Available online at www.sciencedirect.com

ScienceDirect

Journal of Consumer Psychology 27, 2 (2017) 278–286

JCP JOURNAL OF CONSUMER PSYCHOLOGY

Research Report

The role of evaluation mode on the unit effect☆

Dan R. Schley [a,*], Christophe Lembregts [a], Ellen Peters [b]

[a] Department of Marketing, Rotterdam School of Management, Erasmus University, The Netherlands
[b] Department of Psychology, The Ohio State University, United States

Accepted by Cornelia Pechmann and Anirban Mukhopadhyay, Editors; Associate Editor, Steve Posavac

Received 20 February 2014; received in revised form 14 June 2016; accepted 1 July 2016
Available online 15 July 2016

Abstract

Recent research on the unit effect has suggested that consumers tend to ignore relevant unit information and over-rely on numeric magnitudes in judgments (e.g., perceiving the difference between 700 and 900 on a 1000-point quality scale to be larger than the difference between 7 and 9 on a 10-point scale). The current work investigates the nature of the unit effect by studying the role of different modes of evaluation, and types of information processing, on the unit effect. Specifically, three studies demonstrate that the unit effect occurs when options are evaluated simultaneously and attenuated when options are evaluated sequentially. The current article builds on research concerning comparative versus selective information processing. It demonstrates that, when information is processed in a comparative rather than selective manner, common elements in the decision (i.e., units) are more likely to be edited out, resulting in the unit effect.

Keywords: Unit effect; Judgments; Evaluation mode

Normatively speaking, preferences for a given good should not be influenced by the units used to convey the magnitude of the good. That is, a product with a quality score of 7 out of 10 should be perceived as equally attractive as one with a score of 700 out of 1000. In contrast to these normative assumptions, recent research has demonstrated that consumers' judgments are highly influenced by the unit of reference. For example, Pandelaere, Briers, and Lembregts (2011) demonstrated that participants perceived quality differences as larger for two televisions with quality rating scores presented on a 1000-point scale than for two televisions with quality rating scores presented on a 10-point scale. Thus, participants appeared to neglect the reference scale (i.e., unit), relying instead on the magnitude of the presented numeric information (i.e. the unit effect). This and related numerosity effects have been demonstrated in a variety of domains such as currencies (Raghubir & Srivastava, 2002; Wertenbroch, Soman, & Chattopadhyay, 2007), loyalty programs (Bagchi & Li, 2011), and product attributes (Burson, Larrick, & Lynch, 2009; Monga & Bagchi, 2012; Pandelaere et al., 2011). The current research aims to understand factors that produce and attenuate the unit effect.

Consumers should attend to both the numeric differences (e.g., ratings of 7 and 9 vs. 700 and 900) and the unit information (e.g., whether the ratings are on a 10-point scale or 100-point scale) when assessing quantities. However, unit-effect research has demonstrated that consumers rely primarily on the abstract numeric component, often ignoring the relevant unit or scale information (Monga & Bagchi, 2012; Pandelaere et al., 2011; Shen & Urminsky, 2013). For example, Pandelaere et al. (2011) found that consumers evaluated the magnitude of a difference between two dishwasher warranties as subjectively larger when specified in months (84 vs. 108 months) than in years (7 vs. 9 years), which suggests a focus on the size of the numbers and a relative lack of consideration of the contextualizing unit.

☆ Research was supported by NSF Graduate Research Fellowship (Schley), Flemish Science Foundation PhD Fellowship (Lembregts), and NSF SES-1047757 and SES-1155924 (Peters).
* Corresponding author at: Office T10-30 PO Box 1738 3000DR, Rotterdam, The Netherlands.
E-mail address: schley@rsm.nl (D.R. Schley).

http://dx.doi.org/10.1016/j.jcps.2016.07.001
1057-7408/

Consumers only appear to shift attention to the unit under certain circumstances. For example, Shen and Urminsky (2013) demonstrated that increasing the visual salience of the unit increased attention to the unit and decreased the unit effect. Similarly, Monga and Bagchi (2012) found that experimentally evoking an abstract mindset shifted focus to the unit, moderating the unit effect. Although most researchers appear to agree that the unit effect is likely produced by consumers relying on the numbers more than the unit information, the exact mechanisms remain unclear (Adaval, 2013).

Unit (and related) effects have been observed under conditions both of joint evaluation (e.g., Burson et al., 2009; Monga & Bagchi, 2012; Pandelaere et al., 2011) and separate evaluation (e.g., Shen & Urminsky, 2013; Raghubir & Srivastava, 2002; Wong & Kwong, 2000).[1] Because consumers often evaluate more than one alternative, our focus in this article is on joint evaluation. In particular, we turn our attention to a distinction within joint evaluation modes that has not been systematically studied within the unit-effect literature: sequential versus simultaneous evaluation. For instance, when shopping online, a consumer may first view a dishwasher with a 5-year warranty and subsequently click through to another dishwasher with a 7-year warranty (sequential evaluation). Alternatively, many websites offer consumers the ability to compare multiple items simultaneously on a single page.

