Journal of Experimental Psychology: General
2014, Vol. 143, No. 2, 504–509

0096-3445/14/$12.00 DOI: 10.1037/a0033343

BRIEF REPORT

Set-Fit Effects in Choice

Ellen R. K. Evers, Yoel Inbar, and Marcel Zeelenberg
Tilburg University

In 4 experiments, we investigate how the "fit" of an item with a set of similar items affects choice. We find that people have a notion of a set that "fits" together—one where all items are the same, or all items differ, on salient attributes. One consequence of this notion is that in addition to preferences over the set's individual items, choice reflects set-fit. This leads to predictable shifts in preferences, sometimes even resulting in people choosing normatively inferior options over superior ones.

Keywords: sets, choice, gestalt, preferences

Supplemental materials: http://dx.doi.org/10.1037/a0033343.supp

Imagine choosing between two pens, one of which is clearly superior to the other. In all likelihood, you would choose the superior pen. Now imagine making the same choice, with one difference. This time, regardless of what you choose, you also get three extra pens similar to the inferior pen. Would you now choose differently? If you are like the authors, your intuition is "no." However, as studies reported here will show, this intuition is inaccurate. In fact, people's choices between groups of items take into account how well each set "fits" together, leading to predictable shifts in preferences when choosing between sets rather than individual items.

There is some prior work supporting the intuition that certain sets are seen as better-fitting (and therefore more pleasing) than others. Research in Gestalt psychology has uncovered a number of principles that determine how individual items are perceptually grouped into sets (e.g., Wertheimer, 1923; Koffka, 1922; for a recent review article see Wagemans et al., 2012). For example, the Gestalt principle of similarity states that all else being equal, similar stimuli are more likely to be grouped together than dissimilar stimuli. Gestalt perception happens rapidly and automatically, and this feeling of perceptual fluency may make sets that form a good gestalt more pleasing—this would be a specific case of the general rule that easy-to-process stimuli are liked more (Winkielman & Cacioppo, 2001).

In addition to the Gestalt principles, are there any other general criteria of set "goodness"? Philosophers have argued that stimuli that can be described by simpler rules are more aesthetically pleasing (Kintsch, 2012; Schmidhuber, 1997).[1] Similarly, research on the aesthetics of patterns has found that patterns that can be described by simple rules are preferred to those for which no simple descriptive rules exist (Glanzer & Clark, 1963; Garner & Clement, 1963; Garner, 1970). Taken together, this evidence suggests that, all else being equal, sets that follow simple organizing principles will be liked better than sets that do not.

One common organizing principle is the similarity or dissimilarity of a set's members. Sets whose members are all similar or all different can be described by simpler rules than sets where some items are similar and others are different. The latter can only be described in less parsimonious terms—for example, "two items of one kind, three of another, and three more of three different kinds." Of course, "all-similar" or "all-different" are not the only organizing principles that might apply to sets. For example, a set of numbers following the rule "sequentially increasing," as in a straight hand in poker, follows a simple organizing principle, whereas a set of randomly chosen numbers does not. Here, however, given their ubiquity in everyday life, here we focus on all-similar and all-different sets.

We examined the perceived quality of all-similar and all-different sets in a pretest in which we asked participants to rate several sets of pictures of dinosaurs. Some of these sets followed "all similar" or "all different" rules (all dinosaurs were either all of the same type, or all of a different type, and all were either exactly the same color, or all of a different color), and some did not. Consistent with our hypothesis, participants rated sets as better when they followed one of these rules for both shape and color. We also asked the participants how the sets could be improved and

This article was published Online First June 17, 2013.

Ellen R. K. Evers, Yoel Inbar, and Marcel Zeelenberg, Department of Social Psychology & TIBER, Tilburg University, Tilburg, the Netherlands.

The authors would like to thank Dirk Laurijsen for his help conducting Experiment 3. Furthermore, we would like to thank Philippe van de Calseyde, George Loewenstein, Joe Simmons, and Uri Simonsohn for their helpful suggestions on previous drafts of this article.

