Original Communication

When the Others Matter

Context-Dependent Effects on Changes in Appreciation of Innovativeness

Gernot Gerger[1], Helmut Leder[1], Stella J. Faerber[2], and Claus-Christian Carbon[2]

[1]Faculty of Psychology, University of Vienna, Vienna, Austria, [2]University of Bamberg, Germany

Abstract. Although innovativeness is an important variable in product design, we know little about its appreciation. We studied how appreciation of innovativeness and its dynamics depends on the heterogeneity of the context in which it appears. We employed a test-retest design in which appreciation of car interior designs was tested before and after repeated evaluations. We tested heterogeneous stimulus sets (highly *and* lowly innovative designs together; Experiment 1) and homogeneous stimulus sets (highly *or* lowly innovative designs; Experiment 2). The known effect (Carbon, Hutzler, & Minge, 2006; Carbon & Leder, 2005) of a selective increase in attractiveness ratings for highly innovative stimuli after repeated evaluations was only obtained for heterogeneous sets. In homogeneous sets, both highly and lowly innovative interiors were rated similarly and showed similar dynamics. Experiment 3 was a shorter version of Experiment 1, which ruled out differences in experimental design (more ratings and longer duration in Experiment 1) as the cause of the differences. High innovativeness was found to show a specific increase in attractiveness ratings only when innovativeness was made apparent by presenting stimuli in heterogeneous sets. Thus, awareness of variation in innovativeness as a relevant stimulus dimension is a key feature regarding its effect on appreciation.

Keywords: attractiveness, innovativeness, dynamics, repeated evaluation, context, consumer psychology

Product esthetics are essential in modern consumer markets (Hekkert & Leder, 2008). For example, Apple's success has been attributed to its intense focus on attractive and innovative product design. Product design that is attractive through innovativeness (Carbon & Leder, 2005; Hekkert, Snelders, & Van Wieringen, 2003) can obviously be a key for distinguishing between winners and losers in a market (Liu, 2003; Mairesse & Mohnen, 2002). It could be quite costly for a company if innovativeness was not found attractive by customers (Cooper, 2001; Moulson & Sproles, 2000). In consequence, innovativeness is often stressed as being essential for consumer products and a driving force in cultural and industrial progress (see Cox, 2005; and for the arts, Leder, Belke, Oeberst, & Augustin, 2004).

Psychologically, innovativeness is not a very well-defined dimension. The specific nature of innovativeness makes it an interesting dimension regarding attitude formation (Schwarz, 2007). Innovativeness can be defined as "originality by virtue of introducing new ideas" (Carbon & Leder, 2005, p. 587). It involves novel and sometimes unusual stimulus features. Different from novelty, innovativeness remains innovative for some time, while novelty cannot persist (Carbon & Leder, 2005). In design, innovativeness can be extracted from expert knowledge or from concept design studies. However, because what is seen as innovative might differ between perceivers, the effects of innovativeness in empirical studies warrant individual assessment (Carbon & Leder, 2005). In a series of studies using car interiors, systematically varying in innovativeness, Carbon and Leder (2005), Leder and Carbon (2005), and Carbon et al. (2006) found that innovativeness was often rejected at first. When seen for the first time, lowly innovative car interiors were appreciated more than highly innovative car interiors. However, after active elaboration through repeated evaluations (Carbon & Leder, 2005; Carbon et al., 2006), attractiveness selectively increased for the highly innovative car interiors. Critically, in these studies, because highly and lowly innovative stimuli were rated together in one set, innovativeness might have been made apparent through the heterogeneity of the evaluated set. The present study addressed whether an increase in appreciation for highly innovative designs requires explicit awareness of innovativeness as an important dimension, relative to other stimuli. We tested whether heterogeneity in innovativeness is necessary to selectively increase attractiveness for high innovativeness. For example, the perceived attractiveness of a new, innovative car design depends on the other cars that one knows or on the other cars that are presented during the evaluation. Alternatively, innovativeness might also be evaluated independently of either. Understanding under which conditions innovative designs become appreciated would inform us about the nature

DOI 10.1024/1421-0185/a000041

of cognitive-affective evaluations and whether the formation of attitudes is automatic or stimulus-/context-dependent.

