

Less Is Better: When Low-value Options Are Valued More Highly than High-value Options

CHRISTOPHER K. HSEE*

Graduate School of Business, The University of Chicago, USA

ABSTRACT

This research demonstrates a less-is-better effect in three contexts: (1) a person giving a \$45 scarf as a gift was perceived to be more generous than one giving a \$55 coat; (2) an overfilled ice cream serving with 7 oz of ice cream was valued more than an underfilled serving with 8 oz of ice cream; (3) a dinnerware set with 24 intact pieces was judged more favourably than one with 31 intact pieces (including the same 24) plus a few broken ones. This less-is-better effect occurred only when the options were evaluated separately, and reversed itself when the options were juxtaposed. These results are explained in terms of the evaluability hypothesis, which states that separate evaluations of objects are often influenced by attributes which are easy to evaluate rather than by those which are important. © 1998 John Wiley & Sons, Ltd.

Journal of Behavioral Decision Making, **11**: 107–121 (1998)

KEY WORDS: preference reversal; evaluability; dominance violation; joint and separate evaluation; willingness to pay

INTRODUCTION

Imagine that you give each of two friends a gift. The two friends have similar tastes. They do not know what gift the other receives. Suppose that one gift is more expensive than the other. Is it possible that the friend receiving the less expensive gift considers you to be more generous than the friend receiving the more expensive gift? More generally, suppose that, of two options, H and L, H is more valuable than L. These options are presented to and evaluated by two separate groups of people with similar tastes. Is it possible that the low-value option (L) is perceived to be more valuable than the high-value option (H)?

Normative decision theories assume that people have consistent and well-defined preferences regardless of how the preferences are elicited. In contrast, behavioral decision research suggests that

* Correspondence to: Christopher K. Hsee, Center for Decision Research, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637, USA. E-mail: christopher.hsee@gsb.uchicago.edu

Contract grant sponsors: James S. Kempner Foundation Faculty Research Fund and Fourth Quarter Support, Graduate School of Business, University of Chicago.

preferences are constructed *ad hoc* and can be easily influenced by such subtle contextual manipulations as the framing of a problem in terms of either a gain or a loss (e.g. Kahneman and Tversky, 1979; Tversky and Kahneman, 1981), the framing of a task either as acceptance or rejection (e.g. Shafir, Simonson, and Tversky, 1993), the display of information (e.g. Schkade and Johnson, 1989; Schkade and Kleinmuntz; Johnson, Payne, and Bettman, 1988), the addition of a dominated alternative to the choice set (e.g. Huber, Payne, and Puto, 1982; Simonson, 1989), and the scale on which the preferences are elicited (e.g. Mellers, Ordonez, and Birnbaum, 1992; Mellers *et al.*, 1992).

The present article reports a series of studies where a normatively less valuable option is judged more favorably than its more valuable alternative — a phenomenon to be referred to as the ‘less-is-better’ effect. This research is not the first to investigate this phenomenon. For example, Birnbaum and his associates (1992) found that an inferior risky option (e.g. a 5% chance to win \$96 or \$0) can be valued more highly than a superior risky option (e.g. a 5% chance to win \$96 or \$24). They explained the result by a configural-weight theory, according to which, the weight assigned to a given outcome in a risky prospect depends both on its rank and on its size. Medvec, Madey, and Gilovich (1995) reported that athletes who had just won silver medals tended to be less happy than those who had just won bronze medals. They used the counterfactual reasoning notion to explain this finding. Silver medalists counterfactually thought that they had almost won a gold medal and were therefore disappointed with being the second, while the bronze medalists counterfactually thought that they had almost not won anything, so felt lucky with what they had achieved.

This research examines the less-is-better effect in non-gamble contexts with riskless options, where configural-weight theories or other theories of risky choices do not apply. In addition, the contexts of the studies reported herein do not involve counterfactual reasoning, or at least it is not a necessary condition.

Below, I first report a study that demonstrates a less-is-better effect and then propose an explanation. After that I present evidence showing that the less-is-better effect occurs only when the options are evaluated separately, and that the effect is reversed when the options are evaluated jointly.

STUDY 1

Method

Study 1 used a questionnaire that had two between-subject versions. In both versions participants were asked to imagine that they were about to study abroad and had received a good-bye gift from a friend. In one version participants read

It is a wool coat, from a nearby department store. The store carries a variety of wool coats. The worst costs \$50 and the best costs \$500. The one your friend bought you costs \$55.

Alternatively, the other version describes the gift as follows:

It is a wool scarf, from a nearby department store. The store carries a variety of wool scarves. The worst costs \$5 and the best costs \$50. The one your friend bought you costs \$45.

In both conditions participants were asked how generous they thought the friend was. Answers were given on a 0–6 point scale where 0 indicated ‘not generous at all’, and 6 indicated ‘extremely generous’.¹

¹ After answering the generosity question, participants were also asked how happy they were about the gift and how expensive they felt the gift was. Consistent with their responses to the generosity question, those receiving the scarf reported to be happier and considered the gift to be more expensive than those receiving the coat.

Respondents were 83 students from a large Midwestern university who participated in this study as part of their course requirement. Each respondent received one version of the questionnaire and completed it individually.

Results and discussion

A less-is-better effect emerged from this study. Although the \$55 coat is certainly more expensive than the \$45 scarf, those receiving the scarf considered their gift giver to be significantly more generous than those receiving the coat ($M = 5.63$ and 5.00 , respectively, $t = 3.13$, $p < 0.01$).

