

Social Psychology

# "Less Is Better" in Separate Evaluations Versus "More Is Better" in Joint Evaluations: Mostly Successful Close Replication and Extension of Hsee (1998)

Andrew J. Vonasch<sup>1</sup>, Wing Yiu Hung<sup>2</sup>, Wai Yee Leung<sup>2</sup>, Anna Thao Bich Nguyen<sup>2</sup>, Stephanie Chan<sup>2a</sup>, Bo Ley Cheng<sup>2</sup>, Gilad Feldman<sup>2b</sup>

<sup>1</sup> University of Canterbury, Christchurch, New Zealand, <sup>2</sup> Department of Psychology, University of Hong Kong, Hong Kong

Keywords: more is better, replication, judgment and decision making, biases and heuristics, less is better

<https://doi.org/10.1525/collabra.77859>

## Collabra: Psychology

Vol. 9, Issue 1, 2023

We conducted a preregistered close replication and extension of Studies 1, 2, and 4 in Hsee (1998). Hsee found that when evaluating choices jointly, people compare and judge the option higher on desirable attributes as better ("more is better"). However, when people evaluate options separately, they rely on contextual cues and reference points, sometimes resulting in evaluating the option with less as being better ("less is better"). We found support for "less is better" across all studies ( $N = 403$ ; Study 1 original  $d = 0.70$  [0.24, 1.15], replication  $d = 0.99$  [0.72, 1.26]; Study 2 original  $d = 0.74$  [0.12, 1.35], replication  $d = 0.32$  [0.07, 0.56]; Study 4 original  $d = 0.97$  [0.43, 1.50], replication  $d = 0.76$  [0.50, 1.02]), with weaker support for "more is better" (Study 2 original  $d = 0.92$  [0.42, 1.40], replication  $d_z = 0.33$  [.23, .43]; Study 4 original  $d = 0.37$  [0.02, 0.72], replication  $d_z = 0.09$  [-0.05, 0.23]). Some results of our exploratory extensions were surprising, leading to open questions. We discuss remaining implications and directions for theory and measurement relating to economic rationality and the evaluability hypothesis. Materials/data/code: <https://osf.io/9uwms/>

Would people ever think less is worth more? Hsee (1998) demonstrated a "less is better" effect that in some situations people evaluated less valuable options as being more generous, preferred, and better. For example, people evaluate a small but overflowing cup of ice cream to be worth more compared to people who evaluate a larger underfilled cup containing more total ice cream. The less is better effect is frequently cited as a demonstration of the violation of the principle that people have stable, consistent preferences they use to make economic valuations and judgments across contexts (e.g., Bazerman et al., 1999; Kahneman, 2003; Mercier & Sperber, 2011; Shafir & LeBoeuf, 2004; Slovic et al., 2007). It has implications for the efficacy and ethicality of nudging (Davidai & Shafir, 2018; Thaler & Sunstein, 2009), and for the ways in which context shapes preferences for food (e.g., Parrish & Beran, 2014; Pattison & Zentall, 2014). However, to our knowledge, only one of the paradigms used in this research has been directly replicated (Klein et al., 2018), so the broader replicability of this effect remains uncertain.

According to the evaluability hypothesis, the less is better effect occurs when different people make judgments of

difficult-to-evaluate qualities of things (Hsee, 1996; Hsee et al., 1999; but see Sher & McKenzie, 2014 for an alternative explanation). For example, people evaluate a small but overflowing cup of ice cream to be worth more than a larger underfilled cup containing more total ice cream only when they evaluate those cups one at a time. It is relatively difficult to evaluate the amount of ice cream in a cup, but much easier to evaluate how full a cup is, so people tend to simplify their evaluations by comparing the amount of ice cream to the size of cup. Clever choice architects can exploit this cognitive strategy to design situations leading to worse choices being considered better or more valuable. However, when people directly compare the two options, they can discern that more ice cream is worth more (Hsee, 1998). Thus, the evaluability hypothesis makes distinct predictions for situations in which people judge two things jointly versus separately. Separate judgments may reveal a less is better effect, as shown by Hsee (1998), and directly replicated by Klein, et al. (2018). However, joint judgments should reveal a more is better effect, found in Hsee (1998), but to our knowledge never directly replicated until now.

a Shared first author

b Corresponding author: Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR; [gfeldman@hku.hk](mailto:gfeldman@hku.hk)

## Choice of Replication Target: Why Replicate Hsee (1998)?

Replication is a core feature of science that makes scientific knowledge more credible than rumor, dogma, etc. Recent large scale replication efforts found lower than expected replication success rates in psychology (Klein et al., 2018; Nosek et al., 2015) and other fields (Camerer et al., 2016), even of studies published in high impact journals (perhaps especially in high impact journals; Camerer et al., 2018; Schimmack, 2012). Thus, numerous scholars have called for a “credibility revolution” (Vazire, 2018) and “making replications mainstream” (Zwaan et al., 2018), to run additional replication studies to assess and enhance the credibility of psychological research findings (Funder et al., 2014; Gelman & Loken, 2013; Munafò et al., 2017; Nosek & Lakens, 2014).

We chose to replicate three studies in Hsee (1998) because the article provided the original evidence for the “less is better” effect, and thus has been very influential and cited more than 400 times. Moreover, the studies are methodologically clean, simple, elegant examples of the phenomenon. Also, although the original studies were conducted using in-person methods, the designs were readily compatible with online administration via Amazon Mechanical Turk (MTurk), which allows inexpensive, rapid collection of large, high powered samples that, although still WEIRD, are more diverse than the samples of North American college students that once dominated psychological research (Buhrmester et al., 2016; Henrich et al., 2010; Paolacci et al., 2010).

