring effect paradigm). Participants might, for example, be asked whether the length of the Mississippi River is longer or shorter than 3,000 miles and then estimate the length of the Mississippi River. Previous research has consistently found that estimates are biased toward the anchor. In our example, estimates of the length of the Mississippi River would be biased toward 3,000 miles.

CID theory offers one key insight into these anchoring effects, namely that estimates will usually be biased toward comparison-suggested values rather than toward anchor values per se (see Wilson et al., 1996, for a possible exception). This insight should be commonsensical. It bears to reason that if people think that the Mississippi River is shorter than 3,000 miles, then their estimates will not be biased toward 3,000 miles per se, but rather their estimates would be biased toward whatever value they think of as "less than 3,000 miles." That is, their estimates would be biased toward a comparison-suggested difference less than 3,000 miles. Before I describe simulations demonstrating the role that comparison-induced biases might play in creating anchoring effects, it is important to note that CID theory's possible contribution to our understanding of anchoring effects is orthogonal to the contributions offered by previous models of anchoring effects. Following previous models (Anderson, 1965; Huttenlocher et al., 2000), CID theory is a weighted average model of attribute evaluation. Like those previous models, it does not specify the mechanism underlying this weighted averaging. Comparisons could create anchoring effects by changing search strategies (Tversky & Kahneman, 1974), bringing to mind different diagnostic attributes (Strack & Mussweiler, 1997), priming values (Wilson et al., 1996), or changing conversational norms (Schwarz, 1990). Regardless of the particular mechanism that creates this weighted-average biasing, CID theory predicts a slightly different pattern of estimation than the pattern predicted by previous models. CID theory might thereby contribute key insights to our understanding of anchoring effect phenomena, but would not replace these previous models.

As described above, CID theory predicts that comparisons will bias estimates toward comparison-suggested values. If comparison-suggested values are closer to anchor values than are unbiased estimates, then comparisons will bias estimates toward anchor values. In particular, CID theory predicts biases toward anchor values whenever unbiased estimates are farther than a comparison-suggested difference away from the anchor or people judge the unknown value to be approximately the same as the anchor in which case the comparison-suggested value would be the anchor value. Previous research on anchoring effects has typically found such biases toward anchor values. By contrast, however, if comparison-suggested values are farther away from anchor values than are unbiased estimates, then comparisons will bias estimates away from anchor values. In typical anchoring effect scenarios such biases away from anchor values are rare for two reasons. First, because comparison-suggested differences are small relative to the range of people's estimates, most unbiased estimates will be more than a comparison-suggested difference away from the anchor. Most estimates will then be biased toward the anchor value leaving only a few estimates that would not be. Second, many people whose unbiased estimates are less than a comparison-suggested difference away from the anchor will judge the unknown value to be approximately the same as the anchor. These estimates will also be biased toward the anchor. Only in those rare cases in which unbiased estimates are less than a comparison-suggested difference away from the anchor and those unbiased estimates are judged to be different from the anchor (i.e., more or less than the anchor) will estimates be biased away from the anchor.

Even though the vast majority of values are biased toward the anchor value and this result is consistent with both CID theory and alternative theories, CID theory nevertheless makes novel predictions regarding the magnitude of the predicted biases toward the anchor value. In particular, CID theory predicts large biases toward the anchor among unbiased estimates at the extremes, smaller biases toward the anchor for unbiased estimates closer to the comparison-suggested difference away from the anchor, and large biases toward the anchor again as more and more participants start to describe the unknown value as "approximately the same" as the anchor.