One may argue that this result does not represent a less-is-better effect, because a \$45 scarf may be more useful and therefore more valuable than a \$55 coat. However, it is important to note that the dependent variable was generosity of the gift giver, not the usefulness of the gift. Normatively speaking, judgment of a gift giver's generosity should depend on the cost of the gift, not on its usefulness. In other words, the cost of a gift is the only relevant value of concern for generosity. Yet, the result contradicted this intuition, and it is in this sense that it constituted a less-is-better effect. Unlike Study 1, in Studies 2–4, the low-value option (L) will not only cost less than the high-value option (H); it is also strictly worse than H.

The prescriptive implication of Study 1's finding is clear and specific: If gift givers want their gift recipients to perceive them as generous, it is better for them to give a high-value item from a low-value product category (e.g. a \$45 scarf) rather than a low-value item from a high-value product category (e.g. a \$55 coat). The theoretical implication of this result reaches far beyond the context of gift-giving and can shed light on how evaluations are constructed. The following section provides an analysis of the less-is-better effect.

An explanation of the less-is-better effect

The explanation advanced in this section is based on the evaluability hypothesis, which is discussed in greater details in Hsee (1996) and Hsee *et al.* (1997). Briefly, the hypothesis postulates that when a person judges an option in isolation, the judgment is influenced more by attributes that are easy to evaluate than by attributes that are hard to evaluate, even if the hard-to-evaluate attributes are more important. An attribute is said to be hard to evaluate if the decision maker is not aware of its distribution information (e.g. its effective range, its neutral reference point, etc.), and consequently does not know whether a given value on the attribute is good or bad. Conversely, an attribute is said to be easy to evaluate if the decision maker knows its distribution information and thereby knows whether a given value on the attribute is good or bad.

To illustrate, consider an experiment reported in Hsee (1996) involving the evaluations of two job candidates for a computer programmer position. The programmer was expected to use a special computer language called KY. One candidate had completed 70 KY programs and had a college GPA (grade point average) of 2.5; the other had written 10 KY programs and had a college GPA of 4.9. (The experiment was conducted at a university which uses a 5-point GPA system.) When the two candidates were evaluated separately, participants were willing to pay more for the candidate with the higher GPA than the one with more experience. Here, a candidate's experience with KY programs was relatively hard to evaluate. Lacking something to compare it to, most respondents would not know how to evaluate the completion of either 10 or 70 KY programs. In contrast, one's GPA was much easier to evaluate.

In order to apply the evaluability hypothesis to Study 1, we first need to identify the attributes involved. Obviously, one attribute is the actual value of the gift — either \$55 or \$45. This, however, is not the only factor that influences one's judgment. According to norm theory (Kahneman and Miller,

1986; Ritov and Kahneman, 1997), when people evaluate an object in isolation, they often spontaneously think about other objects *in the same category*, and compare the stimulus object to the other objects. Based on this notion, it is assumed that another attribute which respondents used in this study is the relative value of the given gift *in its own product category*, that is, the relative expensiveness of a \$55 wool coat within the wool coat category and the relative expensiveness of a \$45 wool scarf within the wool scarf category.

Thus, the two gifts in Study 1 can be summarized as follows:

	Coat	Scarf
Actual value of the gift	\$55	\$45
The product category	Wool coats	Wool scarves
Relative position of the given gift in its category	Relative position of the \$55 coat in the wool coat category	Relative position of the \$45 scarf in the wool scarf category

Of these factors, the first — the real value of the gift — is hard to evaluate. Without something to compare it to, people would not have a clear idea whether a \$55 (or a \$45) gift is good or bad. The second factor — product category — is a reference, and is not directly evaluated. The last factor — relative position of the given gift in its category — is easy to evaluate. Participants had some distribution information (i.e. price range in this case) of this attribute. Compared with other wool coats ranging in price from \$50 to \$500, a \$55 coat is quite inexpensive. Compared with other wool scarves ranging from \$5 to \$50, a \$45 scarf is quite expensive. The evaluability hypothesis (Hsee, 1996; Hsee *et al.*, 1997) predicts that one's evaluation of the gift would be influenced primarily by the 'relative position' attribute and not by the actual value. The result is clearly consistent with this prediction. It seems that in evaluating a gift, people are neither sensitive to the actual price of the gift, nor to the category of that gift (e.g. whether a coat or a scarf), but they are very sensitive to the relative position of the gift within its category.

The structure of the stimuli in Study 1 can be summarized more generally in terms of the following properties. First, there are two stimulus options to be evaluated, and one option has a higher value. Second, the stimulus options vary on two attributes. One is the actual value of an option (which will be called 'the value of concern'), and the other is the relation of the focal option to some reference(s). In Study 1, the references were the prices of products in the given category, but reference(s) can be other information as well, as will be shown in Studies 2–4. Finally, the references associated with the two options are different. The high-value option is relatively bad compared with its reference(s) and the low-value option is relatively good compared with its own reference(s).

Formally, the preceding analysis can be illustrated as follows,

	H	L
Value of concern:	X_H	X_L
Reference:	R_H	R_L
Relation to reference:	Bad	Good

In the table above, H and L are the high-value and the low-value options, respectively. The value of concern is X , where $X_H > X_L$ ('>' refers to 'is better than' here). The reference is R , where $X_H < R_H$ and $X_L > R_L$. The relation-to-reference attribute describes whether the value of concern (X) appears good or bad in comparison with the reference. It should be noted that reference is not an independent attribute; it can be considered as part of the relation-to-reference attribute.