Simultaneous versus sequential evaluation

Research on simultaneous versus sequential evaluation has primarily investigated the influence of evaluation mode on information search (Diehl & Zauberman, 2005; Gilbert & Mosteller, 1966; Levav, Reinholtz, & Lin, 2012; Shu, 2008; Weitzman, 1979) and order effects (Bruine de Bruin & Keren, 2003; Hogarth & Einhorn, 1992). In particular, these researchers have demonstrated that simultaneous versus sequential evaluation affects attention to and weighting of different attributes in multi-attribute choices; they have left untested whether the difference in evaluation will influence how consumers process the quantity information contained within a particular attribute.

In general, simultaneous and sequential evaluation modes differ in the ease of direct comparisons. As such, it may affect the processing style that consumers adopt in judgments (Kardes, 2013; Mantel & Kardes, 1999). When information is evaluated simultaneously, comparative processing is facilitated by the explicit juxtaposition of the options (Kardes, 2013). This may have important implications for how quantitative information is processed, in particular with regard to the neglect of the contextualizing unit information (Pandelaere et al., 2011; Shen & Urminsky, 2013). Past research has found that when consumers directly compare two options that are identical on a given attribute (e.g., option A: x% chance of winning $10 vs. option B: y% chance of winning $10), they edit out common features (i.e., $10) and make decisions based solely on dissimilar features (i.e., x% vs. y%) (Kahneman & Tversky, 1979; Payne, Bettman, & Johnson, 1992; Thaler & Johnson, 1990). We hypothesize that, because the unit information is identical between options, it may be edited out as a common element.

If common elements in the decision environment can be edited out, then what factors facilitate this editing? Wang and Wyer (2002) found that comparative information processing inhibited the processing of common features in multi-attribute choice. Specifically, they found that participants who were explicitly told to compare two products were more likely to cancel out common features relative to participants who were not instructed to compare the products. If comparative processing leads to editing-out of common features (Wang & Wyer, 2002) and if neglecting unit information produces the unit effect (Pandelaere et al., 2011; Shen & Urminsky, 2013), then simultaneous evaluation, which facilitates comparative processing (Kardes, 2013), should promote the unit effect.

Whereas simultaneous evaluation facilitates comparative processing, sequential evaluation typically lends itself to more step-by-step processing in which global impressions of each alternative are compared in a stepwise fashion (i.e. selective processing; Bruine de Bruin & Keren, 2003). Particularly, the first evaluation is made without explicit comparisons, and consumers are forced to selectively process information in isolation on the basis of its own merits without considering other information (Sanbonmatsu, Posavac, Kardes, & Mantel, 1998). Subsequently, additional judgments in sequential evaluation can involve comparison of the attribute(s) between options. Consumers, however, must rely on memory for any comparisons, making direct comparison more difficult in sequential evaluation; this should attenuate the unit effect. Sequential evaluation is not without comparative processing; rather, it facilitates relatively less comparative processing than simultaneous evaluation.

We test the role of evaluation mode and information processing on the unit effect in three studies. Study 1 investigates whether sequential evaluation attenuates the unit effect relative to simultaneous evaluation. Study 2 investigates what information is processed in both modes using a memory measure. Lastly, Study 3 attempts to test the editing hypothesis by manipulating the mode of information processing, through a comparative evaluation prompt (Wang & Wyer, 2002), and observing its influence on the unit effect.

Study 1

The aim of Study 1 was to test whether evaluation mode moderates the unit effect. To do so, we modified a paradigm used in Study 1 of Pandelaere et al. (2011).

Methods

Online Mechanical Turk participants (n = 204; 41.7% female; M_{age} = 27.7; range 18–75 years) were recruited and paid $0.50 for participation. Participants were asked to imagine that they were shopping for a new television. They were told about two

[1] We should note that, whereas research on sequential versus simultaneous evaluation (both modes of joint evaluation) has not investigated how modes of evaluation influence the processing of quantitative information, extensive research exists related to this topic in joint versus separate evaluation (for a review, see Hsee & Zhang, 2010).

televisions identical in every way except for their quality scores. To manipulate the numeric magnitude of the numbers, the two televisions were presented with quality scores of 7 and 9 on a 10-point scale (i.e., the small-magnitude condition) or 700 and 900 on a 1000-point scale (i.e., the large-magnitude condition). In the simultaneous-evaluation condition, participants were presented the two televisions simultaneously and were asked to indicate how much they would be willing to pay (WTP) for each television.

In the sequential-evaluation condition, participants were presented the two televisions on separate screens presented one immediately after the other. To control for potential order effects in sequential evaluation, we randomized the presentation order of the two televisions. Participants were told that they would only see two televisions. For instance in the small-magnitude condition, a participant would be presented the television with the rating of "7 out of 10" and indicate their WTP. On the next screen, participants would indicate their WTP for the second television with a rating of "9 out of 10".

Results

Because participants differ in their overall WTP for televisions, we analyzed the data using ratios between the two responses. For instance, WTP of $400 and $600 would be coded as 1.5, indicating that the participant would be willing to pay 50% more for the television with the higher rating. Results were comparable using other transformations (e.g., log-transforming raw WTP). In addition, 3% of WTP-ratios were less than 1 (i.e., WTP was lower for the higher rated television) and 1% were greater than 3; these outliers were truncated at 1 and 3 to constrain the large variance produced by outlier responses. In Appendix A, we reanalyze the data of all three studies, first without truncation of outlier responses, and second with outlier responses deleted rather than truncated; results are largely consistent with those below. Descriptive results are presented in Fig. 1.