To demonstrate how comparison-induced distortions could create anchoring effects, 500 values between 500 and 3100 were randomly drawn from a normal distribution to represent simulated unbiased estimates of the length of the Mississippi River. A histogram of the results is presented in the top panel of Figure 7. I then simulated comparison-biased estimates by comparing all 500 values to an anchor of 2,900 miles. For demonstration purposes only, the comparison-suggested difference (Parameter D in Equation 1) was set at 300 miles. To model real data, one would have to measure this value by asking a control group of participants about the length they imagine when they hear about a river that is shorter than 2,900 miles as described in the measuring comparison-suggested differences section above. Also for demonstration purposes only Parameter c in Equation 3 was set at 0.0002. Again to model real data, one would have to measure this value by asking control groups of participants whether they would describe a variety of values as the same as or less than the anchor value, fit Equation 3 to their responses, and find the best-fitting value for Parameter c. With Parameter c set at 0.0002, 18.6% of values were judged to be approximately the same as the anchor value of 2,900 miles. Parameter w in Equation 2 was set at 0.5. To model real data, Equations 1a through 3 would be fit to the data using the empirically measured values for Parameters D and c to find the best-fitting value for Parameter w. The results are presented in the middle panel of Figure 7. Notice that values were biased toward the anchor value of 2,900 miles. Only 2.2% of values were biased away from the anchor value. I also simulated comparison-biased estimates by comparing all 500 values to an anchor of 600 miles. All parameter were given the same values. With Parameter c set at 0.0002, 20.8% of values were judged to be approximately the same as the anchor value. The results are presented in the bottom panel of Figure 7. Notice that values are biased toward the anchor value of 600 miles. Only 2.0% of values were biased away from the anchor value. Note that these results do not depend upon the particular parameter values I used here for demonstration purposes only. The same basic pattern of results is observed no matter what values these parameters take.

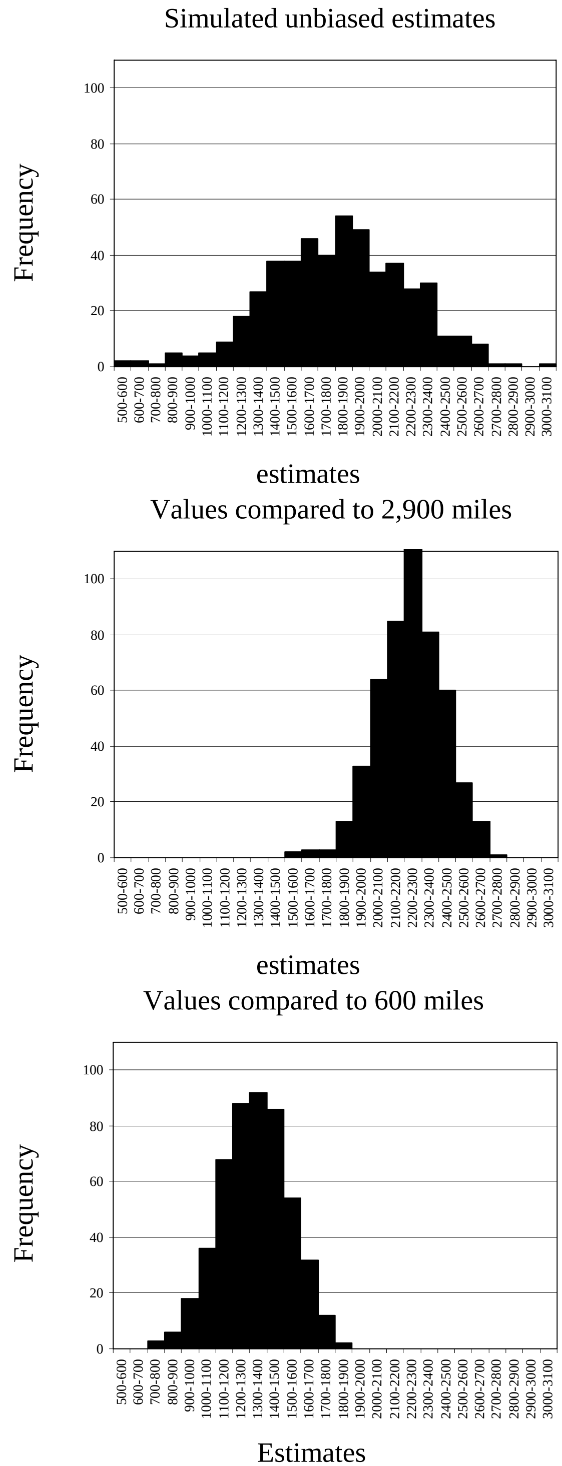

Figure 7. Simulated comparison-induced anchoring effects. The top panel presents a simulated unbiased distribution of estimates of the length of the Mississippi River. The middle panel presents what the distribution would look like, if all values were compared to 2,900 miles. The bottom panel presents what the distribution would look like, if all values were compared to 600 miles