According to the evaluability hypothesis, as long as Attribute X is hard to evaluate, Option L will be valued more highly than Option H in separate evaluation.

Preference reversal between separate and joint evaluations

So far the discussion has been focused only on the separate evaluation mode, where Options H and L are presented separately and evaluated by different individuals. What will happen in the joint evaluation mode, where Options H and L are presented side-by-side and evaluated by the same individuals?

Previous research indicates that people employ different information as their reference in different situations, depending on which information is available (salient, convenient, vivid, etc.) at the time of evaluation (e.g. Tversky and Kahneman, 1974). In separate evaluation, people cannot compare one option against another, and can only compare the given option to whatever reference is available at the time of the evaluation. In joint evaluation, the two options are juxtaposed, and each option becomes the most salient and convenient reference for evaluating the other option. In this case, people will give less weight to or ignore the reference information they would otherwise use in separate evaluation, and use the alternative option as their primary reference (e.g. Hsee, 1996; Hsee *et al.*, 1996). Through such a comparison, the difference between the stimulus options on any attribute will be transparent, regardless of whether the attribute is hard or easy to evaluate in separate evaluation.

In the case of Options H and L, joint evaluation will make it transparent that H is more valuable than L. Accordingly, H will also be valued more positively than L. In other words, there will be a reversal of preference for the two options between separate and joint evaluation.

STUDY 2

Study 2 departs from Study 1 in a number of ways. First, while Study 1 only examined separate evaluations of options, Study 2 also included a joint evaluation condition. Second, while in Study 1 it was equivocal which option was more useful, there was no ambiguity in Study 2. Finally, Study 2 involved a very different context than gift-giving, and used willingness-to-pay instead of perceived generosity as its dependent variable.

Method

The stimulus options were two hypothetical servings of ice cream, one by Vendor H and one by Vendor L. They were illustrated by drawings as reproduced in Exhibit 1. Observe that Vendor H's

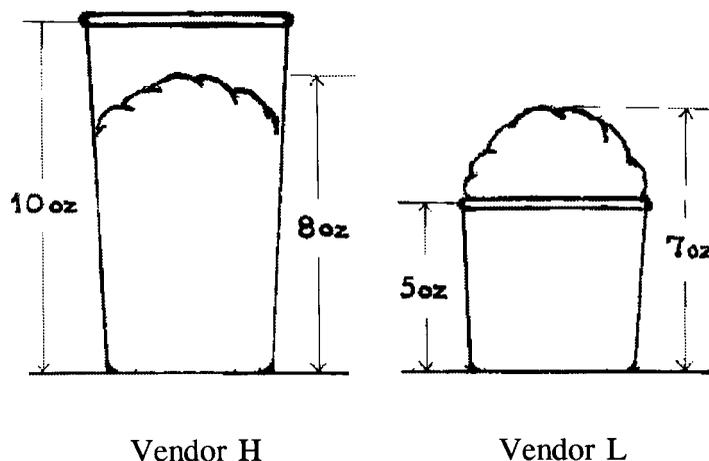


Exhibit 1. Drawings in Study 2

serving was contained in a larger cup than Vendor L's, but Vendor H's serving was underfilled while Vendor L's overfilled. Overall, vendor H's serving contained more ice cream, hence more valuable, than Vendor L's.

The questionnaire for this study had three between-subject versions: separate-H, separate-L and joint. In each separate evaluation version, participants read

Imagine the following scenario: It is summer in Chicago. You are on the beach at Lake Michigan. You find yourself in the mood for some ice cream. There happens to be an ice cream vendor on the beach. She sells Haagen Dazs ice cream by the cup. For each serving, she uses [a 10 oz cup and puts 8 oz of ice cream] {a 5 oz cup and puts 7 oz of ice cream} in it.

The words in the two types of brackets were for the two separate evaluation versions, respectively. These instructions were followed by one of the two drawings reproduced above. Respondents were then asked

What is the most you are willing to pay for a serving?
\$ ____.

The joint evaluation version was parallel to the separate evaluation versions except that participants were told that there were two vendors on the beach. Participants were presented with the information, including the drawings, for both vendors. They were then asked

What is the most you are willing to pay for a serving from each vendor?
Vendor A: \$ ____.
Vendor B: \$ ____.

(On the original questionnaire, Vendor H was labeled 'Vendor B' and Vendor L was labeled 'Vendor A'. The drawing of Vendor L's serving was printed to the right of the drawing of Vendor H's serving. It is for consistency with the other studies reported here that I will always use the label 'H' to denote the high-value option and the label 'L' to denote the low-value option, and always place H to the left of L.)

Respondents were 69 college students recruited from a large Midwestern university. Each respondent received one version of the questionnaire and completed it individually. Upon completion every participant received a cash payment.

Results and discussion

Analysis of the stimulus options

The two ice cream servings can be interpreted as varying on the following dimensions:

	Vendor H's	Vendor L's
Amount of ice cream (value of concern):	8 oz	7 oz
Cup size (reference)	10 oz	5 oz
Filling (relation to reference):	Under (bad)	Over (good)

Here, the amount of ice cream should be the only value of concern to the consumer. Cup size served as a reference in separate evaluation. Whether the cup was overfilled or underfilled represented the relation between the value-of-concern (the amount of ice cream) and the reference (cup size). Note that this structure paralleled that in Study 1.