The present study used a dynamic test paradigm of repeated evaluations (Carbon & Leder, 2005) and compared homogeneous (only highly or only lowly innovative car interiors) and heterogeneous (highly and lowly innovative car interiors) stimulus sets. Differences should reveal whether an appreciation of high innovativeness requires a direct comparison between highly and lowly innovative stimuli or whether innovativeness, when it is seen, produces effects per se, and thus relies on an internal standard of comparison. The latter would be in accordance with effects of an independent, inner standard, similar to specific responses to stimulus features – such as absolute pitch (Takeuchi & Hulse, 1993). The former would be in accordance with effects that depend on differences with other stimuli in the set (Helson, 1948; Parducci, 1995) or, as in paradigms of mismatch negativity, when effects are only measured when a deviating stimulus suddenly appears (Cammann, 1990; Tiitinen, May, Reinikainen, & Näätänen, 1994).

We employed a paradigm devised by Bornstein, Kale, and Cornell (1990), in which set homogeneity was varied with respect to how demanding the visual stimuli were and whether they were shown repeatedly. By using two classes of stimuli – visually demanding optical illusions and simple geometrical line drawings – they found that the effects of stimulus repetition on attractiveness ratings depended on the homogeneity of the stimulus set. Using a between-subjects design in which only one homogeneous stimulus class was shown repeatedly and then evaluated, they found that attractiveness ratings linearly increased for both classes. Additionally, the attractiveness ratings for both classes of stimuli were similar when seen for the first time. However, when both classes of stimuli were shown together in one set when they were evaluated for the first time, the simple figures received lower ratings of attractiveness than the optical illusions. In this heterogeneous set, stimulus repetition resulted in increased attractiveness ratings for the optical illusions, but not for the simple geometric figures. Additionally, after repetition, both classes of stimuli showed a decrease in attractiveness ratings, which was interpreted as the effect of boredom (Berlyne, 1970b; Stang, 1974). Thus, regarding the dimension of visual demands, complexity-dependent changes in attractiveness only emerged when differences in stimulus features were made apparent by simultaneous presentation of the stimuli. Applying a similar design, we test whether the effects of innovativeness (as in Carbon & Leder, 2005) also depend on such context or set effects.

Esthetic appreciation is often studied by measuring attractiveness (Hekkert & Leder, 2008). Attractiveness is a summarizing evaluation representing affective and cognitive aspects (Leder, Augustin, & Belke, 2005; Leder et al., 2004) in which a number of related concepts are involved. Regarding the structure of esthetic evaluations, Faerber, Leder, Gerger, and Carbon (2010) showed how the activation of specific attractiveness-related concepts produces different effects with regard to esthetic appreciation. They tested a semantic network approach to esthetic appreciation by comparing different priming conditions. In these studies, when participants had been primed for innovativeness, changes in attractiveness were observed. However, it is unclear whether these changes depend on the range of innovativeness in the stimulus sets. Different theoretical explanations make different predictions regarding set effects when stimuli are presented in either homogeneous or heterogeneous sets.

The following theories propose that appreciation of innovativeness could depend on a kind of internal, pre-existing standard of comparison: prototype, evolutionary-novelty, and two-factor theory. According to prototype theory (e.g., Halberstadt, 2006; Halberstadt & Rhodes, 2003; Rosch, 1978), each stimulus is matched against an internal prototype (based on previous experiences). Innovative stimuli might be more dissimilar to an internal prototype because they are more dissimilar to familiar (see prototypical) stimuli. If so, then low prototypicality is not preferred (Halberstadt, 2006). Repeated evaluation increases familiarity and might cause minor changes in the internal prototype (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003) toward greater innovativeness (Carbon & Leder, 2005). Therefore, appreciation of highly innovative designs will increase over time. Because stimuli are matched to a pre-existing internal prototype, one might expect to find similar evaluations and dynamics regardless of whether a homogeneous or heterogeneous set is used. However, recent data from the domain of face research have questioned this theory, at least for the process of assessing attractiveness of faces by matching them with an internal prototype (Carbon, Grüter, Grüter, Weber, & Lüschow, 2010).