There was also an additional theoretical reason to replicate these studies. To date, only Study 1’s paradigm has been replicated (Klein et al., 2018) and that paradigm lacked a key control condition necessary to consider the preferences to be truly inconsistent. We therefore chose to replicate Study 1, but also Studies 2 and 4, which included the control conditions needed to test the evaluability hypothesis. We did not replicate Study 3, because Study 3 was very similar to Study 2.

### Extensions to Hsee (1998)

Although the main aim of this project was to directly replicate Hsee (1998), we also added several exploratory extensions. Study 1 added a joint evaluations condition, matching the design of Studies 2 and 4. In addition to a direct replication of Hsee (1998) Studies 1, 2, and 4, we added extension questions to each study. One question asked about the perceived attractiveness of the products. We predicted people would rate more valued products as more attractive. Another question asked people’s affective reaction to the products. We predicted people would have more positive affective reactions to products they preferred and valued.

### Overview of Studies

We closely replicated Studies 1, 2, and 4 of the four studies in Hsee (1998). The procedure was as similar as possi-

ble to the original, including using the same vignettes and main dependent measures. However, the original student samples were replaced with a single online MTurk sample. We added exploratory measures (on a separate page after the main measure to avoid influencing responses to the main measures), and participants completed all three studies in a single session, whereas originally they were run separately. Because they were run together, we took precautions to avoid contaminating the different conditions. All participants were assigned the same condition across the three studies (i.e., high, low, or joint), and the order of studies was counterbalanced. The core predictions were that in the separate condition, participants would judge less is better, but in the joint condition, they would judge more is better.

We made all materials available on the OSF (<https://osf.io/9uwns/>). We crowdsourced the pre-registration analysis with two teams of two coauthors, and we report our analysis using the most conservative and stringent exclusion criteria of the two (discussed below; <https://osf.io/6z32c/>). Data collection was completed before analyzing the data. All collected conditions are reported. All variables collected for this study are reported and included in the provided data.

We note that the above pattern of results was predicted across all variables collected, although two of the preregistrations made these predictions clearer than the others.

### Procedure

All three vignettes are the same as in the original Hsee (1998) study, and instructions were close to identical, except for explaining that there would be three vignettes (instead of one) and that there would be several questions following each vignette (instead of one). The study consisted of three hypothetical scenarios: goodbye gift scenario, ice cream vendor scenario and dinnerware set sales scenario. Participants were randomly assigned to different conditions of each experiment using the Qualtrics randomizer. The order of vignettes was counterbalanced so any order of studies was possible. Participants read each vignette and answered the main dependent measure, followed on the next page by the extension measures for that vignette (except as noted below).

### Participants

A power analysis using G\*power indicated we would achieve 80% power to detect the original effect size with 90 participants in Study 1, and 122 in Studies 2 and 4. For details on power analysis calculations, see the supplemental materials. To increase power and account for possible exclusions, we increased the planned sample to 400.

Four hundred and three American participants were recruited online via MTurk CloudResearch/TurkPrime (Litman et al., 2016) and received payment after completing the Qualtrics survey. Three participants were excluded from data analysis for refusing to participate in the study, leaving a sample size of 400 (192 males; 208 females;  $M_{\text{age}}$  39.3;  $M_{\text{range}}$  18-77). Participants were assigned to the same con-

dition across the three scenarios, with 135 participants in the high condition, 131 in the low condition, and 134 in the joint condition.

### Replication of Study 1: The Gift scenario

In Study 1, participants were randomly assigned into one of three conditions, consistent across the three scenarios. In all conditions, participants were asked to imagine that they were about to study abroad and received a goodbye gift from a friend. In the coat condition, they received a coat; in the scarf condition they received a scarf, and in the joint condition, they were offered a choice between the coat or the scarf. We predicted a successful replication of the less is better effect across the coat and scarf conditions: that people would prefer a less expensive scarf (more expensive from the lower price range) to a more expensive coat (less expensive of the higher price range) when they judge them separately. The original study only included the coat and scarf conditions, not the joint condition, which we added as an extension to mirror the other two scenarios.

Coat condition presented a coat on the lower end of a \$50- \$500 price range ( $N = 135$ ):

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.

The Scarf condition was similar, but instead of a coat they were given a scarf which was on the higher end of a much lower price range \$5 to \$50 ( $N = 131$ ):

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.

We added a third joint condition that was missing in Study 1 but was added in Studies 2 and 4 ( $N = 133$ ). The joint condition read:

Your friend asks that you choose between the following two gift options:  
**Gift 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 is suggesting costs \$45.  
**Gift B: 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.

### Measures

#### Main Measure: Perceived generosity

“How generous do you think the friend is?” (0 – Not generous at all, 6 – Extremely generous).

### Extension measures

Perceived expensiveness (manipulation check)

“How expensive do you feel the gift is?” (0 – *Not expensive at all*, 6 – *Extremely expensive*). This question was presented on the same page but always after the generosity question.

Affect toward receiving the gift

“How would you feel about receiving this scarf/coat gift?” In the joint conditions: “How would you feel about receiving Gift A wool scarf” “How would you feel about receiving Gift B wool coat?” (0 – *Extremely negative*, 6 – *Extremely positive*).

Attractiveness of the gift

In the scarf and coat conditions: “How attractive is this scarf/coat you received as a gift”. In the joint conditions: “How attractive is Gift A wool scarf as a gift?” and “How attractive is Gift B wool coat as a gift?” (0 – *Not attractive at all*, 6 – *Extremely attractive*).

Gift preference

Only in the joint condition, we asked: “Finally, which of the two gifts would you prefer?” as a binary choice “Gift A wool scarf” or “Gift B wool coat.”