Correspondence concerning this article should be addressed to Ellen R. K. Evers, Department of Social Psychology & TIBER, Tilburg University, PO Box 90153, 5000 LE, Tilburg, the Netherlands. E-mail: e.r.k.evers@tilburguniversity.edu

[1] "Simplicity" is defined in terms of Kolmogorov complexity, a term used in algorithmic information theory to denote the minimal amount of resources needed to describe an object. In other words, the Kolmogorov complexity of an object is the shortest possible description of it.

found that a large majority (89.6%) of participants suggested changes that would make the set's items either all-similar or all-different on important attributes.[2]

Based on these results, in the current research we operationalize good sets as those where all items "fit" together according to these simple rules—that is, good sets are those where all items are the same, or all differ, on all salient features. Bad sets are those where one or more items do not "fit"—that is, they violate the all-the-same or all-different principle. We hypothesized that people would choose sets in which items fit together over ones where they did not, even when the bad sets were superior on an item-by-item basis, and that people would be reluctant to choose items that would fit badly with a set of items already in their possession.

Of course, there are cases in which items in a set are complementary, so that the combination of the items adds extra utility. For example, one might prefer a tennis racket over a bowling ball, but a bowling ball and nine pins over a tennis racket and nine pins. However, in the cases we examine, there are no obvious complementarities between goods, so a "good set" is no more useful than a "bad set." Nor, as we see, is the set-fit effect explained by existing phenomena known to affect choice, such as attraction and compromise effects (Simonson, 1989; Simonson & Tversky, 1992), differences in evaluability (Hsee, 1996; Hsee & Zhang, 2010), or variety-seeking (Ratner, Kahn, & Kahneman, 1999; Simonson, 1990). Below, we test the set-fit effect in four experiments.[3]

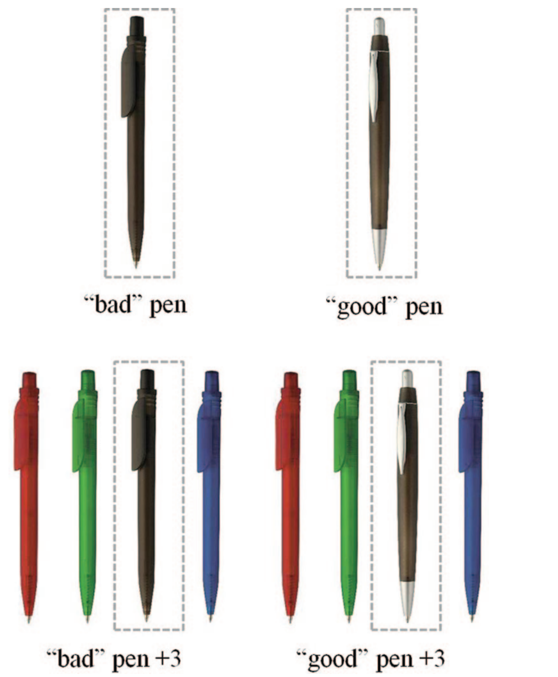

Figure 1. Stimuli in the *individual-choice* condition (top) and *set-choice* condition (bottom) of Experiment 1. Note: Boxes and labels are added for clarity and were not displayed to participants.

Experiment 1

One hundred four participants recruited via Amazon.com's Mechanical Turk (66 females, $M_{age} = 34.79$) were randomly assigned to one of two conditions. In the *individual-choice* condition, participants chose between a superior metal-accented pen and an inferior all-plastic pen. In the *set-choice* condition, participants also chose between these two pens, but with the same three all-plastic pens added to each option (see Figure 1). We expected that participants would prefer the superior over the inferior pen in the *individual-choice* condition, and indeed 78.8% (41/52) chose the superior pen. In the *set-choice* condition, we expected more participants to choose the set including the inferior pen (which fit with the other pens) over the set including the superior pen (which did not). Indeed, 50.0% (26/52) chose the set including the inferior pen, a significant difference from the individual-choice condition, $\chi^2(1, N = 104) = 9.44$, $p < .01$, $\varphi = .30$.

Experiment 2

In Experiment 2, we extend the first study in three ways: First, we attempted to conceptually replicate the set-fit effect using different stimuli. Second, all choices were between sets of items, in order to rule out alternative explanations involving extraneous differences in how people choose between sets and single items (e.g., participants evaluating individual choices more locally and sets more globally; Kimchi & Palmer, 1982).Finally, participants made real choices between goods (Experiment 1's were hypothetical).