SOCIAL COMPARISONS

Festinger (1954) suggested that people evaluate their own personal-attribute values by comparing their attribute values to the attribute values of their peers. For example, people might evaluate their income, their intelligence, their appearance, their body size, and so forth by comparing themselves to their friends and others who are similar to them. Verbal comparisons of these personal attribute values are likely to produce the same biases as other verbal comparisons. That is, if actual differences are the same size as comparison-suggested differences, then there will be no biases in evaluation. If actual differences are smaller than comparison-suggested differences, then people will overreact to the difference. That is, evaluations will be biased toward the larger comparison-suggested difference. The smaller value will be evaluated smaller than it would have been evaluated otherwise and the larger value will be evaluated larger than it would have been evaluated otherwise. If actual differences are larger than comparison-suggested differences, then people will under react to the difference. That is, evaluations will again be biased toward the comparison-suggested difference, but this time the comparison-suggested difference will be smaller. The smaller value will then be evaluated larger than it would have been otherwise and the larger value will be evaluated smaller than it would have been otherwise. The overall effect of these biases would be such that moderate downward comparisons would produce higher evaluations of the self than would extreme downward comparisons and moderate upward comparisons would produce lower evaluations of the self than would extreme upward comparisons.

Consider, for example, Jennifer and Michelle, the two fictional DePaul University undergraduate students described in Table 1. If Jennifer whose income was $7,868.00 were to compare her income to Michelle's income of $9,000.00, then there would be no biases. The $1,132.00 difference in their salaries would be exactly the same size as the $1,132.00 comparison-suggested difference so reliance on comparison information would create no biases (recall that these were the incomes imagined by actual DePaul University undergraduate women for full-time female students working part time jobs to help pay for school expenses, notice how much larger the incomes imagined by DePaul University undergraduate men were for full-time male students). However, if Michelle had an income of $8,000.00, instead of $9,000.00, and Jennifer were to describe her $7,868.00 income as less than Michelle's $8,000.00 income or Michelle's $8,000.00 as more than her $7,868.00 income (a difference of $132.00), then Jennifer would likely overreact to the difference (unless she were to describe these salaries as "approximately the same" in which case she would under react to the difference), because her evaluations would be biased toward the larger $1,132.00 comparison-suggested difference. Her income would be evaluated smaller than it would have been otherwise and Michelle's income would be evaluated larger than it would have been otherwise. By contrast, if Michelle had an income of $10,000, instead of $9,000.00, and Jennifer were to describe her $7,868.00 income as less than Michelle's $10,000 income (a difference of $2,132.00), then Jennifer would likely under react to the difference. Once again her evaluations would be biased toward the $1,132.00 comparison-suggested difference, but this time the $1,132.00 comparison-suggested difference will be smaller than the $2,132.00 actual difference. Her $7,868.00 income would then be evaluated larger than it would have been evaluated otherwise and Michelle’s $10,000.00 income would be evaluated smaller than it would have been evaluated otherwise.

Simulated biases created when Jennifer compares her own and other people’s salaries are presented in Figure 8. Jennifer’s comparison-biased evaluations of other people’s salaries (Figure 8 left panel) would follow an s-shaped evaluation function. This function is almost identical to the evaluation function shown in Figure 2 Salary Evaluations. The only difference would be that her reference point (RP) would be her own $7,868.00 income, rather than the RP of $50,000.00 shown in Figure 2 Salary Evaluations. Like Figure 2 Salary Evaluations, all other incomes would be evaluated relative to her RP. The top line represents evaluations of other people’s salaries when she describes their salaries as more than her own $7,868 salary. The middle line represents evaluations of other people’s salaries when she describes their salaries as approximately the same as her own salary. The bottom line represents evaluations of other people’s salaries when she describes their salaries as less than her own salary. The Stevens’ (1961) Law formalization of the evaluation function (Kahneman & Tversky, 1979) is also plotted. CID theory predicts a continuous function similar to Stevens’ Law when evaluations from less than, approximately the same, and more than comparisons are averaged together (see Figure 3).