### Results and Discussion

#### Preregistered Hypotheses: Separate Evaluations (predicted less is better effect)

The less is better effect replicated successfully. As predicted, participants rated the scarf ( $M = 5.45$ ,  $SD = 0.95$ ) as a more generous gift than the coat ( $M = 4.18$ ,  $SD = 1.55$ ), Welch’s  $t(223) = 8.11$ ,  $p < .001$ , ( $d = 0.99$ , 95% CI: [0.72, 1.25]). This successful replication had a larger effect size than in the original study (original  $d = 0.70$ , 95% CI: [0.24, 1.15]) and the ManyLabs replication ( $d = 0.78$ , 95% CI: [0.74, 0.83]; Klein et al., 2018).

#### Extensions: Separate evaluations

Participants in the separate conditions rated the scarf as more expensive than the coat (Table 1). This result was unexpected, because the coat was literally more expensive than the scarf. However, it suggests participants perhaps interpreted the question to mean how expensive the gift is relative to other gifts of its kind.

Participants in all conditions rated the scarf more attractive and felt better about receiving it (Table 1), which is consistent with the less is better effect, insofar as “more attractive” and “positive feelings” carry a similar meaning (i.e., “better gifts”) to the main dependent variable, generosity.

**Table 1. Study 1 extensions**

		Scarf M (SD)	Coat M (SD)	t	p	df	d or d <sub>z</sub> [95%CI]
Attractive	Separate	5.06 (1.29)	4.09 (1.51)	5.65	<.001	259.58	.69 [.44, .94]
	Joint	4.69 (1.49)	3.83 (1.73)	4.52	<.001	133	.39 [.21, .57]
Affect	Separate	5.27 (1.18)	4.85 (1.34)	2.74	.007	261.39	.34 [.09, .58]
	Joint	4.90 (1.29)	4.11 (1.54)	5.06	<.001	133	.44 [.26, .61]
Expensive	Separate	4.95 (1.04)	2.61 (1.39)	15.57	<.001	248.72	1.91 [1.57, 2.23]
Generosity	Separate	5.45 (0.95)	4.18 (1.55)	8.11	<.001	222.87	0.99 [0.72, 1.25]
Preference	Joint	n=85 63%	n=49 37%	Binomial	.002	N=134	95%CI scarf: [55-72%] 95%CI coat: [28-45%]

Note. *t* = Welch's *t* for separate, Paired-samples *t* for joint. We used JAMOVI to calculate 95% CIs around Cohen's *d* in each study. Across all studies, we report Cohen's *d* for between subjects tests and *d<sub>z</sub>* for paired-samples tests (Lakens, 2013).

### Extensions: Joint evaluations

We predicted, based on the evaluability hypothesis, that in the joint condition we would observe “more is better” effects across domains. The evaluability hypothesis predicts this specifically in the domain of generosity, but related constructs also seemed likely to produce more is better effects: preferences between gifts, ratings of the attractiveness of the gifts, and feelings about the gifts. We did not ask participants in the joint condition to rate which gift was more expensive, because we presumed they would simply compare the prices and observe that the coat was literally more expensive. We also asked participants about generosity in the joint condition— yet there was an oversight in the way we framed the question: We asked about the friend's overall generosity, rather than about the generosity of each potential gift. Thus, we were unable to test any predictions about generosity in the joint evaluations condition.

We note there was also an accidental inconsistency in the phrasing of the joint condition. For the scarf, the friend is “suggesting” an option, whereas for the coat the friend has “bought” one. This could potentially have influenced the results below, but we do not think there is any clear pattern of results that would follow from this inconsistency. If anything, we would have thought participants who misconstrued this as meaning the coat was already purchased might find themselves trying to justify keeping it, to avoid having to rudely reject the already purchased gift—but this was not supported by the data. Any future research using this paradigm should correct this inconsistency.

### Exploratory Analyses

#### Attractiveness and positive affect

Counter to expectations, participants in the joint condition rated the scarf as more attractive and felt better about receiving it than the coat. Insofar as these constructs carry a similar meaning to generosity (as argued above), this is consistent with a less is better effect in the joint condition. This is surprising, because for generosity Hsee (1998)'s model would predict a more is better effect, and these seem like similar constructs.

#### Preference

Counter to expectations, but consistent with ratings of attractiveness and positive affect, participants in the joint condition preferred the scarf (63%) to the coat (37%; binomial: *p* = .002).

### Discussion

The less is better effect replicated successfully in the separate evaluations condition. As predicted, in separate judgments, people believed a \$45 scarf (a product ranging from \$5 to \$50) was a more generous gift than a \$55 coat (a product ranging from \$50 to \$500). This is consistent with the evaluability hypothesis, and a successful replication of Hsee (1998)'s Study 1.

However, the extensions to Study 1 yielded surprising results that introduce some nuance about the evaluability hypothesis as tested by this paradigm. In the separate evaluations condition, we found that people's judgments of generosity also tracked their judgments of the other variables: people thought the gift was more generous, but also more attractive, and they felt more positive affect upon receiving it. However, surprisingly, we also found that in the joint condition, people rated the less expensive gift as more attractive, felt more positive affect upon receiving it, and indicated they preferred it. There are thus two possibilities: either generosity in the joint condition would have behaved as expected by the evaluability hypothesis (but different from the other variables), or it would have behaved like the other variables (contradicting the evaluability hypothesis).