Participants' WTP-ratios were submitted to a two-way ANOVA, with unit (10-point vs. 1000-point scales) and evaluation mode (simultaneous vs. sequential) treated as between-participant variables. The unit effect is assessed based on sensitivity to the differences between the two televisions (i.e., greater WTP-ratios) between unit presentations. Results indicated no main effect of unit or evaluation mode, $ps > .17$. As expected, a significant interaction was observed, $F(1, 200) = 4.66$, $p = .03$. Consistent with past unit-effect findings, in the simultaneous-evaluation condition we observed the unit effect, such that participants were more sensitive to the difference between televisions with ratings of 700 and 900 out of 1000 ($M_{\text{WTP-ratio}} = 1.75$, SD = 0.35) than to televisions with ratings of 7 and 9 out of 10 ($M_{\text{WTP-ratio}} = 1.53$, SD = 0.48, $F(1, 100) = 7.20$, $p = .01$). Conversely, in the sequential-evaluation condition, we observed no difference in evaluations of televisions with ratings of 700 and 900 out of 1000 ($M_{\text{WTP-ratio}} = 1.57$, SD = 0.45) and televisions with ratings of 7 and 9 out of 10 ($M_{\text{WTP-ratio}} = 1.61$, SD = 0.51, $F(1, 100) = 0.27$, $p = .61$). Thus, Study 1 provided the first known evidence that the unit effect is observed in simultaneous evaluation but attenuated when options are evaluated sequentially. In the next two studies, we attempt to replicate this finding and examine potential underlying mechanisms.

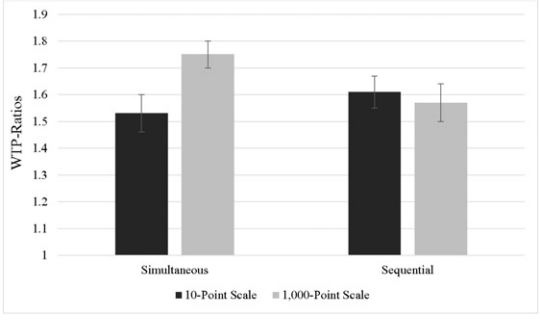

Fig. 1. WTP-ratios by condition in Study 1. Error bars reflect standard error of the means. The unit effect is evidenced by larger gray bars than black bars.

Study 2

In Study 2, we used a wider range of stimuli to test generalization of the effect. In addition, we investigated participants' memory for the stimuli to assess whether evaluation mode affects the information processed. If simultaneous evaluation fosters more comparative information processing and editing-out of common features (relative to sequential evaluation), then participants should recall less information about the unit in the simultaneous-evaluation condition.

Methods

Online Mechanical Turk participants ($n = 201$; 43.5% female; $M_{\text{age}} = 33.7$; range 18–72 years) were recruited and paid $0.50 for participation. Participants indicated WTP for four pairs of products, with the pairs presented in random order. Participants were presented dishwashers with warranties in years (3 vs. 5) or months (36 vs. 60), televisions with screens in inches (47 vs. 55) or centimeters (119 vs. 140), ground beef in pounds (1 vs. 2) or ounces (16 vs. 32), and kitchen knives with customer ratings on 10-point (7 vs. 9) or 1000-point scales (700 vs. 900). In addition, participants evaluated the pairs either simultaneously or sequentially. Unit and evaluation mode were manipulated between participant and product type was within participant.

To assess what information participants processed when making their judgments, at the end of the task and after a short filler task (see Methodological Details Appendix), participants were asked to freely recall any and all information about the products they saw. Free recall was chosen to avoid potential demand effects that may occur when using cued-recall or recognition-memory tasks. Two independent coders coded for mention of numbers, units, and product type. For instance, for the participant who recalled "Dishwasher warranty 32 or 60 months/television centimeters 119 or 140," the coder indicated that the participant recalled 4 numbers, 2 units, and 2 items. Because we were interested in the types of information processed, we coded according to categories even if the specific information was incorrect (e.g., 32 was coded as a

number even though the correct number was 36). Inter-rater reliability was reasonable ($\alpha_{Krippendorff} = 0.891$), and differences were handled by averaging the two ratings.

Results

WTP

Responses were handled in the same manner as Study 1. Descriptive results are presented in Table 1. Participants' WTP-ratios were predicted by unit, evaluation mode, and their interaction in a mixed-model, allowing for random effects for the different products. Results indicated no main effect of unit, $F(1, 603) = 1.50$, $p = .22$, a significant main effect of evaluation mode, $F(1, 603) = 5.84$, $p = .02$, and a significant interaction, $F(1, 603) = 5.01$, $p = .03$. Results largely replicate those of Study 1. In the simultaneous-evaluation condition, participants demonstrated a significant unit effect, exhibiting greater sensitivity to the differences between, for instance, the 36- and 60-month dishwasher warranties ($M_{WTP\text{-}ratio} = 1.30$, SD = 0.24) than the 3- and 5-year warranties ($M_{WTP\text{-}ratio} = 1.20$, SD = 0.23) (contrast test using mixed model, $F(1, 291) = 4.97$, $p = .03$). Consistent with Study 1, results again indicated no significant unit effect when products were evaluated sequentially, $F(1, 312) = 0.54$, $p = .46$.