It was predicted that Vendor L's serving would be valued more in separate evaluation, and Vendor H's serving would be valued more in joint evaluation. The reasons for these predictions are as follows. In separate evaluation, even though the amount of ice cream was the only value of concern, this attribute was hard to evaluate independently. Lacking a comparison, it was difficult to determine whether a serving with 8 oz ice cream (or with 7 oz ice cream) was good or not. The size of the cup provided a reference point. Presumably, respondents compared the amount of ice cream against that reference. The resulting 'relation' attribute (over- versus under-filling) was easy to evaluate: Most consumers would find an overfilled serving appealing, and an under-filled serving unappealing. Thus, the over-filled serving (Vendor L's) would be valued more highly in separate evaluation. In joint evaluation, respondents could compare one option against the other, could recognize that Vendor H's serving contained more ice cream than Vendor L's, and would make their valuations accordingly.

Exhibit 2. WTP prices for Vendor H's and Vendor L's servings in Study 2

Evaluation mode	Vendor H's	Vendor L's	<i>t</i> -value
Separate evaluation	\$1.66	\$2.26	2.47, $p < 0.05$
Joint evaluation	\$1.85	\$1.56	4.31, $p < 0.01$

Willingness-to-pay prices

As predicted, there was a clear less-is-better effect in separate evaluation, and a clear preference reversal between joint and separate evaluations (see Exhibit 2). In the separate evaluation versions, Vendor L's serving was valued significantly more than for Vendor H's, even though Vendor L's serving contained less ice-cream. In joint evaluation, the effect was reversed: Vendor H's serving was valued significantly more.²

The results of Study 2 corroborate many existing studies showing that people exhibit inconsistent, even reversed preferences, when using different elicitation methods. Slovic and Lichtenstein (1968; Lichtenstein and Slovic, 1971) provided the first systematic demonstration of preference reversals. When evaluating two gambles — one having a large chance of winning a small payoff (the p-bet) and the other a smaller chance of winning a larger payoff (the \$-bet) — people prefer the p-bet if asked to choose a gamble, but prefer the \$-bet if asked to set a selling price for each gamble. Preference reversals have also been demonstrated between other response conditions, such as choice versus matching (e.g. Tversky, Sattath, and Slovic, 1988; Slovic, Griffin, and Tversky, 1990), and attractiveness ratings versus pricing (e.g. Mellers *et al.*, 1992).

The preference reversal demonstrated in Study 2 departs from those conventionally studied reversals in two ways. First, in the previous studies, the stimulus options typically involve a tradeoff; for example, the p-bet and the \$-bet involve a tradeoff between probability and payoff. In Study 2, Serving H was strictly better than Serving L. Second, the conventionally studied reversals typically occur between conditions that use different evaluation scales, for example, acceptability (as in choice) versus money (as in pricing). In Study 2, the evaluation scale was held constant — always willingness to pay.

² Responses in Studies 2, 3, and 4 were not made on bounded scales. To control for extreme responses, the following criterion was set *a priori*: any responses over 3 standard deviations from the mean would be excluded. Only one response in Study 4 and none in any of the other studies was excluded under this criterion.

Recently, preference reversals between joint and separate evaluation using a constant evaluation scale have been documented in a number of studies (e.g. Bazerman, Loewenstein, and White, 1992; Bazerman *et al.*, 1994; Hsee, 1996; Kahneman and Ritov, 1994; Irwin *et al.*, 1993). In an experiment by Bazerman *et al.* (1994), MBA students evaluated various hypothetical job offers. Some of the offers entailed higher salary to oneself but were perceived to be unfair because the company would pay others with similar background even more; other options entailed less salary to oneself but did not seem unfair because the respondent would be paid as much as others. In joint evaluation, the greater-payoff/unfair options were deemed more acceptable, but in separate evaluation, the smaller-payoff/fair options were rated more acceptable. As another example, consider an experiment by Hsee (1996) involving the evaluations of two used music dictionaries. One contained 20,000 entries and had a torn cover; the other contained 10,000 entries and was like new. When the two dictionaries were evaluated jointly, the more comprehensive dictionary was valued more. When they were evaluated separately, the newer-looking dictionary was favored.

As in the traditionally studied preference reversals, the stimulus options used in studies to demonstrate preference reversals between joint and separate evaluation also involve tradeoffs along different attributes. For example, the two dictionaries in Hsee (1996) involved a tradeoff between number of entries and cosmetic condition. The two options in Bazerman *et al.* (1992) involved a tradeoff between absolute payoff and fairness. However, the stimulus options in Study 2 involved no tradeoffs. It should be noted that, unlike fairness in dispute resolutions or job offers, the filling of an ice-cream serving is not a normatively relevant factor in deciding how much to pay for the serving. Fairness in dispute resolution or job offers — e.g. whether the company would pay me as much as my colleague — implies consequences beyond monetary payoffs, e.g. my future social status in the company, my relationship with the colleague, etc. Therefore, fairness is a normatively legitimate consideration in deciding whether to accept the job. The filling of an ice cream serving, on the other hand, does not imply anything other than the amount of ice cream. Study 2 shows that even in situations without a tradeoff between the two options, the low-value option can still be favored over the high-value option in separate evaluation.