There are a few possible interpretations of these surprising findings. One is that these judgments of attractiveness, positive affect, and preference are importantly different from ratings of generosity. From this perspective, one might rate the same gift as more attractive, as preferred, and as making you feel better to receive it—but also rate that gift as less generous than the preferred gift. If so, more research distinguishing these seemingly similar judgments would be useful to help understand why generosity judgments are so unexpectedly dissimilar to preference, attractiveness, and affective judgments. Another interpretation

is that people would likely rate the gift they prefer, think is more attractive, and feel more positively about as more generous. People may think the cheapest coat is such low quality that it is not particularly worth having, particularly compared to a top-quality scarf—consistent with the price quality heuristic (e.g., Gerstner, 1985). If this latter interpretation is correct, that would undermine the validity of this paradigm for testing the evaluability hypothesis, because it would mean this paradigm does not produce a reversal of judgments from the separate evaluations to the joint evaluations conditions.

Unfortunately, due to our oversight in designing the generosity measure in the joint condition, we were unable to test which interpretation is correct in this study. We believe it is important that future research test this because Study 1 is the most widely used paradigm for testing the less is better effect and the only one to be directly replicated (prior to the present research)—it was used in the ManyLabs replication project (Klein et al., 2018). If it turns out this paradigm does not demonstrate a reversal of evaluations, it would mean the evidence testing the evaluability hypothesis is considerably more limited than previously thought. This heightens the importance of testing the other paradigms below.

### Replication of Study 2: Ice Cream Scenario

This was a close replication of Hsee (1998)'s Study 2. Participants were randomly assigned to one of three vignettes about ice cream vendors: two separate evaluation conditions and one joint evaluation condition. We predicted a less is better effect for separate evaluations of less ice cream overflowing a small cup versus more ice cream filling less of a bigger cup, but a more is better effect for joint evaluations. As two extensions, we further predicted that perceived attractiveness of the ice cream and positive feelings toward its price would have the same pattern as willingness to pay for the ice cream.

Underfilled ice cream condition ( $n = 133$ ):

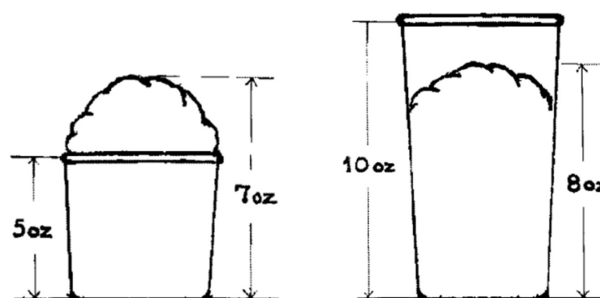
It is summer in Chicago. You are on the beach at Lake Michigan. You find yourself in the mood for some ice cream. There happens to be an ice cream vendor on the beach. She sells Haagen Dazs ice cream by the cup. For each serving, she uses a 10 oz cup and puts 8 oz of ice cream in it.

Overfilled ice cream condition ( $n = 131$ ):

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 5 oz cup and puts 7 oz of ice cream in it.

Joint evaluation condition ( $n = 133$ ; see [Figure 1](#)):

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 two ice cream vendors on the beach.



**Figure 1. Sizes of ice creams in the Study 2 vignette**

Note. Images are from Hsee (1998)

They both sells Haagen Dazs ice cream by the cup. For each serving, Vendor A uses a 5 oz cup and puts 7 oz of ice cream in it. And Vendor B uses a 10 oz cup and puts 8 oz of ice cream in it.

### Measures

#### Main Measure: Willingness to pay

“What is the most you are willing to pay for a serving? (in USD)”.

#### Extensions

##### Attractiveness of ice cream

“It just so happens that the vendor is selling the ice-cream for (amount they were willing to pay) USD, please rate the attractiveness of the ice cream offered.” on a scale from (0 not attractive at all) to (6 extremely attractive).

##### Affect toward the price

“It just so happens that the vendor is selling the ice-cream for (amount they were willing to pay) USD, please rate your feelings about the price of ice cream offered.” (0 = Extremely negative; 6 = Extremely positive).

##### Preference (joint condition only)

Would you be willing to pay more to Vendor A or Vendor B?

#### Study Specific Exclusion Criteria

The two teams designing the preregistrations chose slightly different exclusion criteria. One group specified additional exclusions criteria in Study 2, to match the criteria of Hsee (1998), but the other group did not specify this. The lead author, who was not part of either team, independently decided that matching the exclusion criteria of Hsee (1998) was the most appropriate, as it removes two outliers listing extreme values of willingness to pay for an ice cream cup. However, the results do change somewhat depending on criteria chosen, so we have elected to first report the results using Hsee's criteria, and second to report the results without excluding the outliers. Hsee's criteria: Answers over three standard deviations from the mean will-

ingness to pay across all conditions were excluded (responses over \$10.39 for an ice cream). Two responses in the joint condition and two in the separate conditions were excluded for this reason. Additionally, participants giving contradictory responses would have been excluded, but no responses were contradictory. Exploratory analyses inspired by a reviewer comment revealed varying the exclusion criteria in this study slightly did not have any meaningful effect on the final analyses' significant or effect sizes.

## Results

### ***Separate evaluation condition (preregistered prediction: less is better effect)***

The less is better effect successfully replicated. In the separate evaluation condition, participants were willing to pay more for the smaller cup ( $M = 3.99$ ,  $SD = 1.49$ ) than for the larger cup ( $M = 3.54$ ,  $SD = 1.34$ ), Welch's  $t(257.95) = 2.59$ ,  $p = .010$ , ( $d = 0.32$ , 95% CI [0.07, 0.56]). The effect size was smaller than in the original study ( $d = 0.74$ , 95% CI [0.12, 1.35]).

*Without excluding the two outliers*, the less is better effect did not successfully replicate. In the separate evaluation condition, we found no support for willingness to pay more for the smaller cup ( $M = 3.99$ ,  $SD = 1.49$ ) than for the larger cup ( $M = 3.86$ ,  $SD = 2.98$ ), Welch's  $t(198.52) = 0.47$ ,  $p = .640$ , ( $d = 0.06$ , 95% CI [-0.18, 0.30]).