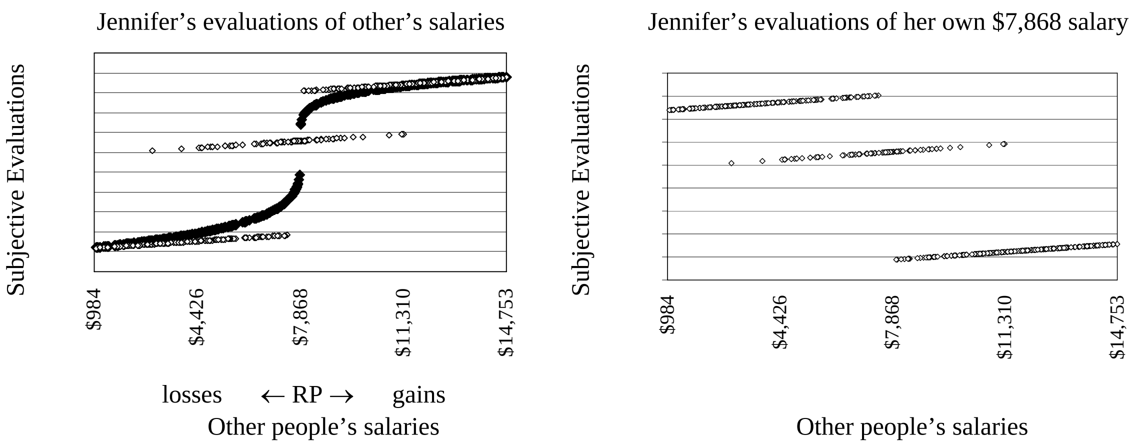

Figure 8. Biases created by social comparisons. The left panel presents comparison-biased evaluations of other people’s salaries. The right panel presents comparison-biased evaluations of Jennifer’s own salary. The middle lines in both panels represent biases created when Jennifer describes her salary as approximately the same as other people’s salaries. The right most lines (top line on the left panel; bottom line on the right panel) represent biases created when Jennifer describes other people’s salaries as more than her salary or her salary as less than other people’s salaries. The left most lines (bottom line on the left panel; top line on the right panel) represent biases created when Jennifer describes other people’s salaries as less than her salary or her salary as more than other people’s salaries.

Jennifer's comparison-biased evaluations of her own salary are plotted in Figure 8 right panel. The top line represents evaluations of her own $7,868 salary when she describes it as more than other people's salaries. The middle line represents evaluations of her salary when she describes it as approximately the same as other people's salaries. The bottom line represents evaluations of her salary when she describes it as less than their salaries. Notice the unique predictions of this CID theory account of social comparisons. This account predicts that comparisons to salaries that are moderately lower than her own salary will cause her to evaluate her salary higher than comparisons to salaries that are much lower than her own salary. Likewise, comparisons to salaries that are moderately higher than her own salary will cause her to evaluate her salary lower than comparisons to salaries that are much higher than her own salary. This pattern of biases might not be observable using category rating measures of evaluation (but see Qian & Brown, 2005), but has been observed in my laboratory using recall of values from memory as the dependent measure.

DECOY EFFECTS

Consumers typically have to pick one option among several alternatives. If they are purchasing an airline ticket, they typically have several flight options to choose from and they have to pick one. If they are renting an apartment, they typically have to decide which of several apartments to rent. If they are purchasing pickles, they typically have to decide which of several brands, flavors, and jar sizes to purchase. These options typically vary along multiple attributes and choices between them often involve trade-off between desirable attribute values. One flight option might leave later in the morning allowing the consumer to sleep later but arrive at the destination later and have a longer layover. One apartment might have a shorter commute to work but cost more in rent. To study how people choose among multi-attribute alternatives such as these, researchers have studied the effects of decoys on choice. Decoys are options inserted into a choice set that people typically do not choose. Despite the fact that people do not choose these decoy options, they affect which of the remaining options people chose.