Evolutionary accounts claiming that stimulus novelty (see innovativeness) results in ambiguity or uncertainty (Lee, 2001; Robinson & Elias, 2005) make the same predictions. A perceiver cannot know whether a novel (see innovative) stimulus is potentially harmful. As a consequence, approach and avoidance behaviors would be triggered simultaneously, resulting in attenuated attractiveness judgments. If this initial ambiguity is overcome through repeated evaluations, then attractiveness judgments eventually increase. Regardless of set combination, one might find lower attractiveness ratings for highly innovative than for lowly innovative stimuli when rated for the first time. After repeated evaluation, the attractiveness of innovative stimuli would increase.

The two-factor theory of Berlyne (1970b) and Stang (1974) arrives at the same conclusions. According to this theory, the more a stimulus is embedded in our cognitive system, the more positively it will be evaluated until boredom sets in, which then affects evaluations. Embedding occurs through repeated evaluations and results in positive habituation (Berlyne, 1970b) and increased familiarity (Zajonc, 2001). Moreover, processing fluency of the stimulus also increases (Bornstein & D'Agostini, 1994; Reber, Schwarz, & Winkielman, 2004). All of these factors increase attractiveness evaluations (Carbon, 2010) until boredom sets in, at which point the positive evaluations begin to wane (Berlyne, 1970b; Stang, 1974). Because of their relative novelty, highly innovative stimuli are less embedded in our cognitive system than lowly innovative stimuli. Consequently, the two-factor theory would predict increases for highly but not for lowly innovative stimuli, regardless of whether innovative stimuli are shown within a heterogeneous or homogeneous set.

However, if appreciation of innovativeness depends on a relative standard of comparison, this would be in accordance with a different rationale of the two-factor model (Berlyne, 1970b; Stang, 1974) or arousal theory (Berlyne, 1970a). According to the two-factor theory, effects of habituation and boredom on attractiveness ratings could also be *relative* depending on the stimulus set. After repeated evaluations, the highly innovative stimuli might be perceived as less boring compared to the lowly innovative stimuli. But if only one set is rated, then the boredom effects within the set will be similar and independent of level of innovativeness. Thus, different dynamics might develop when the range of innovativeness is different.

Arousal theories (Berlyne, 1970a) also predict that attractiveness ratings depend on relative differences between the stimuli. They assume that medium levels of arousal result in the highest attractiveness ratings. Importantly, according to the arousal account, evaluations critically depend on the *relative* arousal level induced by different stimuli. Highly innovative stimuli – because of their novelty, unexpectedness, and unusualness – when seen for the first time might produce higher suboptimal arousal levels than lowly innovative stimuli. Through repeated evaluations, this initially high arousal might be reduced to a medium level, while the arousal level of the lowly innovative stimuli might be reduced to a suboptimal level. Thus, when both innovativeness levels are judged together, arousal differences due to innovativeness might be highly apparent and determine their attractiveness. On the other hand, in a homogeneous set, the arousal levels associated with the stimuli might be similar, which would result in more similar attractiveness evaluations and dynamics. Evidence for such changes in arousal was also found in Carbon, Michael, and Leder (2008), who measured electrodermal activity indicative of arousal. Highly innovative material showed physiological effects in accordance with maintaining positive arousal after repeated evaluations.

Thus, the present experiments will test the following hypotheses: If prototype, novelty, or two-factor explanations account for the effects of innovativeness, we should find similar effects in heterogeneous and homogeneous stimulus sets. However, if the appreciation of innovativeness depends on relative differences as suggested by arousal or based on the relative boredom level (according to the two-factor theory), then effects should differ between the conditions.

Experiment 1

A heterogeneous stimulus set was used in Experiment 1. It was based on Carbon and Leder's (2005) experimental paradigm, the "repeated evaluation technique" (RET). Car interiors were judged for attractiveness and innovativeness before and after a phase of repeated stimulus evaluations. Thus, Experiment 1 served as a baseline replication of Carbon and Leder but used photorealistic instead of line-drawing stimuli.