### ***Joint evaluation condition (preregistered prediction: more is better effect)***

The more is better hypothesis was supported. In the joint condition, as predicted, participants were willing to pay more for the larger amount of ice cream ( $M = 4.28$ ,  $SD = 1.63$ ) than for the smaller amount ( $M = 3.76$ ,  $SD = 1.53$ ),  $t(131) = 6.92$ ,  $p < .001$ , ( $d_z = 0.60$ , 95% CI [.42, .79]). The effect size was smaller than in the original study ( $d = 0.92$ , 95% CI [0.42, 1.40]).

*Without excluding outliers*, the more is better hypothesis was still supported. In the joint condition, as predicted, participants were willing to pay more for the larger amount of ice cream ( $M = 4.35$ ,  $SD = 1.72$ ) than for the smaller amount ( $M = 3.88$ ,  $SD = 1.82$ ),  $t(133) = 5.61$ ,  $p < .001$ , ( $d_z = 0.48$ , 95% CI [.31, .66]). The effect size was smaller than in the original study.

## Extensions

In the two extensions testing the attractiveness of the ice cream and participants' feelings toward the price we found no support for differences between the conditions (Table 2). There was no support for differences in attractiveness or in affect in any condition. Thus, contrary to our predictions, we found no evidence of a pattern similar to willingness to pay for attractiveness and affect.

## Discussion

Both replication predictions were supported when using Hsee's (1998) exclusion criteria. When making separate

evaluations, the less is better effect occurred: people were willing to pay more for the smaller overfilled ice cream than the larger under-filled one. However, when making joint evaluations, the more is better effect occurred: people were willing to pay more for a larger amount of ice cream if they could jointly compare the smaller and larger amounts. Thus, the prediction that the less is better effect is specific to the context of separate evaluations was supported. Unlike Study 1, we failed to find support for our extension predictions regarding attractiveness and feelings.

Two teams designed this study separately, and there were slight variations in the exclusion criteria for the preregistered analyses. The senior authors decided to use the stricter exclusion criteria endorsed by Hsee (1998), though this did change the significance level of one analysis. Using a more inclusive criterion allowing two major outliers into the analyses weakened the less is better effect in Study 2. This highlights the importance of both removing flexibility in data analyses, and in using stringent criteria to clean data whenever possible.

### **Replication of Study 4: Dish Sets Scenario**

This was a close replication of Hsee (1998) Study 4. Participants were randomly assigned to one of three vignettes about dish sets. We predicted a successful replication: a less is better effect for separate evaluations of a small set of dishes versus a bigger set containing some broken dishes, but a more is better effect for joint evaluations of the same sets. As two extensions, we further predicted that perceived attractiveness of the dishes and positive feelings toward them would exhibit the same pattern of judgments as willingness to pay for the dishes.

## Vignettes

### **Joint condition**

The vignettes of the two sets of dishes used in the experiment are provided in Table 3. Participants read the following: "Imagine that you are shopping for a dinnerware set and that there is a clearance sale in a local store where dinnerware regularly runs between \$30 and \$60 per set." They were then presented with both sets of dishes.

### **Lower value condition**

In the separate lower value condition, participants only read about lower value Set B.

### **Higher value condition**

In the separate higher value condition, participants only read about higher value Set A. We note that Set A is higher value than Set B in the sense that it includes all of the items in Set B, plus additional items, though some of those are broken.

**Table 2. Study 2 results**

		Underfilled M (SD)	Overfilled M (SD)	t	p	df	d or $d_z$ [95%CI]
Attractive	Separate	4.76 (1.14)	4.84 (1.04)	0.60	.551	260.54	.07 [-.17, .31]
	Joint	4.50 (1.17)	4.61 (1.19)	0.95	.344	131	.08 [-.09, .25]
Affect	Separate	4.62 (1.23)	4.80 (1.06)	1.31	.192	257.69	.16 [-.08, .40]
	Joint	4.52 (1.12)	4.65 (1.08)	1.27	.208	131	.11 [-.06, .28]

Note. *t* = Welch's *t* for separate, Paired-samples *t* for joint.

**Table 3. Study 4: Dish sets**

	Higher value: Set A includes 40 pcs	Lower value: Set B 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	/

## Measures

### Main Measure: Willingness to pay

“What is the most you are willing to pay for Set A/Set B? (in USD)”.

### Extensions

Joint condition: Willingness to pay more

In the joint condition, participants were asked “Would you pay more for Set A or Set B?” (on the same page as willingness to pay).

Participants were asked two follow up questions on their willingness to pay:

### Attractiveness of the dinnerware

“It just so happens that the store is selling Set A/Set B for (amount they were willing to pay) USD, please rate the attractiveness of the dinnerware Set A offered” (0 = *Not attractive at all*; 6 = *Extremely attractive*).

### Affect toward the price

“It just so happens that the store is selling Set A/Set B for (amount they were willing to pay) USD, please rate your feelings about the price of the dinnerware Set A offered” (0 = *Extremely negative*; 6 = *Extremely positive*).

## Results

### Separate Evaluations Condition (preregistered prediction: less is better effect)

The less is better effect was supported, see [Table 4](#). As predicted, in the separate evaluation condition, partici-

pants were willing to pay more for the smaller set of unbroken dishes ( $M = \$33.70$ ,  $SD = 11.45$ ) than for the larger set that included some broken dishes ( $M = \$23.96$ ,  $SD = 13.94$ ), Welch's  $t(257.03) = 6.23$ ,  $p < .001$ , ( $d = 0.76$ , 95% CI [0.50, 1.02]). The effect size was smaller than in the original study ( $d = 0.97$ , 95% CI [0.43, 1.50]).