STUDY 3

Study 3 was a replication of Study 2 with two main variations. First, in Study 2 participants were shown the exact measures (in oz) of the cup size and the amount of ice cream. In reality, however, people rarely know these numerical values. In Study 3, the ice cream servings were still illustrated by drawings, but these drawings were not accompanied by size measures; instead, they were in real size.

Second, in Study 2, the amount of ice cream never equaled the size of the cup; the cup was either overfilled or underfilled. In Study 3, one of the servings was neither overfilled nor underfilled.

Method

The questionnaire for Study 3 was essentially identical to that for Study 2. The drawings of the two servings are reproduced in Exhibit 3.

Note that Vendor H's serving was neither overfilled nor underfilled, and Vendor L's serving was overfilled. Again, Vendor H's serving contained more ice cream.

The original drawings were larger than those shown here. The two cups were approximately 3.5 inches and 2.1 inches tall, respectively, and the ice cream was approximately 3.5 inches and 3.1 inches tall, respectively. Participants were told that the drawing(s) represented the actual size of the ice cream serving(s). As in Study 2, respondents were assigned to either the separate evaluation or the

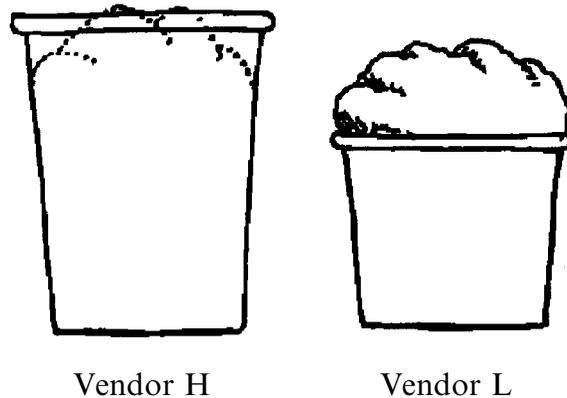


Exhibit 3. Drawings in Study 3. (The original drawings are larger than the ones reproduced here)

joint evaluation version. Those in the separate evaluation versions were given information for only one of the servings and indicated their willingness-to-pay price. Those in the joint evaluation version learned about both servings and indicated their willingness-to-pay prices.³

Respondents were 98 students recruited from a large Midwestern university to complete this and other questionnaires in exchange for a chance to win a \$100 lottery.

Results and discussion

Analysis of the stimulus options

The two servings of ice cream can be interpreted as:

	Vendor H's	Vendor L's
Amount of ice cream (value of concern):	3.5 inches	3.1 inches
Cup size (reference):	3.5 inches	2.1 inches
Filling (relation to reference):	None (neutral)	Over (good)

In other words, the structure of the two servings took the form of:

	H	L
Value of concern:	X_H	X_L
Reference:	R_H	R_L
relation to reference:	Neutral	Good

where $X_H > X_L$, $X_H = R_H$ and $X_L > R_L$. Note that unlike Studies 1 and 2, only one stimulus option in Study 3 differed from its reference; the other option was at the same level as its reference.

³ In the joint evaluation version of Studies 3 and 4, participants were first asked for which option they were willing to pay more before being asked how much they were willing to pay for each option. Because responses to the first question were redundant with those to the second, they will not be discussed further. One respondent in Study 3 and two in Study 4 gave contradictory answers, i.e. said that they would pay more for one option but gave a higher WTP value to the other, and these responses were excluded.

Willingness-to-pay prices

Once again, there was a less-is-better effect as well as a preference reversal between separate and joint evaluation (see Exhibit 4): In separate evaluation, participants were willing to pay more for Vendor L's serving. In joint evaluation, they were willing to pay more for Vendor H's serving.

Exhibit 4. WTP prices for Vendor H's and Vendor L's servings in Study 3

Evaluation mode	Vendor H's	Vendor L's	<i>t</i> -value
Separate evaluation	\$1.80	\$2.23	2.16, $p < 0.05$
Joint evaluation	\$2.02	\$1.81	2.55, $p < 0.05$

A comparison of Study 3 with Study 2 suggests that it was not just the underfilling of Vendor H's serving that drove the less-is-better effect in Study 2; the overfilling of Vendor L's serving alone would render the same effect. In addition, it was not necessary to provide numerical measures; the same result was obtained even if respondents only saw the real-size drawings of the ice cream servings.

STUDY 4

Method

Study 4 sought to replicate the findings of the other studies in yet another context. The stimulus options were two sets of dinnerware. Set H included more pieces but some of them were broken; Set L had fewer pieces and all were intact. They were described as follows:

	Set H includes 40 pcs	Set L includes 24 pcs
Dinner plates:	8, all in good condition	8, all in good condition
Soup/salad bowls:	8, all in good condition	8, all in good condition
Dessert plates:	8, all in good condition	8, all in good condition
Cups:	8, 2 of them are broken	
Saucers:	8, 7 of them are broken	

Note that Set H included all the pieces included in Set L, plus several more intact pieces; thus, set H was better than Set L. (On the original questionnaire, Set H was referred to as Set A, and Set L as Set B.)