Memory

Results of the free-recall coding are presented by condition in Table 2. Participants recalled an average of 2.29 numbers (out of 8 possible), 1.33 units (out of 4), and 2.56 products (out of 4). For each type of memory, we predicted the frequency of recall with evaluation mode, the unit condition, and their interaction. As hypothesized, participants recalled significantly less information about units when options were evaluated simultaneously ($M = 1.13$, SD = 1.27) than when evaluated sequentially ($M = 1.51$, SD = 1.35; $F(1, 197) = 4.33$, $p = .04$); no interaction existed between evaluation mode and unit, $p = .96$. There was also a significant effect of evaluation mode on the frequency of numbers recalled, $F(1, 197) = 3.88$, $p = .05$. Evaluation mode did not influence participants' memory for the type of products, $p = .81$, nor did it influence the likelihood of not recalling anything, $p = .35$, suggesting that evaluation mode only influenced recall of the unit and numeric information.

Discussion

Results of Study 2 replicated the attenuation of the unit effect in sequential evaluation and demonstrated that the effect (and its effect size) generalized across multiple product types.

Table 2
Frequency of items coded to each category and standard errors from the free-recall memory measure in Study 2.

	Simultaneous evaluation	Sequential evaluation
Units (out of 4 possible)	1.13	1.51
	(0.13)	(0.13)
Numbers (out of 8 possible)	1.90	2.65
	(0.28)	(0.28)
Product type (out of 4 possible)	2.58	2.54
	(0.11)	(0.11)
Recall nothing (0 or 1)	0.09	0.06
	(0.28)	(0.23)

We should note that, although the simultaneous-evaluation condition resulted in less recall of units, it also resulted in less recall of numbers. Participants tended not to mention numbers without specifying units (e.g., participants didn't state "36 and 60 somethings"), resulting in a correlation of .87 between the recall of numbers and units. We prefer not to interpret this result as anything other than an artifact of participants' response tendencies. One possibility is that sequential evaluation may simply increase elaboration of all information. Although this appears not to be the case in Study 2 (i.e., no influence of evaluation mode on recall of item type), we acknowledge this as a limitation and moderate our conclusions accordingly.

Study 3

Study 2 provided initial evidence that simultaneous evaluation inhibits the processing of the common feature (i.e., the unit information). However, the manipulation of evaluation mode may change a host of informational processes in addition to comparative processing. To investigate the role of comparative processing independent of other potential processes, Study 3 manipulates comparative processing within an evaluation mode. Specifically, we reason that when consumers in sequential evaluation are prompted to make evaluations comparatively (Kardes, 2013; Wang & Wyer, 2002), the unit effect should reemerge.

Methods

Online MTurk participants ($n = 404$; 39.9% female; $M_{age} = 31.4$; range 18–64 years) were paid $0.50 for participation. To assess the unit effect, participants were presented dishwashers with warranties in years (3 vs. 5) or months (36 vs. 60). In addition, participants were randomly assigned to one of four evaluation-mode conditions. As in previous studies, participants were assigned to evaluate the dishwashers simultaneously or

Table 1
WTP-ratios and standard errors for each product presented by condition in Study 2.

Evaluation mode	Dishwasher warranties		Television screen sizes		Ground beef		Customer review scores for knives		Average	
	Years	Months	Inches	Centimeter	Pounds	Ounces	10-point scale	1000-point scale	Big unit/small numbers	Small unit/ big numbers
Simultaneous	1.20 (0.03)	1.30 (0.04)	1.30 (0.02)	1.36 (0.04)	1.75 (0.03)	1.87 (0.03)	1.52 (0.07)	1.62 (0.07)	1.44 (0.03)	1.54 (0.03)
Sequential	1.19 (0.03)	1.23 (0.04)	1.33 (0.04)	1.24 (0.03)	1.75 (0.05)	1.74 (0.05)	1.49 (0.05)	1.45 (0.06)	1.44 (0.02)	1.41 (0.03)

sequentially. Although we have suggested that our manipulation of simultaneous versus sequential presentation influences the extent of comparative information processing, it is nevertheless possible that manipulating evaluation mode produces the apparent attenuation of the unit effect via other mechanisms. To mollify such concerns, in Study 3 we attempted to manipulate comparative information processing independent of the evaluation mode. In the sequential-comparative condition, we used prompts aimed at facilitating comparisons (Wang & Wyer, 2002). Before evaluating the first dishwasher, we added "We are interested in how people compare product attributes such as warranties". After evaluating the first dishwasher, participants were told to "compare this dishwasher to the one you saw on the previous screen. Consider the length of the warranty and how much you were willing to pay for that dishwasher when judging the following dishwasher." Because the comparative prompts may induce greater elaboration of all information, we created a fourth condition where participants were instructed to "take a moment to really think about this dishwasher and an appropriate price for it." Thus, Study 3 manipulated units (small vs. large) and evaluation mode (simultaneous vs. sequential-control vs. sequential-comparative vs. sequential-elaboration).