I will discuss two such decoy effects in this section: the asymmetric dominance effect and the phantom decoy effect. Both effects start with two options that typically vary along two dimensions (e.g., two apartment options that vary in rent and the length of the commute between the apartment and work). One option is better on one dimension, but the other option is better on the other dimension. For example, one apartment might have a shorter commute but cost more in rent than the other apartment. In Figure 9, apartment option 1 has a better commute time (25 minutes compared to 45 minutes for option 2), but apartment option 2 costs less in rent ($575 per month compared to $700 per month). Researchers try to set these values so that people are indifferent between these options. If the options in Figure 9 were set right (the correct values would differ across different groups and would need to be pre-tested), the value of the shorter commute from one apartment would be exactly equal to the extra rent one would pay for that apartment so that if those two options were the only ones presented people would be indifferent between these alternatives. This indifference is represented in Figure 9 by the indifference curve.

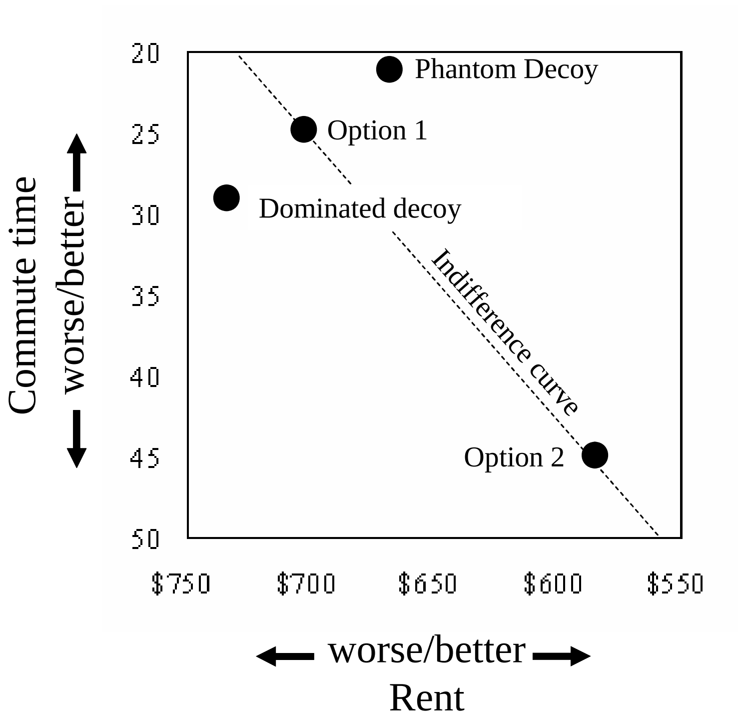

Figure 9. Decoy effects. Options 1 and 2 are designed to be equally attractive (because the benefit of the better value on one dimension is exactly equal to the detriment of the worse value on the other dimension) and so they lie on the indifference curve. Adding a dominated decoy to the option set causes people to choose the option that is similar to the dominated decoy (an effect called the asymmetric dominance effect). Adding a phantom decoy to the option set and marking it unavailable also causes people to choose the option that is similar to the phantom decoy (an effect called the phantom decoy effect).

In the asymmetric dominance effect (Huber et al., 1982), researchers place a dominated decoy in the option set (see Figure 9). This dominated decoy is similar to but worse than one of the other options so that no one would actually choose it. The dominated decoy in Figure 9, for example, is similar to option 1, but it has a longer commute and costs more in rent so that no one would actually choose it if option 1 were available. Nevertheless, its inclusion into the choice set affects the choices people make such that people will be more likely to choose the option that is similar to and better than the dominated decoy. That is, the inclusion of the dominated decoy in the choice set changes people's preferences so that they are no longer indifferent between option 1 and option 2, pushing option 1 above the indifference curve, pulling the indifference curve down, or both.

In the phantom decoy effect (Highhouse, 1996; Pratkanis & Farquhar, 1992), researchers place a phantom decoy in the option set (see Figure 9). This dominating decoy is similar to but better than one of the other options. The phantom decoy in Figure 9, for example, is similar to option 1, but it has a shorter commute and costs less in rent. If this option were available, almost everyone would choose it; but it is marked unavailable, sold out, discontinued, etc. so that it cannot be chosen. Nevertheless, its inclusion into the choice set affects the choices people make such that people will be more likely to choose the option that is similar to the dominating decoy despite the fact that it is not as good as the dominating decoy. The inclusion of the phantom decoy then, like the inclusion of the dominated decoy in the asymmetric dominance effect, changes people's preferences so that they are no longer indifferent between option 1 and option 2. The inclusion of the phantom decoy pulls option 1 above the indifference curve, pushes the indifference curve down, or both.