### Joint Evaluations Condition (preregistered prediction: more is better effect)

The more is better effect was not supported, see [Table 4](#). Although participants were willing to pay more for the larger set of dishes ( $M = \$31.67$ ,  $SD = 15.96$ ) than for the smaller set ( $M = \$30.35$ ,  $SD = 12.75$ ), the difference was not significant,  $t(133) = 1.26$ ,  $p = .210$ , ( $d_z = 0.11$ , 95% CI [-0.06, 0.28]). The effect size was smaller than in the original study ( $d = 0.37$ , 95% CI [0.02, 0.72]).

### Extensions

Willingness to pay more

In the joint evaluation condition, most participants (59%) were willing to pay more for the larger set than for the smaller one (41%), binomial test:  $p = .047$ .

Separate evaluations

As predicted, in the separate evaluation conditions, participants' ratings of the dishes' other qualities matched their willingness to pay. Participants rated the larger set ( $M = 3.91$ ,  $SD = 1.41$ ) as less attractive than the smaller set ( $M = 4.63$ ,  $SD = 1.08$ ), Welch's  $t(250.07) = 4.70$ ,  $p < .001$ ,  $d = 0.58$ , 95%CI: [.32, .82]. They also had less positive feelings about the larger set's price ( $M = 4.01$ ,  $SD = 1.44$ ) than about the smaller set's price ( $M = 4.79$ ,  $SD = 1.05$ ), Welch's  $t(244.43) = 5.05$ ,  $p < .001$ ,  $d = 0.62$ , 95%CI: [.36, .87]. Thus, participants across conditions found the smaller set more attrac-

**Table 4. Study 4: Replication analyses**

	Study 4 (Dinnerware Set) Original Study (N = 104)				Study 4 (Dinnerware Set) Replication (N = 400)					
	Dish Set A Mean	Dish Set B Mean	<i>t</i>	<i>p</i>	Dish Set A Mean	Dish Set B Mean	<i>t</i>	<i>p</i>	<i>df</i>	<i>d</i> or <i>d<sub>z</sub></i> [95%CI]
Separate Evaluation	\$23.25	\$32.69	3.91	<.001	\$23.96	\$33.70	6.23	<.001	257	0.76 [.50, 1.02]
Joint Evaluation	\$32.03	\$29.70	2.15	.039	\$31.67	\$30.35	1.26	.210	133	0.11 [-.06, .28]

tive and had more positive feelings toward its price, echoing the pattern of willingness to pay.

#### Joint evaluations

Against our predictions, participants in the joint evaluation condition rated the larger set ( $M = 3.57$ ,  $SD = 1.43$ ) as less attractive than the smaller set ( $M = 4.19$ ,  $SD = 1.15$ ),  $t(133) = 4.44$ ,  $p < .001$ ,  $d_z = 0.38$ , 95%CI: [.21, .56]. They also had less positive feelings about the larger set's price ( $M = 3.66$ ,  $SD = 1.40$ ) than about the smaller set's price ( $M = 4.22$ ,  $SD = 1.21$ ),  $t(133) = 4.59$ ,  $p < .001$ ,  $d_z = 0.40$ , 95%CI: [.22, .57]. Thus, while they were willing to pay the same amount, or possibly more, for the larger set, they were less sanguine about other aspects of the set.

#### Exploratory exclusion criteria

Hsee did not mention exclusion criteria for Study 4, so we did not initially use any. However, it would not be unreasonable to apply the same criteria as used in Study 2: exclude anyone willing to pay more than 3SD above the mean. Doing this would exclude two participants in the joint condition and zero participants in the separate condition. However, excluding these participants did not substantively change any of the analyses. All significant analyses remain significant, and non-significant analyses remain non-significant.

## Discussion

The less is better effect for separate evaluations successfully replicated. However, the more is better effect for joint judgments did not—although in the predicted direction, the difference was not significant.

An extension found that people's judgments of the attractiveness of the dishes and people's feelings toward their prices corresponded closely to their willingness to pay for them in the separate evaluations condition. However, in the joint evaluations condition, though people were willing to pay similar or more for the bigger set of dishes, they felt more negative toward their prices and thought they were less attractive. Perhaps people dislike the unequal number of unbroken cups and saucers in the large set condition, though the present research did not test this.

### General Discussion

Overall, we found support for four of the five effects in Hsee (1998), with no support for the fifth effect, which was

in the right direction (Table 5). Our extensions resulted in insightful findings. We conclude robust evidence for the less is better effect, with effect size varying when compared to the original, depending on the vignette used. However, there was weaker evidence for the more is better effect, with only Study 2 showing significant support for it. This means there was weaker than anticipated evidence for the evaluability hypothesis, despite successfully replicating its most surprising findings. Below, we discuss each effect, as well as some open questions about theory and measurement relating to economic rationality and the evaluability hypothesis.

#### Less is better effect

Three variations of the less is better effect (Hsee, 1998) replicated successfully: When making separate evaluations, people thought \$45 scarves were more generous gifts than \$55 coats, smaller overflowing ice creams were worth more than larger underfilled ice creams, and smaller dish-sets were more valuable than larger ones with some broken pieces. Thus, the less is better effect, as tested in these paradigms, is on very strong ground as a reliable effect.