Participants (199 Tilburg University [Tilburg, the Netherlands] students; 137 female, $M_{age} = 21.03$) chose one of two sets of paperclips to take home. In the *color* condition participants chose between a set of four plastic-covered colored paperclips (red, green, yellow, and blue) and a set of three colored clips (the red, green, and blue clips from the first set) plus one ordinary metal clip. In the *metal* condition, participants chose between a set of the ordinary metal clip plus three metal animal-shaped clips, and a set of the same three metal animal-shaped clips plus a yellow clip (see Figure 2). Thus, the ordinary metal clip and the yellow clip appeared in each condition, but as part of different sets. As predicted by the set-fit account, the set containing the metal clip was preferred by 63% of participants (63/100) in the *metal* condition (i.e., when the metal clip fit with the rest of the set); but by only 18% of participants (18/99) in the *color* condition (i.e., when the yellow clip fit with the rest of the set), $\chi^2(1, N = 199) = 41.40$, $p < .001$, $\varphi = .46$.

Experiment 3

Experiment 2 demonstrated the set-fit effect in real choices, but perhaps not particularly consequential choices. We therefore conducted Experiment 3 with young children using (sets of) marbles. Marbles are of considerable value to Dutch schoolchildren, espe-

[2] For a complete description of the pretest, please see the online supplemental materials.

[3] Termination of data collection was decided in advance, based either on a set length of time or a set total amount of participants. All conditions we ran are reported. Finally, only choice was measured as dependent variable. For exploratory reasons we also asked the participants to explain their choices in Experiments 2 and 3 (the two experiments in which the participants interacted directly with experimenters). We did not ask for these reasons with any statistical analysis in mind and therefore these "results" are only addressed in the General Discussion.

cially during the spring "marble season" when this experiment was run.

Children from the first through third grades of the elementary school "De Bracken" (in Rijen, the Netherlands; $N = 140$; 75 females; $M_{age} = 7.51$) volunteered and participated individually. The design paralleled that of Experiment 1. In the *individual-choice* condition, participants chose between a blue marble with white stripes and a monochrome red marble (pretesting revealed that children saw the former as more desirable). In the *set-choice* condition, participants also chose between these two marbles, but with the same three monochrome marbles added to each option (see Figure 3). As predicted by the set-fit account, 89.6% of participants (43/48) preferred the superior striped marble over the monochrome red marble in the *individual-choice* condition, but the set including this marble was preferred by only 27.3% (12/44) in the *set-choice* condition, $\chi^2(1, N = 92) = 37.1$, $p < .001$, $\varphi = .63$.[4]

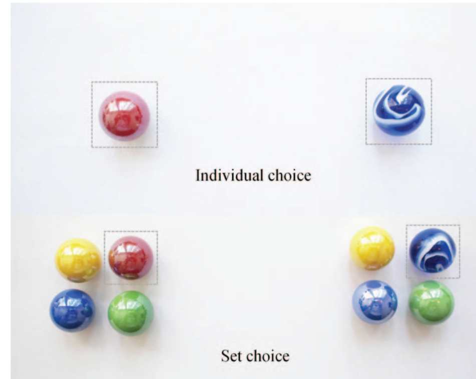

Figure 3. Choice options in the *individual-choice* condition (top) and *set-choice* condition (bottom) of Experiment 3. Choice options were physically shown to participants.

Experiment 4

Experiments 1–3 showed that people are attracted to sets that fit well together, so much so that they sometimes choose a set of inferior items over one including a superior item. Experiment 4 tests whether making an item fit better or worse with a set of already endowed goods can increase and decrease its choice share, respectively.

Participants (207 Fontys University Tilburg [Tilburg, the Netherlands] students; 154 female; $M_{age} = 19.7$) imagined buying specialty beer for a party. Participants were told that they had already chosen three beers and that they could choose one more from an assortment of four, which included a beer from Rochefort Abbey (see Figure 4). In the *control* condition, participants were given no further information and simply chose which beer they would purchase. In the remaining two conditions, participants saw the three beers they had (allegedly) already chosen: In the *all-similar* condition, these were three beers from Rochefort Abbey; in the *all-different* condition, these were beers from three different brewers including Rochefort. We expected that compared to those in the control condition, participants in the *all-similar* condition would be more likely to choose the Rochefort beer (because it would fit best with the "all-similar" set), and those in the *all-different* condition would be *less* likely to choose this beer (because it would not fit with the "all -different" set). Indeed, compared to the control condition, in which 11/74 (14.9%) chose the Rochefort beer, participants in the *all-similar* condition were more likely to choose the Rochefort (55/69; 79.7%), $\chi^2(1, N = 143) = 60.41$, $p < .001$, and those in the *all-different* condition were less likely to (2/64; 2.9%), $\chi^2(1, N = 138) = 5.54$, $p < .001$. Note that although there are alternative explanations for why Rochefort's choice share would increase in the *all-similar* condition (e.g., telling participants that they had already chosen three Rocheforts may have been seen as an implicit recommendation or as informative about their or their host's preferences) none of these explanations would explain why Rochefort's choice share would *decrease* (relative to control) in the *all-different* condition.