Method

Participants

Twenty-seven participants (19 female, 8 male) enrolled in various introductory psychology courses at the University of Vienna, Austria, participated in the experiment for partial course credit. The participants' mean age was 21.7 years (range: 18 to 28 years).

Figure 1. Examples of stimuli used: a highly innovative car interior (left) and a less innovative car interior (right).

Table 1
Mean attractiveness and innovativeness ratings for experiments 1, 2, and 3, separately for T1 and T2

	Experiment 1 heterogeneous set – long		Experiment 2 homogeneous set		Experiment 3 heterogeneous set – short	
	Attractiveness					
	T1	T2	T1	T2	T1	T2
Lowly innovative	3.35 (1.00)	3.32 (0.85)	2.96 (1.00)	3.60 (0.80)	3.15 (0.80)	3.47 (0.85)
Highly innovative	2.90 (0.72)	3.75 (1.18)	3.26 (1.02)	3.77 (0.80)	3.06 (1.16)	4.98 (1.00)
	Innovativeness					
Lowly innovative	3.11 (0.76)	3.05 (0.69)	3.41 (0.97)	3.62 (0.69)	3.20 (0.91)	3.40 (1.16)
Highly innovative	4.44 (1.05)	4.63 (0.94)	3.58 (1.00)	3.84 (1.02)	3.87 (0.90)	4.44 (0.78)

Note. Standard deviations are in brackets.

Stimuli

Eighteen photorealistic grayscale depictions of car interiors were created in Adobe Photoshop 7. These differed in two levels of innovativeness (Carbon & Leder, 2005) as confirmed by pretests (for examples, see Figure 1). As in Faerber et al. (2010), the two levels of innovativeness (low, high) were fully crossed with levels of complexity and curvature on 3 × 3 levels (low, medium, high) (for a detailed description of these dimensions, see Carbon & Leder, 2005, and Leder & Carbon, 2005). Several previous studies ensured equal degrees of complexity and curvature for the two levels of innovativeness using 7-point Likert scales (Carbon & Talker, 2006).

Apparatus

The experiment was administered using PsyScope PPC 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993) on Apple eMac computers. Stimuli were centrally presented on a 17-inch monitor at a size of 678 × 438 pixels with a screen resolution of 1024 × 768 pixels.

Procedure

Experiment 1 consisted of three consecutive parts. In an initial rating phase (Test phase 1: T1), participants first rated all of the car interiors with respect to their attractiveness and then their innovativeness. In the second part, all stimuli were evaluated on 25 different scales. This repeated exposure phase was used to ensure that participants actively elaborated the stimuli. The stimuli were rated on the following dimensions (as in Carbon & Leder, 2005; German terms in parentheses): repellent (abschreckend), pleasant (angenehm), appealing (ansprechend), unsophisticated (bieder), carefully designed (durchdacht), inviting (einladend), elegant (elegant), overwhelming (erdrückend), extravagant (extravagant), funky (flippig), futuristic (futuristisch), dignified (gediegen), tasteful (geschmacksvoll), of high quality (hochwertig), tacky (kitschig), easily operated (komfortabel), conservative (konservativ), luxurious (luxuriös), modern (modern), plain (nüchtern), functional (praktisch), stylish (stilvoll), cluttered (unübersichtlich), ornamental (verspielt), and overloaded (überladen). The presentation order of these scales was randomized across participants. Finally, the participants rated all of the stimuli again for attractiveness and innovativeness (Test phase 2: T2). All ratings were made on 7-point Likert scales anchored with 1 = *hardly* (*wenig*) and 7 = *very* (*sehr*). The presentation order of the stimuli was randomized within each scale. Participants were not given time constraints for their ratings.