#### More is better effect

Two close replications tested the more is better effect, yielding less consistent results. Although the means were in the predicted direction in both replications, and were significant in Study 2, the results were weaker and non-significant in Study 4. People were willing to pay more for more ice cream, but not willing to pay significantly more for a larger set of dishes including broken ones. An extension on the original Study 4 suggests one reason people may have been unwilling to pay more for the dishes is they felt more negatively toward their prices and thought they were likely less attractive. Some dishes being broken likely signaled a low quality of dishes. Of course, one failed replication of the more is better effect in Study 4 does not necessarily mean the paradigm is invalid, but it does suggest that the evidence is possibly weaker than previously thought.

#### Open questions about the evaluability hypothesis

Following these replication studies, there are some open questions about the evaluability hypothesis. First, in our review process we realized that there is some theoretical ambiguity about the meaning of the less is better effect. In



**Table 5. Comparison of original and replication effects**

Effect	Study	Original/ManyLabs <i>d</i> [CIL, CIH]	Replication <i>d</i> or <i>d<sub>z</sub></i> [CIL, CIH]	Replication Status
<i>Less is better</i>				
	Study 1	0.70 [0.24, 1.15]	0.99 [0.72, 1.25]	Signal, larger effect
	ManyLabs	0.78 [0.74, 0.83]		
	Study 2	0.74 [0.12, 1.35]	0.32 [0.07, 0.56]	Signal, smaller effect
	Study 4	0.97 [0.43, 1.50]	0.76 [0.50, 1.02]	Signal, smaller effect
<i>More is better</i>				
	Study 2	0.92 [0.42, 1.40]	0.60 [.42, .79]	Signal, smaller effect
	Study 4	0.37 [0.02, 0.72]	0.11 [-0.06, 0.28]	No signal, smaller effect

our extensions, we had included measures of preference, attractiveness, and positive versus negative feelings toward the items in these studies, thinking these would be appropriate subsidiary tests of the main hypothesis—i.e., that less would be perceived as better. A reviewer pointed out that in Hsee (1998), “better” was meant to be about value, and was only operationalized as perceived generosity (Study 1) and willingness to pay (Studies 2 and 4), so using other variables such as preference would not be an appropriate test of the less is better effect. They were correct in arguing that perceived generosity and willingness to pay may deviate from other seemingly similar measures. For example, one might judge that an expensive but hideous coat is a generous gift, but not a gift one would prefer to receive. Indeed, our results suggest that willingness to pay and perceived generosity relate quite inconsistently with preference, attractiveness, and feelings toward items, suggesting there are likely important theoretical distinctions between these measures in these “less is better” paradigms.

This suggests the need for conceptual clarity in the literature on the less is better effect. While Hsee (1998) used perceived generosity and willingness to pay to operationalize value, other terms including “judged more favorably,” “is better than,” “appealing,” and “preference” were used to describe the effect—which are conceptually very similar to “preference,” “attractiveness,” and “positive feelings about.” Moreover, follow-up research has continued to use the concept broadly—applying the less is better effect to describe diverse findings including preferences for portion size on plates of varying sizes (Parrish & Beran, 2014). Identifying the boundary conditions on the effect seems important given the inconsistency with which people responded in these studies to different measures in different paradigms—and developing a detailed conceptual framework to understand when less is more and when it isn’t would also be useful. Thus, an open question remains: Under what circumstances can researchers expect preferences to reverse?

Second, the meaning of the less is better effect depends on the mechanism behind it. The effect is meant to be part of a wider literature critiquing the idea that economic valuations are based on stable, consistent preferences (e.g., Bazerman et al., 1999; Kahneman, 2003; Mercier & Sperber, 2011; Shafir & LeBoeuf, 2004; Slovic et al., 2007). The less is better effect has been interpreted as evidence of unsta-

ble, inconsistent preferences across contexts—that people’s judgments of and preferences for two different products reverse depending on how they evaluate them—separately and independently judging their qualities, or jointly comparing them (Hsee, 1998; Hsee et al., 1999). Indeed, Study 2 provides solid, successfully replicated evidence of this.

However, we are now less confident that the paradigms in Studies 1 and 4 provide evidence of these preference reversals, given the less compelling results for the more is better effect. An alternative explanation for these paradigms is possible, especially in Study 1: Hsee (1998) may have discovered some situations in which people think less really is better, regardless of whether those judgments are made jointly or separately. For example, in Study 1, perhaps the price-quality heuristic led people to think inexpensive meant low-quality (e.g., Gerstner, 1985). Thus, there remain open questions about the paradigms in Studies 1 and 4—although they do show a less is better effect, is that evidence of preference reversals?

### Limitations and future directions

There are some limitations. Being a direct replication, the results speak to the replicability of these paradigms, but are less informative about the generalizability of the effect. Future conceptual replications could help establish generalizability. Also, although the current mTurk sample is more diverse than the original student sample, the participants are WEIRD, and future research may aim to replicate this effect on a more culturally diverse sample.

Given the within-subjects design, it is conceivable that participants could have guessed the purpose of the study in ways that could have influenced their responses to subsequent decisions—but this is unlikely because participants were placed in the same conditions across studies. Moreover, because study order was counterbalanced, it is unlikely this would undermine the conclusions drawn. It is also possible that participants’ responses to the conceptually related conditions they saw previously may have influenced their responses later on. We conducted a comparison of the means of the full sample against the means of only the first study displayed and found a very similar pattern, suggesting that order did not seem to substantially impact the results. Nonetheless, future studies could address this

limitation by either running each study on separate participants or by including filler tasks between studies.

Scholars who wish to build on the less is better effect should be careful in choosing paradigms for their own research—crucially, researchers should include joint valuation conditions to ensure the less is better effect happens only when people evaluate options separately. Study 2's paradigm seems most likely to yield strong results consistently. Additional direct replications of conceptually related effects would be useful to increase confidence in the generalizability of the less is better effect.

---

### Competing Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

### Funding

The author(s) received no financial support for the research and/or authorship of this article.