Choplin and Hummel (2002) argued that CID theory could explain the asymmetric dominance effect; and Choplin and Hummel (2005) discovered a one dimensional analogue of the asymmetric dominance effect and argued that unlike other models in the literature CID theory could account for this one dimensional effect. The basic idea underlying the CID theory account of the asymmetric dominance effect is that verbal comparisons (e.g., better, closer, less expensive, worse, farther, more expensive) between the option that is similar to the dominated decoy (option 1 in Figure 9) and the dominated decoy bias the evaluation of the option that is similar to the dominated decoy. Because the differences between the dominated decoy and the option that is similar to the dominated decoy are typically quite small (i.e., typically smaller than comparison-suggested differences), comparisons bias evaluations apart making the option appear better (and the dominated decoy worse) than it otherwise would and so people choose the dominating option. To simulate these predictions, I modeled the asymmetric dominance effect scenario presented in Figure 9. The choice set included apartment option 1 with a commute of 45 minutes and rent of $575.00 per month, apartment option 2 with a commute of 25 minutes and rent of $700.00 per month, and the dominated decoy with a commute of 29 minutes and rent of $750.00 per month. To model the effects of comparisons on evaluations, I set the comparison-suggested difference (Parameter D in Equation 1) for commute times at 8 minutes (any value larger than 4 minutes would create essentially the same pattern of bias) and the comparison-suggested difference for rent at $100.00 (any value larger than $50.00 would create essentially the same pattern of bias). With these values for Parameter D, CID theory predicts an asymmetric dominance effect as long as Parameter c in Equation 3 takes a value greater than 0.173 for commute times (the probability of describing a 4-minute difference in commute times as "approximately the same" is less than 50.0%) and Parameter c takes a value greater than 0.014 for rents (the probability of describing a $50.00 difference in rent as "approximately the same" is less than 50.0%). These predictions hold regardless of what value the weighting parameter (Parameter w in Equation 2) takes as long as that value is greater than 0.0.

Given this CID theory explanation of the asymmetric dominance effect it might appear at first that CID theory would be unable to account for phantom decoy effects (Highhouse, 1996; Pratkanis & Farquhar, 1992). Since the differences between phantom decoys and the items they dominate are also typically quite small (i.e., they also are typically smaller than comparison-suggested differences), comparison between phantom decoys and options should make the option that is similar to the phantom decoy appear worse. Why then would people choose the similar option?

CID theory could simultaneously account for both the asymmetric dominance effect and the phantom decoy effect only if the unavailability of the phantom decoy were to make people more likely to describe the difference as "approximately the same." Recent empirical results in my laboratory have found evidence that people are more likely to describe options as "approximately the same as" unavailable options than to describe them as "approximately the same as" available options. To simulate these predictions, I modeled the phantom decoy effect scenario presented in Figure 9. As in the simulation of the asymmetric dominance effect described above, the choice set included apartment option 1 with a commute of 45 minutes and rent of $575.00 per month and apartment option 2 with a commute of 25 minutes and rent of $700.00 per month. Instead of the dominated decoy with a commute of 29 minutes and rent of $750.00 per month, however, the phantom decoy with a commute of 21 minutes and a rent of $650.00 was placed in the set, but was not available. To model the effects of comparisons on evaluations, I used the same values for Parameter D described in the simulation of the asymmetric dominance effect above. With these values, CID theory predicts a phantom decoy effect as long as Parameter c in Equation 3 takes a value less than 0.173 for commute times (the probability of describing a 4-minute difference in commute times as "approximately the same" is greater than 50.0%) and Parameter c takes a value less than 0.014 for rents (the probability of describing a $50.00 difference in rent as "approximately the same" is greater than 50.0%). Again, these predictions hold regardless of what value the weighting parameter (Parameter w in Equation 2) takes as long as that value is greater than 0.0. If people were not more likely to describe decoys as "approximately the same" as options when decoys were marked unavailable than when they were not, CID theory would not be able to simultaneously account for both asymmetric dominance effects (Huber et al., 1982) and phantom decoy effects (Highhouse, 1996; Pratkanis & Farquhar, 1992) at the same time.