As in previous studies, the questionnaire for this study had two separate evaluation versions and one joint evaluation version. In all the versions, participants were asked to imagine that they were shopping for a dinnerware set and that there was a clearance sale in a local store where dinnerware regularly ran between \$30 and \$60 per set. In the separate evaluation condition, participants were presented with the information of and indicated their WTP prices for either Set H or Set L; in the joint evaluation condition, for both sets.

Respondents were 104 college students from a large Midwestern university who participated in this and other experiments for payment.

Results and discussion

Analysis of the stimulus options

The two dinnerware sets can be interpreted as follows:

	Set H	Set L
Number of intact pieces (value of concern):	31	24
Total number of pieces (reference):	40	24
Relation to reference:	Bad	Neutral

Here, the only attribute that should be of concern to the consumer was the number of intact pieces. In separate evaluation, participants presumably used the total number of pieces in the set as a reference. For Set H, the number of intact pieces (31) was worse (fewer) than its reference (40); for Set L, the number of intact pieces (24) was the same as its reference (24).

Note that like Study 3, the value of only one option in Study 4 differed from its reference. But instead of the value of Option L being better than its reference, it was the value of Option H that was worse than its reference. In other words, the stimuli in Study 4 took the form of:

	H	L
Value of concern:	X_H	X_L
Reference:	R_H	R_L
Relation to reference:	Bad	Neutral

where $X_H > X_L$, $X_H < R_H$ and $X_L = R_L$.

Willingness-to-pay prices

Again, the low-value option was valued more favorably than the high-value alternative in separate evaluation: WTP values were significantly higher for Set L than for Set H although Set L included fewer intact pieces than Set H. In joint evaluation, the effect was reversed: WTP values were significantly higher for Set H.

Exhibit 5. WTP prices for dinnerware set H and dinnerware set L in Study 4

Evaluation mode	Set H	Set L	<i>t</i> -value
Separate evaluation	\$23.25	\$32.69	3.91, $p < 0.001$
Joint evaluation	\$32.03	\$29.70	2.15, $p < 0.05$

Although Study 4 was concerned with a very different product category than Study 3, the structures of the stimuli in those two studies were very similar. The only difference was that in Study 3, one of the options was *superior* to its reference while the other equalled its reference, but in Study 4, one of the options was *inferior* to its reference while the other equalled its reference. Despite the difference, similar results were obtained in both studies. The findings of these studies extended the observations from Studies 1 and 2 by demonstrating that the less-is-better effect could occur not only when both options differ from their respective references, but also when only either one of the options differs from its reference.

GENERAL DISCUSSION

This research compares valuations of options between separate evaluation and joint evaluation where one of the options is always more valuable or better than the other. It demonstrates that preferences were reversed between these two evaluation modes, and, more interestingly, that the low-value option was valued more highly than the high-value alternative in separate evaluation.

Strictly speaking, a higher valuation of the low-value option in separate evaluation is not a violation of the dominance principle. It is only from the omniscient perspective of the experimenter that one option is more valuable than the other. Participants in the separate evaluation conditions did not realize this relation, and therefore did not knowingly 'prefer' the low-value option.

Nevertheless, these findings are of both theoretical and practical importance. Theoretically, they corroborate the growing evidence in the descriptive decision literature (see Slovic, 1995, for a review) by showing that preferences are neither consistent nor stable; they are constructed ad hoc and depend heavily on whatever comparison information is available at the time of the evaluation. Specifically, people use different information as their reference points in the joint evaluation mode than in the separate evaluation mode. Even in separate evaluation, the reference point associated with one option often differs from that associated with the other option. If the reference associated with the high-value option is better than the high-value option itself, and/or if the reference associated with the low-value option is worse than the low-value option itself, the less-is-better effect may emerge.

The observations from the current studies also lend support to the evaluability hypothesis, suggesting that in separate evaluation, the impact an attribute has on one's judgment may not be proportional to the normative importance of the attribute, but rather to its evaluability: An easy-to-evaluate attribute, such as the overfilling or underfilling of a serving of ice cream, may have much greater impact on one's judgment than a hard-to-evaluate attribute, such as the actual amount of ice cream, even though any reasonable consumer would consider the latter to be more important.

In all of the studies reported in this article, the value of concern was hard to evaluate, and its relationship with the reference — whether better than, worse than, or the same as the reference — was easy to evaluate. Therefore, the valuations of the stimulus options in separate evaluation were dominated by the 'relation-to-reference' attribute.

Two clarifications are in order here. First, the less-is-better effect will be less likely to occur if the value of concern itself is easy to evaluate. For example, for people in the ice cream business, the actual amount of ice cream in a serving may be easy to evaluate. They know how much a serving with a given amount of ice cream is worth, without having to compare it to the size of the cup. If those people had served as participants in Study 2 or 3, Vendor L's ice cream may have been valued *less* than Vendor H's even in separate evaluation.

Second, the less-is-better effect occurs not only because the 'relation-to-reference' attribute is easy to evaluate, but also because the stimulus options have *different* relations with their respective references — one is better than the reference and the other is worse than the reference, or one is better (or worse) than the reference and the other is equal to the reference. If the stimulus options have similar relations with their respective references — for example, both better than their references, then the less-is-better effect will be less likely to occur. For instance, in the ice cream case, if both servings were overfilled, there would probably not be a less-is-better effect in separate evaluation, and not be a preference reversal between joint and separate evaluations, either. These propositions are examined in greater details elsewhere (Hsee *et al.*, 1997).