### Authorship Declaration

Andy verified and reanalyzed the data, rewrote the manuscript, and compiled all materials. Gilad led the replication effort, supervised each step in the project, conducted the pre-registrations, and ran data collection. Gilad and Andy jointly finalized manuscript for submission.

Stephanie Chan, Wing Yiu Hung, Wai Yee Michelle Leung, and Anna Thao Bich Nguyen wrote the pre-registration, conducted the replication and data analysis, and wrote an initial report. Bo Ley Cheng guided and assisted the replication effort.

### Acknowledgements

We thank Yiyu Chen, Mannix Chan, Shanshan Peng, and Yaqi Jin for their help in serving as our external red-team providing an in-depth assessment and evaluation of our work with important error detection and recommendations on how to improve.

Submitted: January 04, 2023 PDT, Accepted: May 25, 2023 PDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## References

- Bazerman, M. H., Moore, D. A., Tenbrunsel, A. E., Wade-Benzoni, K. A., & Blount, S. (1999). Explaining how preferences change across joint versus separate evaluation. *Journal of Economic Behavior & Organization*, 39(1), 41–58. [https://doi.org/10.1016/s0167-2681\(99\)00025-6](https://doi.org/10.1016/s0167-2681(99)00025-6)
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2016). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality data? In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 133–139). American Psychological Association. <https://doi.org/10.1037/14805-009>
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433–1436. <https://doi.org/10.1126/science.aaf0918>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., ... Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, 2(9), 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Davidai, S., & Shafir, E. (2018). Are ‘nudges’ getting a fair shot? Joint versus separate evaluation. *Behavioural Public Policy*, 1–19.
- Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Sansone, C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology: Recommendations for research and educational practice. *Personality and Social Psychology Review*, 18(1), 3–12. <https://doi.org/10.1177/1088868313507536>
- Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time*. Department of Statistics, Columbia University.
- Gerstner, E. (1985). Do higher prices signal higher quality? *Journal of Marketing Research*, 22(2), 209–215. <https://doi.org/10.1177/002224378502200210>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466(7302), 29–29. <https://doi.org/10.1038/466029a>
- Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3), 247–257. <https://doi.org/10.1006/obd.1996.0077>
- Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, 11(2), 107–121. [https://doi.org/10.1002/\(sici\)1099-0771\(199806\)11:2](https://doi.org/10.1002/(sici)1099-0771(199806)11:2)
- Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, 125(5), 576–590. <https://doi.org/10.1037/0033-2909.125.5.576>
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720. <https://doi.org/10.1037/0003-066x.58.9.697>
- Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahnik, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., ... Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 1(4), 443–490. <https://doi.org/10.1177/2515245918810225>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 2(49), 433–442.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(2), 57–74. <https://doi.org/10.1017/s0140525x10000968>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 1–9. <https://doi.org/10.1038/s41562-016-0021>
- Nosek, B. A., Aarts, A. A., Anderson, J. E., Kappes, H. B., & Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- Nosek, B. A., & Lakens, D. (2014). A method to increase the credibility of published results. *Social Psychology*, 45(3), 137–141. <https://doi.org/10.1027/1864-9335/a000192>
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5), 411–419. <https://doi.org/10.1017/s1930297500002205>

- Parrish, A. E., & Beran, M. J. (2014). When less is more: Like humans, chimpanzees (*Pan troglodytes*) misperceive food amounts based on plate size. *Animal Cognition*, *17*(2), 427–434. <https://doi.org/10.1007/s10071-013-0674-3>
- Pattison, K. F., & Zentall, T. R. (2014). Suboptimal choice by dogs: When less is better than more. *Animal Cognition*, *17*(4), 1019–1022. <https://doi.org/10.1007/s10071-014-0735-2>
- Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, *17*(4), 551–566. <https://doi.org/10.1037/a0029487>
- Shafir, E., & LeBoeuf, R. A. (2004). Context and conflict in multiattribute choice. In *Blackwell Handbook of Judgment and Decision Making* (pp. 339–359). <https://doi.org/10.1002/9780470752937.ch17>
- Sher, S., & McKenzie, C. R. (2014). Options as information: Rational reversals of evaluation and preference. *Journal of Experimental Psychology: General*, *143*(3), 1127–1143.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. <https://doi.org/10.1016/j.ejor.2005.04.006>
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, *13*(4), 411–417. <https://doi.org/10.1177/1745691617751884>
- Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*, 120. <https://doi.org/10.1017/s0140525x17001972>

## Supplementary Materials

### Peer Review History

Download: [https://collabra.scholasticahq.com/article/77859-less-is-better-in-separate-evaluations-versus-more-is-better-in-joint-evaluations-mostly-successful-close-replication-and-extension-of-hsee-19/attachment/163322.docx?auth\\_token=S2zm0XtfA67dONvMgPzA](https://collabra.scholasticahq.com/article/77859-less-is-better-in-separate-evaluations-versus-more-is-better-in-joint-evaluations-mostly-successful-close-replication-and-extension-of-hsee-19/attachment/163322.docx?auth_token=S2zm0XtfA67dONvMgPzA)

---

### Supplementary Materials

Download: [https://collabra.scholasticahq.com/article/77859-less-is-better-in-separate-evaluations-versus-more-is-better-in-joint-evaluations-mostly-successful-close-replication-and-extension-of-hsee-19/attachment/163323.docx?auth\\_token=S2zm0XtfA67dONvMgPzA](https://collabra.scholasticahq.com/article/77859-less-is-better-in-separate-evaluations-versus-more-is-better-in-joint-evaluations-mostly-successful-close-replication-and-extension-of-hsee-19/attachment/163323.docx?auth_token=S2zm0XtfA67dONvMgPzA)

---