CONCLUSION

Simulation results demonstrated that biases created by verbal comparisons are capable of explaining or providing insights into a variety of phenomena in the psychology of judgment and decision making. In particular, simulation results demonstrated that biases created by verbal comparisons might help explain or provide insights into s-shaped evaluation functions (Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979), distribution-density effects (Birnbaum, 1974; Niedrich et al., 2001; Parducci, 1965, 1995; Riskey et al., 1979), anchoring effects (Tversky & Kahneman, 1974), social comparisons (Festinger, 1954), and decoy effects such as the asymmetric dominance effect (Huber et al., 1982) and the phantom decoy effect (Highhouse, 1996; Pratkanis & Farquhar, 1992). Because verbal comparisons are ubiquitous whenever people decide between alternatives, future research may find that biases created by verbal comparisons might help explain or provide insights into other decision-making phenomena as well.

Simulation results demonstrated that biases created by verbal comparisons might help explain why evaluation functions are s-shaped (Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979) when people compare to-be-evaluated values to a standard or reference point value. Consistent with previous research on s-shaped evaluation functions, comparisons to a standard or reference point would cause people to over emphasize the importance of small differences and under emphasize the importance of large differences. In addition, some people might describe small differences as "approximately the same." Describing small differences as "approximately the same" would cause people to under emphasize the importance of small differences. A continuous, monotonic function would be created by averaging across the different types of comparisons people make (i.e., less than, approximately the same, and more than comparisons). This continuous function would be concave downward for positive changes from the reference point and convex (concave upward) for negative changes from the reference point. Affective factors cause people to be loss averse (disliking losses more than they like gains). Biases created by verbal comparisons might mediate some of these loss aversion effects, if affective factors cause people to use different comparison words to describe losses than to describe gains, and these comparisons, in turn, affect evaluations.

Simulation results also demonstrated that reliance on information from verbal comparisons might help explain distribution-density effects (Birnbaum, 1974; Niedrich et al., 2001; Parducci, 1965, 1995; Riskey et al., 1979), if people compare to-be-evaluated attribute values to other values drawn from the distribution (i.e., compare them to smaller values sometimes, larger values sometimes, and frequent values—middle values in normal distributions, small values in positively skewed distributions, or large values in negatively skewed distributions—most frequently of all). Because these comparisons would also cause people to over emphasize the importance of small differences and under emphasize the importance of large differences, comparisons to these other values would cause the evaluation function to be s-shaped for values drawn from a normal distribution, concave downward for values drawn from positively skewed distributions, convex (or concave upward) for values drawn from negatively skewed distributions, inverted s-shaped for values drawn from bimodal distributions, and linear for values drawn from uniform distributions.

Simulation results demonstrated that considering the effects of verbal comparisons on estimates might provide insights into anchoring effects. In particular, if estimates are biased by verbal comparisons then when a participant notes that the unknown value is more than (or less than) the anchor value estimates will not be biased toward the anchor value per se, but rather estimates will be biased toward the comparison-suggested value more than (or less than) the anchor value. This CID theory insight into anchoring phenomena is orthogonal to the insights provided by most current accounts of anchoring effects (Schwarz, 1990; Strack & Mussweiler, 1997; Tversky & Kahneman, 1974; Wilson et al., 1996), but is important in its own right.

Biases produced by verbal social comparisons might also provide insights into social comparison phenomena. In particular, simulation results demonstrate that moderate downward comparisons will produce higher evaluations of the self than will extreme downward comparisons and moderate upward comparisons will produce lower evaluations of the self than will extreme upward comparisons.

Because consumers typically compare alternatives whenever they make a choice, biases produced by reliance on verbal comparisons might help explain some consumer decision-making phenomena. In particular, biases produced by verbal comparisons might help explain the effects of decoys on consumer choice. Choplin and Hummel (2002) demonstrated that biases produced by verbal comparisons might explain the asymmetric dominance effect (Huber et al., 1982) and Choplin and Hummel (2005) demonstrated a one-dimensional version of the asymmetric dominance effect and argued that CID theory provided the best account of this phenomenon. Simulation results presented here demonstrated that biases produced by verbal comparisons would be able to simultaneously account for both the asymmetric dominance effect and the phantom decoy effect only if the unavailability of the phantom decoy makes people more likely to describe other options as "approximately the same" as the decoy.