The reader may wonder whether the preference reversal documented herein can be accounted for by the prominence principle (Tversky *et al.*, 1988). This principle was designed originally to explain preference reversals between choice (in which respondents choose between two alternatives) and matching (in which respondents are presented with both alternatives but some information about one of

the alternatives is missing, and the respondents' task is to fill in the missing information to make the two options equally attractive). According to the prominence principle, the most prominent attribute in the choice set looms larger in choice than in matching. This is a convincing explanation for choice-matching reversals as well as for many other choice-judgment reversals, but it may not be the most viable explanation for the findings of the present research. First, the prominence principle is concerned with options that involve tradeoffs on multiple dimensions; the stimulus options in the current research do not. The prominence principle does not account for the less-is-better effect. Moreover, choice-matching preference reversals are very different from preference reversals between joint and separate evaluations. For example, in matching the decision maker is presented with both stimulus options and can perform careful tradeoff analyses (Tversky *et al.*, 1988), in separate evaluation the decision maker is presented with only one option and cannot perform tradeoff analyses. Therefore, the mechanics involved in the present studies are perforce distinct from those accounted for by the prominence principle.

The studies presented in this article also yield important practical implications. As Study 1 suggests, if a person wants to impress her gift recipient with her generosity, she should give a top-of-the-line product from a low-price product category rather than a bottom-of-the-line product from a high price product category. Likewise, Studies 2 and 3 illustrate that people are insensitive to the actual value of the stimulus option, or to the reference or frame associated with the stimulus option, but are sensitive to the relative position of the stimulus within that frame. Thus, for instance, if a conference speaker estimates that only about 50 people will attend his talk, but wants to create the impression that his talk is very popular, he should give the talk in a small conference room with 50 or fewer seats than a large conference room with, say, more than 100 seats. A small room will easily be 'overfilled' with the 50 attendees and look crowded. A large room will be 'underfilled' with the 50 attendees and look empty.⁴

Study 4 implies that a perfect product will be valued more favorably than an imperfect product that is objectively more valuable. Consider, for example, two boxes of candies. One contains 24 pieces, all of which are exquisite chocolates. The other box contains 40 pieces, which are the same 24 exquisite chocolates plus 16 low-quality candies. Even though the second box is more costly and normatively better than the first box, the first box will probably cast a better impression on the consumer.

The above examples are based on the assumption that the alternative scenarios are presented separately and cannot be compared directly. If these scenarios were presented jointly, then the proposed effect would be reversed.

Another practical implication of the current research concerns the relative validity between joint evaluation and separate evaluation as a method to predict consumer satisfaction. Suppose, for example, that a marketer has developed two alternative versions of a certain product (e.g. two alternative versions of a certain software application), and wants to predict which version consumers will prefer. The marketer has two ways to run this test: (a) giving samples of the two versions to two different groups of consumers and then comparing their satisfaction ratings (separate evaluation), or (b) giving samples of the two versions to a single group of consumers and asking them to indicate which option is more satisfying (joint evaluation). The current research implies that these methods are likely to yield different conclusions.

Which is the more valid method? The first (separate evaluation) method is apparently imperfect, because, as we now know, the option favored in separate evaluation may well be different from the option that is objectively more valuable or better. Then, is the second (joint evaluation) method more valid? Not necessarily, either. The answer depends on what it means by 'more valid'. If the criterion of validity is for the evaluations to be consistent with the objective quality of the stimulus options, then

⁴ This example was inspired by a discussion with Max Bazerman.

joint evaluation is indeed more valid than separate evaluation, because in joint evaluation people are more likely to discover subtle differences between the stimulus options which may be masked in separate evaluation, and more likely to know which option is objectively more valuable.

However, if the criterion of validity is for the evaluations to be consistent with the consumer's ultimate experience at the time of consumption, then the first (separate evaluation) method may be more valid. More often than not, the consumption of a product is in separate evaluation mode; for example, a person receiving a coat will not automatically compare it with a scarf. People who bought and enjoy an overfilled serving of ice cream do not actively think about what they would feel if the ice cream is served in a larger cup and the cup is underfilled. People who enjoy a 10-choice Sunday brunch only focus on the quality of the given foods and do not think about what they would feel if there were more choices and some of the choices were mediocre. In those cases, preferences elicited in the second (joint evaluation) method may be at odds with consumers' experiences at the time of consumption. For example, suppose that a novelist has written two versions of a novel but can publish only one version. She asks a critic to read both versions (joint evaluation) and to tell her which is better. Even if the critic has the same taste as future readers of the novel, it is possible that the critic would recommend publishing one version, while it is the other version that would enjoy a greater success if published instead. The reason for this dilemma is simple. The critic is in a joint evaluation mode when making the recommendation, but the future readers will be in a separate evaluation mode when reading the novel.

In short, preferences elicited in joint evaluation may differ from preferences elicited in separate evaluation. Preferences elicited in joint evaluation are often more consistent with the objective quality of the evaluated options, but preferences elicited in separate evaluation are often more predictive of consumers' ultimate experience. As for ultimate experience, an objectively inferior or less valuable option may be better than an objectively superior or more valuable alternative.

ACKNOWLEDGEMENTS

The author benefited from discussions with Sanjay Dhar, Rick Larrick and Joe Nunes at different stages of this project. This research was supported by the James S. Kempner Foundation Faculty Research Fund, and Fourth Quarter Support from the Graduate School of Business, the University of Chicago.