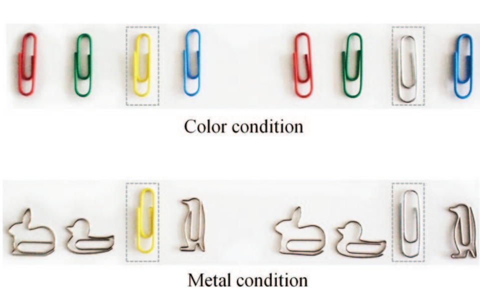

Figure 2. Choice sets in the *color* condition (top) and *metal* condition (bottom) of Experiment 2. These sets were physically shown to participants. Note that the yellow and metal clips appear once per set in each condition.

General Discussion

Four experiments revealed that when choosing between (or adding to) sets of items, people's choices reflect set-fit in addition to preferences over the set's individual items. This can cause predictable preference reversals when an individually preferred item is bundled with other items with which it fits poorly.

The notion of set-fit can explain behavior that is un- or mispredicted by well-known accounts of nonnormative influences on choice, such as attraction and compromise effects (Simonson, 1989), or contrast or assimilation (Wänke, Bless, & Schwarz, 1999). Attraction or compromise effects occur when choosers must make trade-offs between domains (e.g., price and quality) when choosing between roughly equally attractive options. These effects do not apply to cases where one choice option is clearly inferior, whereas the set-fit account predicts that inferior options can be preferred when they fit well with other bundled items. Nor can contrast or assimilation explain our results: If participants in Experiments 1 and 3 contrasted the superior items with the inferior items in the set, this should have increased the choice share of this set by making the superior item seem even more attractive. Sim-

[4] We also ran one other control condition ($n = 48$) where children chose between a set of four blue/white striped marbles and a set of three blue/white striped marbles and one monochrome red marble. As expected, a large majority of participants in this condition (75%) preferred the set of four blue/white striped marbles.

ilarly, an assimilation account would not predict that people would prefer a well-fitting set of inferior items (as was the case in Experiment 3) and would make no predictions regarding Experiments 2 and 4, where no item or set was of obviously higher quality.

Even though our approach builds on Gestalt theory's "similarity principle," according to which items similar in shape or hue are perceptually grouped (Wertheimer, 1923), set-fit is not just an application of Gestalt theory. First, the similarity principle explains a perceptual phenomenon, not preferences or choices; second, of course, it would not predict (as we do) that good sets can also follow an "all-different" principle. That said, we do believe our findings fit with the larger "Gestaltian" notion that people naturally structure groups of stimuli in order to reduce complexity. Furthermore, we believe that the similarity and difference principles we investigated here are not the only ones that determine perceptions of set-fit (and thus affect choices). Theoretically, any organizing principle that reduces complexity should affect perceptions of set-fit. Some possibilities include other Gestalt principles, such as common fate (moving together; Wertheimer, 1923), numbering (e.g., a set consisting of 1, 2, 3, 4, 5, 6, 7 should be preferred over one consisting of 1, 4, 8, 2, 13, 6) and perceptual symmetry (Garner, 1970). Testing these possibilities goes beyond the scope of the current research, but we expect that any organizing principle that reduces set complexity should function in the same way as the principles we have focused on here.

One challenge for any alternative to the set-fit account would be to explain why people prefer similarity under some circumstances but variety under others. As we showed in Experiment 4, fit (or lack thereof) with either an all-similar or all-different set can affect the attractiveness of a choice option. We ran an additional test of this idea, including all hypothesized effects within one experiment. We ran four conditions in which participants chose between two sets of mugs (see Figure 5). These sets were designed in such a way that in each condition the only difference between the two choices was a green vs. an orange mug. Adding three green mugs to both options (*all-similar green*) we found that the majority

Figure 4. Choice options and endowed sets from Experiment 4. The top panel shows the choice options. The bottom panels show the beers participants in each condition were said to have already selected.