Results

Responses were handled in the same manner as the previous studies and displayed in Fig. 2. Participants' WTP-ratios were submitted to a two-way ANOVA, with unit (years vs. months), evaluation mode (simultaneous vs. sequential-control vs. sequential-comparative vs. sequential-elaboration), and their interaction as predictors. Results indicated a marginal main effect of unit, $F(1, 396) = 2.94$, $p = .09$, a significant main effect of evaluation mode, $F(3, 396) = 13.35$, $p < .0001$, and a significant interaction, $F(3, 396) = 3.45$, $p = .02$. Using pairwise analyses, we again observed a significant unit effect in the simultaneous condition where participants were more sensitive to the differences between the dishwashers with warranties in months ($M_{\text{WTP-ratio}} = 1.35$, SD = 0.16) than in years ($M_{\text{WTP-ratio}} = 1.25$, SD = 0.16, $F(1, 100) = 9.23$, $p = .003$. Consistent with previous studies, we also observed an attenuation of the unit effect when evaluated in the sequential-control condition, $p = .40$. More importantly, when we focus on the sequential-control and sequential-comparative conditions, we observe a significant condition × unit interaction, $F(1, 197) = 4.23$, $p = .04$. Consistent with our predictions, participants exhibited the unit effect in sequential evaluation when they were instructed to make direct comparisons across the focal attribute. Specifically, sequential-comparative participants were more sensitive to the differences between dishwashers when warranties were presented in months ($M_{\text{WTP-ratio}} = 1.26$, SD = 0.17) than in years ($M_{\text{WTP-ratio}} = 1.20$, SD = 0.14, $F(1, 98) = 4.13$, $p = .04$). Sequential-elaboration participants ($M_{Years} = 1.21$ vs. $M_{Months} = 1.19$), on the other hand, did not demonstrate a different-sized unit effect than sequential-control participants ($M_{Years} = 1.18$ vs. $M_{Months} = 1.15$, $F(1, 198) = 0.00$, $p = .95$), suggesting that prompting participants to make explicit comparisons did not simply make them elaborate more on the information. Sequential-comparative participants showed a marginally greater unit effect than sequential-elaboration participants, $F(1, 197) = 2.93$, $p = .09$. Although the unit effect was nominally larger in the simultaneous condition ($M_{Years} = 1.25$ vs. $M_{Months} = 1.35$) than in the sequential-comparative condition ($M_{Years} = 1.20$ vs. $M_{Months} = 1.26$), the difference was not significant $F(1, 198) = 0.56$, $p = .46$.

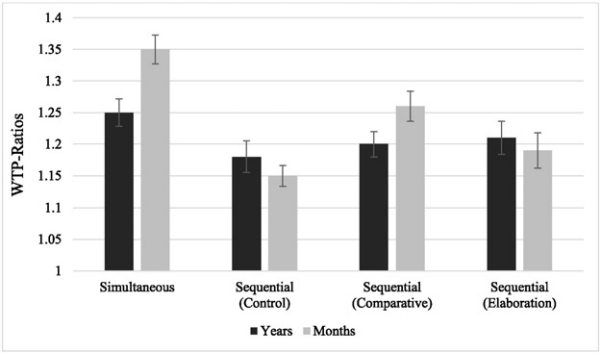

Fig. 2. WTP-ratios by condition in Study 3. Error bars reflect standard error of the means. The unit effect is evidenced by larger gray bars than black bars.

Discussion

Results of Study 3 provide additional evidence for the role of selective versus comparative information processing in the unit effect. In particular, relatively more selective processing in sequential evaluation leads to an attenuation of the unit effect. However, when options were evaluated sequentially, but participants were explicitly prompted to make explicit comparisons, the unit effect reemerged, suggesting an important causal role for comparative information processing.

General discussion

In three studies, we demonstrated that the unit effect occurs in simultaneous evaluation and attenuates in sequential evaluation. Study 2 further provided initial evidence of an editing process. In particular, participants were less likely to freely recall the unit information when options were evaluated simultaneously compared to sequentially. Study 2, however, did not directly assess the comparative process thought to drive the editing-out of the common unit feature. Therefore, Study 3 tested our proposed comparative-processing mechanism independent of evaluation mode. In particular, we observed the expected unit effect in the simultaneous-evaluation condition, the expected attenuation in the sequential-control-evaluation condition, and an emergence of the unit effect in the sequential-comparative-evaluation condition. Although the current studies employed only Mechanical Turk samples, earlier pretests with university students demonstrated largely consistent results.

Given evidence that comparative information processing is one factor related to the unit effect, the potential malleability of comparative information processing is consequential. Study 3 demonstrates that comparative prompts can facilitate the unit effect in sequential evaluation, presumably by increasing comparative processing. A possible alternative approach would be to present

stimuli sequentially and then have participants evaluate options on the same page. It may also be possible to alter the extent of comparative information processing within simultaneous evaluation. Past research has found that, when evaluating products with multiple attributes, horizontal presentation of the products facilitates more holistic product-based (and perhaps selective) information processing, whereas vertical presentation facilitates more attribute-based (and perhaps comparative) processing (Bettman & Kakkar, 1977; Bettman, Luce, & Payne, 1998).