The simulation results presented here demonstrate that the CID theory account of these phenomena is feasible and that biases produced by verbal comparisons are either sufficient to explain phenomena or else provide important insights into phenomena. These simulation results have not shown that biases produced by verbal comparisons are necessary to explain these phenomena nor have they shown that CID theory provides a better account of these phenomena than alternative theories. Arguments that CID theory provides a better explanation of these phenomena than other theories are presented elsewhere.

Importantly, however, the fact that CID theory can account for such a large variety of phenomena is itself evidence in favor of CID theory. Far too often, researchers develop theories to account for one phenomenon or class of phenomena, rather than proposing theories that generalize across a wide variety of phenomena. CID theory follows Helson's (1964) adaptation-level theory and Parducci's (1965; 1995) range-frequency theory in proposing an account of attribute evaluation that generalizes across a wide variety of phenomena. Biases in evaluation produced by verbal comparisons are predicted whenever people compare alternatives. Since such comparisons are ubiquitous, this influence is likely to be pervasive.

REFERENCES

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology, 70*, 394-400.

Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics, 15*, 89-96.

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

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

Choplin, J. M., & Hummel, J. E. (2005). Comparison-induced decoy effects. *Memory & Cognition*, *33*, 332-343.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140.

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

Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology, 78*, 764-771.

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

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

Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes, 65,* 68-76.

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.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*(2), 220-241.

Irving, L. M. (1990). Mirror images: Effects of the standard of beauty on the self- and body-esteem of women exhibiting varying levels of bulimic symptoms. *Journal of Social and Clinical Psychology, 9*, 230-242.

Janiszewski, C., & Lichtenstein, D. R. (1999). A Range Theory Account of Price Perception. *Journal of Consumer Research, 25*(25), 353-368.

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

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.

Kalyanaram, G., & Little, J. D. C. (1994). An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research, 21*, 408-418.

Lavine, H., Sweeney, D., & Wagner, S. H. (1999). Dipicting women as sex objects in television advertising: Effects on body dissatisfaction. *Personality and Social Psychology Bulletin, 25*, 1049-1058.

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

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology, 16*, 148-156.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*, 472-489.

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

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

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

Phelps, L., Johnston, L. S., Jimenez, D. P., Wilczenski, F. L., Andrea, R. K., & Healy, R. W. (1993). Figure preference, body dissatisfaction, and body distortion in adolescence. *Journal of Adolescent Research, 8*, 297-310.

Pratkanis, A. R., & Farquhar, P. H. (1992). A brief history of research on phantom alternatives: Evidence for seven empirical generalizations about phantoms. *Basic and Applied Social Psychology, 13*, 103-122.

Qian, J., & Brown, G. D. A. (2005). *Similarity-based sampling: Testing a model of price psychophysics.* Paper presented at the Twenty-seventh annual conference of the cognitive science society, Stresa, Italy.

Richins, M. L. (1991). Social Comparison and the Idealized Images of Advertising. *Journal of Consumer Research, 18*, 71-83.

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.

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

Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In C. Hendrick & M. Clark, S. (Eds.), *Research Methods in Personality and Social Psychology (Review of Personality and Social Psychology, vol. 11)* (pp. 98-119). Beverly Hills, CA: Sage.

Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*, 1317-1323.

Sokolov, A., Pavlova, M., & Ehrenstein, W. H. (2000). Primacy and frequency effects in absolute judgments of visual velocity. *Perception & Psychophysics, 62*, 998-1007.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science, 133*, 80-86.

Stevens, S. S. (1962). The surprising simplicity of sensory metrics. *American Psychologist, 17*, 29-39.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53*, 1-26.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73*, 437-446.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1130.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Roherer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 279-294). New York: Harper & Row.

Wansink, B. (2004). Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition, 24*, 455-479.

Wedell, D. H., & Parducci, A. (2000). Social Comparison: Lessons from basic research on judgment. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 223-252). New York: Plenum.

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 and Human Decision Processes, 67*, 326-344.

Wilson, T. D., Houston, C. E., Brekke, N., & Etling, K. M. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125*, 387-402.