Figure 5. Participants chose between two sets of mugs (see supporting materials for details). Note that the only difference between the two choice options is always the green versus orange mug. Boxes are added to the pictures for clarification but were not part of the original stimuli.

(60.4%) chose the set with the fourth green mug. Adding three orange mugs to both options (*all-similar orange*) reversed preferences—now the majority (62.3%) chose the set of four orange mugs ($p = .03$). These two conditions reveal *the all-similar* effect. In the other two conditions we found the *all-different* effect. In the first condition we added a blue, pink, and orange mug to both options. In this case the majority (84.3%) chose the set with the green mug. When we added a blue, pink, and green mug to the two original options, the majority (94.1%) chose the set with the orange mug ($p < .001$).[5]

Moderators and Boundary Conditions

One obvious precondition for set-fit effects is that a group of items is actually perceived as one set. Based on the Gestalt principle of proximity (Wertheimer, 1923), we would expect the shapes in Figure 6A to be perceived as one poorly fitting set of three triangles and a circle, whereas the same shapes in Figure 6B should be more likely to look like one well-fitting set of three triangles plus an extra circle. Thus, subtle perceptual features may have large effects on the influence of set-fit and, therefore, on the attractiveness of sets.

The set-fit effect seems to rely more on intuition than deliberation. Even though investigating how intuitive the effect is was not the primary goal of this research, we did ask participants in Experiments 2 and 3 why they chose as they did. Participants in Experiment 2 often looked confused when they were asked to explain their choices. Some literally said "I don't know" others said "I guess because these seem to belong together" or "the yellow [clip] just really doesn't seem to belong together with the other ones." Coming up with a reason for the choice also generally took much longer than making the choice itself. The children in Experiment 3 reacted similarly, mentioning that they felt some marbles "belonged" together or—pointing at the set they did not choose—that the nonfitting marble "just felt weird."

These reactions are consistent with the idea that set-fit effects are the result of basic perceptual grouping principles (Wertheimer, 1923) that are applied relatively automatically and uncontrollably. Following Gestalt psychologists, we believe that positive evalua-

[5] For a complete description of this experiment, please see the supplemental materials.

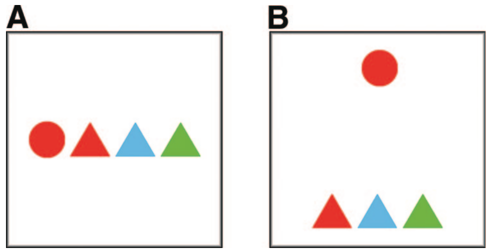

Figure 6. A. A "bad set" of four three triangles and a circle. B. A "good set" of three triangles and an additional circle.

tions of good sets do indeed occur rapidly and effortlessly. However, this does not mean that better-fitting sets will always be preferred in choice. Whereas the evaluation of the set and its configuration occurs automatically, its effect on choices may be overridden by more deliberate processing of the options or other judgment processes. For example, four boys are a "better set" than three boys and one girl, but we do not believe a pregnant mother of three boys would wish for a fourth in order to have a more pleasing set of children. In this case, other, more deliberative influences—e.g., the anticipated pleasure to be gained from another boy versus a girl, or a commitment to loving a child of either gender equally—are likely to greatly outweigh the influence of set-fit. Set-fit effects on evaluations of set quality *per se* are thus always expected, but set-fit effects on actual choices are likely to be most pronounced when people rely on their immediate intuitive responses, and weaker when decisions are based on deliberative reasoning. This also means that when investigating set-fit effects it is important to keep such other considerations stable.