In the current studies, we randomized the presentation order in sequential evaluation to study differences between simultaneous and sequential evaluation independent of order effects, but we did not intend to analyze order and, indeed, did not capture it in our experiments. However, because order can alter consumers' preferences (Li & Epley, 2009; Mantonakis, Rodero, Lesschaeve, & Hastie, 2009), it is possible that the unit effect interacts with serial position. In particular, because the first judgment in sequential evaluation is akin to separate evaluation, and unit effects influence judgments in separate evaluation (Raghubir & Srivastava, 2002; Shen & Urminsky, 2013), it is likely that the first judgment made in sequential evaluation depends on the unit. Consumers then may apply a judgment strategy (e.g., Bayesian inference; Biswas, Zhao, & Lehmann, 2011; McKenzie, 1994) to evaluate the second stimulus relative to their reference point (i.e., their first judgment). If the unit effect influences the first judgment, then it must influence the second judgment (via the influence of the first judgment). Alternatively, the unit may alter how consumers update from the first to the second judgment. Future research should consider how serial position and the unit effect interact.

Unit-effect research has focused on how two options described with smaller units (i.e., larger numbers) are judged as more different relative to the identical options described with larger units. If the unit effect reflects biased information processing in the construction of judgments, one might ask where the effects are largest as an indication of the possible origin of the effect. In particular, across three studies, the differences in the unit-effect size between evaluation modes was driven primarily by the stimuli with the smaller units (see Table 1). If sequential evaluation attenuates the unit effect, then the current studies provide some initial evidence that the unit effect might be exaggerated through the use of smaller units. Of course, in the present studies, it may simply be that the smaller units are less familiar and that less familiar judgments are more malleable to situational influences such as evaluation mode during the judgment/preference construction process (Lichtenstein & Slovic, 2006; Slovic, 1995). Another explanation might be that because smaller units involve larger numbers and larger numbers are more difficult to process than smaller numbers (Dehaene, 2011; Lembregts & Pandelaere, 2013; Schley & Peters, 2014), judgments of stimuli with smaller units (larger numbers) will be more labile. The hypothesis that the unit effect might be driven by smaller units, and potential mechanisms for such a finding, have yet to be properly tested, offering yet another relevant future research avenue.

The current findings provide a more nuanced understanding of some of the factors companies should consider when advertising quantitative product attributes. That is, in situations in which comparative processing is likely to occur (e.g., simultaneous evaluation), it may be better to specify attributes in smaller units when their product provides superior value on that attribute. Also, similar to the manipulation in study 3, marketers could instigate comparative processing by encouraging their consumers with slogans such as "compare our warranty to anyone's!"

Appendix A. Reanalysis of outlier data

Note to the reader: we used truncation to deal with outliers in the data. For example, if a participant indicated that they would pay $500 for a dishwasher with a 3-year warranty but $300 for the same dishwasher with a 5-year warranty, this implies that the participant gains negative utility from an additional benefit. As a consequence we truncated the WTP-ratio from the calculated 0.6 to 1.0 (i.e., valuing the two dishwashers equally). Similarly, we truncated surprisingly high ratios to 3 to avoid large outlier effects. For example, if a participant indicated they would pay $300 for a dishwasher with a 3-year warranty but $1500 for the same dishwasher with a 5-year warranty, this would result in a WTP-ratio of 5.0. So to deal with these potential outliers we truncated the WTP-ratios. The following results are those without any truncation, allowing for larger outlier effects. Subsequently, we present the results with outliers deleted instead of truncated.

A.1. Reanalyzing studies without data truncation specified in the text

Table S1
Reanalyzing Study 1 without data truncation specified in the text. Mean WTP-ratios with standard deviations in parentheses.

	10-point scale	1000-point scale
Simultaneous	1.53	1.75
	(0.48)	(0.35)
Sequential	1.62	1.58
	(0.60)	(0.59)

Results indicated no main effect of unit or evaluation mode, ps > .20. As expected, a marginal interaction was observed, $F(1, 200) = 3.36$, $p = .07$. The slightly higher p-value is to be expected since these outlier responses are not truncated and thus add to the sum of squared errors of the models. Consistent with past unit-effect findings, in the simultaneous-evaluation condition we observed the unit effect, such that participants were more sensitive to the difference between televisions with ratings of 700 and 900 out of 1000 than to televisions with ratings of 7 and 9 out of 10 ($F(1, 100) = 7.20$, $p = .01$). Conversely, in the sequential-evaluation condition, we observed no difference in evaluations of televisions with ratings of 700 and 900 out of 1000 and televisions with ratings of 7 and 9 out of 10 ($F(1, 100) = 0.11$, $p = .74$).

Table S2
Reanalyzing Study 2 without data truncation specified in the text. WTP-Ratios and standard errors for each product presented by condition in Study 2.