REFERENCES

- Bazerman, M. H., Loewenstein, G. F., and White, S. B. 'Reversals of preference in allocation decisions: Judging an alternative versus choosing among alternatives', *Administrative Science Quarterly*, **37** (1992), 220–40.
- Bazerman, M. H., Schroth, H. A., Shah, P. P., Diekmann, K. A., and Tenbrunsel, A. E. 'The inconsistent role of comparison others and procedural justice in reactions to hypothetical job descriptions: Implications for job acceptance decisions', *Organizational Behavior and Human Decision Processes*, **60** (1994), 326–52.
- Birnbaum, M. H., Goffey, G., Mellers, B. A., and Weiss, R. 'Utility measurement: Configural-weight theory and the judge's point of view', *Journal of Experimental Psychology: Human Perception and Performance*, **18** (1992), 331–46.
- Hsee, C. K. 'The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives', *Organizational Behavior and Human Decision Processes*, **67** (1996), 247–57.
- Hsee, C. K., Blount, S., Loewenstein, G. F., and Bazerman, M. H. 'Preference reversals between joint and separate evaluation of options: A theoretical analysis', Working paper, University of Chicago, 1997.
- Huber, J., Payne, J. W., and Puto, C. 'Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypotheses', *Journal of Consumer Research*, **9** (1982), 90–98.
- Irwin, J. R., Slovic, P., Lichtenstein, S., and McClelland, G. H. 'Preference reversals and the measurement of environmental values', *Journal of Risk and Uncertainty*, **6** (1993), 5–18.

- Johnson, E. J., Payne, J. W., and Bettman, J. R. 'Information displays and preference reversals', *Organizational Behavior and Human Decision Processes*, **42** (1988), 1–21.
- Kahneman, D. and Miller, D. T. 'Norm theory: Comparing reality to its alternatives', *Psychological Review*, **93** (1996), 136–53.
- Kahneman, D. and Ritov, I. 'Determinants of stated willingness to pay for public goods. A study in the headline method', *Journal of Risk and Uncertainty*, **9** (1994), 5–38.
- Kahneman, D. and Tversky, A. 'Prospect theory: An analysis of decision under risk', *Econometrica*, **47** (1979), 263–91.
- Lichtenstein, S. and Slovic, P. 'Reversal of preferences between bids and choices in gambling decisions', *Journal of Experimental Psychology*, **89** (1971), 46–55.
- Medvec, V. H., Madey, S., and Gilovich, T. 'When less is more: Counterfactual thinking and satisfaction among Olympic medalists', *Journal of Personality and Social Psychology*, **69** (1995), 603–10.
- Mellers, B. A., Chang, S., Birnbaum, M. H., and Ordóñez, L. D. 'Preferences, prices and ratings in risky decision making', *Journal of Experimental Psychology: Human Perception and Performance*, **18** (1992), 347–61.
- Mellers, B. A., Ordóñez, L. D., and Birnbaum, M. H. 'A change-of-process theory for contextual effects and preference reversals in risky decision making', *Organizational Behavior and Human Decision Processes*, **52** (1992), 331–69.
- Ritov, I. and Kahneman, D. 'How people value the environment: Attitudes versus economic values', in Bazerman, M. H. *et al.* (eds), *Environment, Ethics and Behavior*, San Francisco: New Lexington Press, 1997.
- Schkade, D. A. and Kleinmuntz, D. N. 'Information display and choice processes: differential effects of organization, form, and sequence', *Organizational Behavior and Human Decision Processes*, **57** (1994), 319–37.
- Schkade, D. A. and Johnson, E. J. 'Cognitive processes in preference reversals', *Organizational Behavior and Human Decision Processes*, **44** (1989), 203–31.
- Shafir, E., Simonson, I., and Tversky, A. 'Reason-based choice', *Cognition*, **49** (1993), 11–36.
- Simonson, I. 'Choice based on reasons: the case of attraction and compromise effects', *Journal of Consumer Research*, **16** (1989), 158–74.
- Slovic, P. 'The construction of preference. 102nd Annual Convention of the American Psychological Association Distinguished Scientific Contribution Award Address', *American Psychologist*, **50** (1995), 364–71.
- Slovic, P., Griffin, D., and Tversky, A. 'Compatibility effects in judgment and choice', in Hogarth, R. M. (ed.), *Insights in Decision Making: Theory and Applications*, Chicago: University of Chicago Press, 1990.
- Slovic, P. and Lichtenstein, S. 'The relative importance of probabilities and payoffs in risk taking', *Journal of Experimental Psychology Monographs*, **78** (1968), (3, Pt 2).
- Tversky, A. and Kahneman, D. 'Judgment under uncertainty: Heuristics and biases', *Science*, **185** (1974), 1124–31.
- Tversky, A. and Kahneman, D. 'The framing of decisions and the psychology of choice', *Science*, **211** (1981), 453–58.
- Tversky, A., Sattath, S., and Slovic, P. 'Contingent weighting and judgment and choice', *Psychological Review*, **95** (1988), 371–84.

Author's biography:

Christopher K. Hsee (PhD, Yale University) is Associate Professor of Behavioral Sciences, the Graduate School of Business, University of Chicago. His areas of interest include judgment and decision making, consumer behavior, and cross-cultural research.

Author's address:

Christopher K. Hsee, Center for Decision Research, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637, USA.