To what extent does variety-seeking conflict with maximizing set-fit? Note that the processes underlying set-fit and variety-seeking are very different, so they are likely to exert independent effects on choices. For goods or experiences that are chosen and consumed sequentially over time (e.g., Ratner et al., 1999), we would not expect set-fit effects, as it is unlikely that these choices would spontaneously be construed as a set. When choices are bracketed together (Read & Loewenstein, 1995) set-fit effects may *amplify* variety-seeking. If people desire at least some diversification and therefore start by choosing different items, a concern with set-fit should lead them to favor greater diversity (i.e., an all-different set a partially different set). Of course, in cases where the set is all-similar, set-fit and variety-seeking effects operate in opposite directions: Set-fit increases the chance of a similar item being chosen, whereas variety-seeking increases the chance of a different item being chosen. In these situations the outcome depends on the relative strength set-fit and variety-seeking. In the experiments described in this article, considerations that amplify variety-seeking—such as satiation (McAlister, 1982), overestimations of satiation due to time contraction (Galak, Kruger, & Loewenstein, 2011), or spreading of risk (Simonson, 1990)—were minor or nonexistent, and set-fit predominated, leading people to choose the all-similar set. In choices where these considerations play more of a role—for example, when physical satiation creates a strong desire for variety—we would expect variety-seeking to predominate.

Conclusion

The results reported here add to a large and growing body of work showing that—counter to the assumptions of classical economics—people's preferences are often unstable and subject to minor contextual factors (Hsee, 1996; Lichtenstein & Slovic, 1971; McGraw, Shafir, & Todorov, 2010). In the case of set-fit, this may have a number of important consequences, including the fact that improvements to products that are seen as part of a set may unexpectedly decrease their appeal when the improvements make the overall set worse. One real-life cautionary example is the case of a Dutch publisher who attempted to improve the design of a popular series of books by changing the covers, paper, and typography; these changes were the result of much careful research regarding consumers' preferences. However, the publisher did not anticipate that many existing customers would be upset because the new, redesigned books would not "fit" with those already purchased. Finally—after many angry letters—the publisher decided to release the books in both old and new formats.

References

Galak, J., Kruger, J., & Loewenstein, G. (2011). Is variety the spice of life? It all depends on the rate of consumption. *Judgment and Decision Making, 6,* 230–238.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58,* 34–42.

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning & Verbal Behavior, 2,* 446–452. doi:10.1016/S0022-5371(63)80046-8

Glanzer, M., & Clark, W. H. (1963). Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning & Verbal Behavior, 1,* 289–299. doi:10.1016/S0022-5371(63)80008-0

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,* 247–257. doi:10.1006/obhd.1996.0077

Hsee, C. K., & Zhang, J. A. (2010). General evaluability theory. *Perspectives on Psychological Science, 5,* 343–355. doi:10.1177/1745691610374586

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 521–535. doi:10.1037/0096-1523.8.4.521

Kintsch, W. (2012). Musings about beauty. *Cognitive Science, 36,* 635–654. doi:10.1111/j.1551-6709.2011.01229.x

Koffka, K. (1922). Perception: An introduction to the gestalt-theorie. *Psychological Bulletin, 19,* 531–585. doi:10.1037/h0072422

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling situations. *Journal of Experimental Psychology, 89,* 46–55. doi:10.1037/h0031207

McAlister, L. (1982). A dynamic attribute satiation model of variety-seeking behavior. *Journal of Consumer Research, 9,* 141–150. doi:10.1086/208907

McGraw, A. P., Shafir, E., & Todorov, A. (2010). Valuing money and things: Why a $20 item can be worth more and less than $20. *Management Science, 56,* 816–830. doi:10.1287/mnsc.1100.1147

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research, 26,* 1–15. doi:10.1086/209547

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separate choices. *Journal of Experimental Psychology: Applied, 1,* 34–49. doi:10.1037/1076-898X.1.1.34

Schmidhuber, J. (1997). Low-complexity art. *Leonardo, Journal of the International Society for the Arts, Sciences, and Technology, 30,* 97–103.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174. doi:10.1086/209205

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research, 27,* 150–162. doi:10.2307/3172842

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281–295. doi:10.2307/3172740

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138,* 1172–1217. doi:10.1037/a0029333

Wänke, M., Bless, H., & Schwarz, N. (1999). Assimilation and contrast in brand and product evaluations: Implications for marketing. *Advances in Consumer Research, 26,* 95–98.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II [Investigation on gestalt principles: II]. *Psychologische Forschung, 4,* 301–350. doi:10.1007/BF00410640

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81,* 989–1000. doi:10.1037/0022-3514.81.6.989

Received November 11, 2012
Revision received May 1, 2013
Accepted May 6, 2013 ■