Evaluation mode	Dishwasher warranties		Television screen sizes		Ground beef		Customer review scores for knives		Average	
	Years	Months	Inches	Centimeter	Pounds	Ounces	10-point scale	1000-point scale	Big unit/ small numbers	Small unit/big numbers
Simultaneous	1.20	1.30	1.30	1.36	1.75	1.87	1.52	1.64	1.44	1.54
	(0.03)	(0.04)	(0.02)	(0.04)	(0.03)	(0.03)	(0.07)	(0.07)	(0.03)	(0.03)
Sequential	1.19	1.34	1.33	1.24	1.75	1.74	1.49	1.45	1.44	1.44
	(0.03)	(0.15)	(0.04)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.02)	(0.04)

Results indicated no main effect of unit or evaluation mode, ps > .09. Without truncation the interaction was not statistically reliable, $F(1, 603) = 2.01$, $p = .16$. The slightly higher p-value is to be expected since these outlier responses are not truncated and thus add to the sum of squared errors of the models. Consistent with past unit-effect findings, in the simultaneous-evaluation condition we observed the unit effect, such that participants were more sensitive to the difference between quantities in the large-magnitude condition than in the small-magnitude condition ($F(1, 291) = 5.04$, $p = .03$). Conversely, in the sequential-evaluation condition, we observed no difference between unit conditions ($F(1, 312) = 0.00$, $p = .99$).

Table S3
Reanalyzing Study 3 without data truncation specified in the text. Mean WTP-ratios with standard deviations in parentheses.

	Years	Months
Simultaneous	1.25	1.35
	(0.16)	(0.16)
Sequential	1.17	1.15
	(0.40)	(0.11)
Sequential comparative	1.20	1.27
	(0.14)	(0.22)
Sequential elaboration	1.32	1.20
	(0.98)	(0.33)

Results indicated no effect of unit, $F(1, 396) = 0.01$, $p = .92$, a nonsignificant main effect of evaluation mode, $F(3, 396) = 2.03$, $p = .10$, and a nonsignificant interaction, $F(3, 396) = 1.54$, $p = .20$. Using pairwise analyses, we observed a significant unit effect in the simultaneous condition where participants were more sensitive to the differences between the dishwashers with warranties in months than in years ($F(1, 100) = 9.23$, $p = .003$. Consistent with previous studies, we also observed an attenuation of the unit effect when evaluated in the sequential-control condition, $p = .63$. When we focus on the sequential-control and sequential-comparative conditions, we observe a nonsignificant condition × unit interaction, $F(1, 197) = 2.17$, $p = .14$. Again, the higher p-value is to be expected since these outlier responses are not truncated and thus add to the sum of squared errors of the models.

We also reanalyzed the studies with the outliers deleted instead of truncated and present these analyses below:

A.2. Reanalyzing studies with outliers deleted

Table S4
Reanalyzing Study 1 with outliers deleted. Mean WTP-ratios with standard deviations in parentheses.

	10-point scale	1000-point scale
Simultaneous	1.53	1.75
	(0.48)	(0.35)
Sequential	1.62	1.57
	(0.47)	(0.39)

Results indicated no main effect of unit or evaluation mode, *p*s > .16. As expected, a significant interaction was observed, $F(1, 192) = 5.21$, $p = .02$. Consistent with past unit-effect findings, in the simultaneous-evaluation condition we observed the unit effect, such that participants were more sensitive to the difference between televisions with ratings of 700 and 900 out of 1000 than to televisions with ratings of 7 and 9 out of 10 ($F(1, 100) = 7.20$, $p = .01$). Conversely, in the sequential-evaluation condition, we observed no difference in evaluations of televisions with ratings of 700 and 900 out of 1000 and televisions with ratings of 7 and 9 out of 10 ($F(1, 92) = 0.37$, $p = .54$).

Table S5
Reanalyzing Study 2 with outliers deleted. WTP-Ratios and standard errors for each product presented by condition in Study 2.

Evaluation mode	Dishwasher warranties		Television screen sizes		Ground beef		Customer review scores for knives		Average	
	Years	Months	Inches	Centimeter	Pounds	Ounces	10-point scale	1000-point scale	Big unit/ small numbers	Small unit/ big numbers
Simultaneous	1.20	1.29	1.30	1.35	1.75	1.86	1.52	1.59	1.44	1.52
	(0.03)	(0.04)	(0.02)	(0.04)	(0.03)	(0.03)	(0.07)	(0.06)	(0.03)	(0.03)
Sequential	1.20	1.19	1.34	1.27	1.78	1.78	1.49	1.48	1.45	1.43
	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.05)	(0.06)	(0.06)	(0.02)	(0.03)

Results indicated no main effect of unit or evaluation mode, *p*s > .06. The interaction was statistically reliable, $F(1, 586) = 4.69$, $p = .03$. Consistent with past unit-effect findings, in the simultaneous-evaluation condition we observed the unit effect, such that participants were more sensitive to the difference between quantities in the large-magnitude condition than in the small-magnitude condition ($F(1, 290) = 4.52$, $p = .03$). Conversely, in the sequential-evaluation condition, we observed no difference between unit conditions ($F(1, 296) = 0.67$, $p = .42$).

Table S6
Reanalyzing Study 3 with outliers deleted. Mean WTP-ratios with standard deviations in parentheses.

	Years	Months
Simultaneous	1.25	1.35
	(0.16)	(0.16)
Sequential	1.17	1.15
	(0.13)	(0.11)
Sequential comparative	1.20	1.27
	(0.14)	(0.22)
Sequential elaboration	1.20	1.18
	(0.15)	(0.16)

Results indicated an effect of unit, $F(1, 388) = 4.62$, $p = .02$, a main effect of evaluation mode, $F(1, 388) = 16.37$, $p < .0001$, and a significant interaction, $F(3, 388) = 3.52$, $p = .02$. Using pairwise analyses, we observed a significant unit effect in the simultaneous condition where participants were more sensitive to the differences between the dishwashers with warranties in months than in years ($F(1, 100) = 9.23$, $p = .003$. Consistent with previous studies, we also observed an attenuation of the unit effect when evaluated in the sequential-control condition, $p = .55$. When we focus on the sequential-control and sequential-comparative conditions, we observe a significant condition × unit interaction, $F(1, 194) = 3.89$, $p = .05$.

Appendix B. Methodological Details Appendix

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jcps.2016.07.001.

References

Adaval, R. (2013). Numerosity and consumer behavior. *Journal of Consumer Research*, *39*(5), xi–xiv.

Bagchi, R., & Li, X. (2011). Illusionary progress in loyalty programs: Magnitudes, reward distances, and step-size ambiguity. *Journal of Consumer Research*, *37*(5), 888–901.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research*, 233–240.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187–217.

Biswas, D., Zhao, G., & Lehmann, D. R. (2011). The impact of sequential data on consumer confidence in relative judgments. *Journal of Consumer Research*, *37*(5), 874–887.

Bruine de Bruin, W. B., & Keren, G. (2003). Order effects in sequentially judged options due to the direction of comparison. *Organizational Behavior and Human Decision Processes*, *92*(1), 91–101.

Burson, K. A., Larrick, R. P., & Lynch, J. G. (2009). Six of one, half a dozen of the other: Expanding and contracting numerical dimensions produces preference reversals. *Psychological Science*, *20*, 1074–1078.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* Oxford University Press.

Diehl, K., & Zauberman, G. (2005). Searching ordered sets: Evaluations from sequences under search. *Journal of Consumer Research*, *31*(4), 824–832.

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the maximum of a sequence. *Journal of American Statistical Association*, *61*, 35–73.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*(1), 1–55.

Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, *5*(4), 343–355.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*, 363–391.

Kardes, F. R. (2013). Selective versus comparative processing. *Journal of Consumer Psychology*, *23*(1), 150–153.

Lembregts, C., & Pandelaere, M. (2013). Are all units created equal? The effect of default units on product evaluations. *Journal of Consumer Research*, *39*(6), 1275–1289.

Levav, J., Reinholtz, N., & Lin, C. (2012). The effect of ordering decisions by choice-set size on consumer search. *Journal of Consumer Research*, *39*(3), 585–599.

Li, Y. E., & Epley, N. (2009). When the best appears to be saved for last: Serial position effects on choice. *Journal of Behavioral Decision Making*, *22*(4), 378–389.

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference.* Cambridge University Press.

Mantel, S. P., & Kardes, F. R. (1999). The role of direction of comparison, attribute-based processing, and attitude-based processing in consumer preference. *Journal of Consumer Research*, *25*(4), 335–352.

Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in choice effects of serial position on preferences. *Psychological Science*, *20*(11), 1309–1312.

McKenzie, C. R. (1994). The accuracy of intuitive judgment strategies: Covariation assessment and Bayesian inference. *Cognitive Psychology*, *26*(3), 209–239.

Monga, A., & Bagchi, R. (2012). Years, months, and days versus 1, 12, and 365: the influence of units versus numbers. *Journal of Consumer Research*, *39*(1), 185–198.

Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research*, *38*, 308–322.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, *43*(1), 87–131.

Raghubir, P., & Srivastava, J. (2002). Effect of face value on product valuation in foreign currencies. *Journal of Consumer Research*, *29*, 335–347.

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin & Review*, *5*(2), 197–220.

Schley, D. R., & Peters, E. (2014). Assessing "economic value" symbolic-number mappings predict risky and riskless valuations. *Psychological Science*, *25*(3), 753–761.

Shen, L., & Urminsky, O. (2013). Making sense of nonsense: The visual salience of units determines sensitivity to magnitude. *Psychological Science*, *24*(3), 297–304.

Shu, S. B. (2008). Future-biased search: The quest for the ideal. *Journal of Behavioral Decision Making*, *21*(4), 352–377.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, *36*(6), 643–660.

Wang, J., & Wyer, R. S., Jr. (2002). Comparative judgment processes: The effects of task objectives and time delay on product evaluations. *Journal of Consumer Psychology*, *12*(4), 327–340.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, *47*(3), 641–654.

Wertenbroch, K., Soman, D., & Chattopadhyay, A. (2007). On the perceived value of money: The reference dependence of currency numerosity effects. *Journal of Consumer Research*, *34*(1), 1–10.

Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes*, *82*(2), 314–333.