Results and Discussion

In Experiment 1, participants judged a set of stimuli that were heterogeneous in innovativeness (as in Carbon & Leder, 2005). As can be seen in Table 1, lowly innovative stimuli were rated higher on attractiveness than highly innovative stimuli at T1. However, after repeated evaluations, only highly innovative stimuli showed an increase in attractiveness at T2 (see Table 1 and Figure 2). As for the innovativeness ratings, the data showed a clear pattern (see Table 1): Highly and lowly innovative stimuli differed at T1 and T2. In order to analyze these effects, we calculated repeated measures ANOVAs with the factors time (T1, T2) and innovativeness (high, low) separately for the attractiveness and innovativeness ratings. For a sample of 27 participants, an effect size of $f = 0.36$ (i.e., between a medium, $f = 0.25$, and a large, $f = 0.40$, effect as defined by Cohen, 1988) can be detected with a probability of $1 - \alpha = .95$ (Faul, Erdfelder, Buchner, & Lang, 2009). For the follow-up dependent *t*-tests, a posteriori effect sizes were calculated according to equation 3 in Dunlap, Cortina, Vaslow, and Burke (1996).

Attractiveness Ratings

A 2 × 2 (time × innovativeness) repeated measures ANOVA for the attractiveness ratings revealed a significant main effect of time, $F(1, 26) = 10.99$, $p < .01$, $\eta_p^2 = 0.30$. Importantly, the effect of time was qualified by a significant interaction of

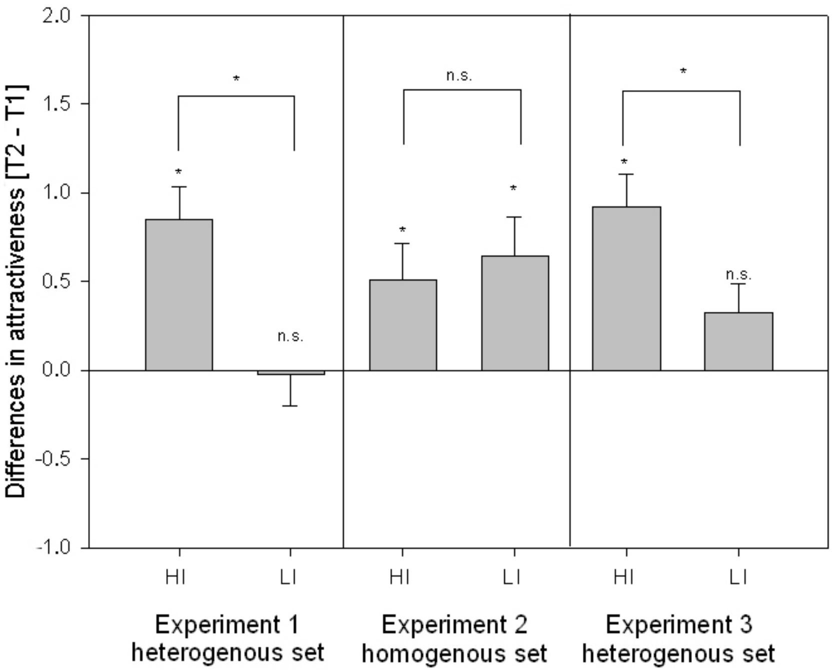

Figure 2. Interaction and main effects of the repeated measures ANOVAs expressed as change scores (T2 minus T1) in the attractiveness ratings of highly innovative (HI) and lowly innovative stimuli (LI). *significant differences at a $p = .05$ level. Error bars depict one standard error of the mean.

time and innovativeness, $F(1, 26) = 10.55, p < .01, \eta_p^2 = 0.29$, reflecting attractiveness dynamics that were dependent on the innovativeness level of the stimuli (see Figure 2). Dependent *t*-tests showed that this interaction was due to the highly innovative stimuli receiving significantly higher attractiveness ratings at T2 than at T1, $t(26) = 4.50, p < .01, d = 0.81$; it was not due to changes in the attractiveness of the lowly innovative stimuli, $t(26) = 0.14$, *ns*, $d = 0.03$. The results regarding the attractiveness ratings replicated (using more realistic stimuli) Carbon and Leder's (2005) findings. Highly innovative stimuli were initially rejected, but following repeated evaluations, they received higher attractiveness ratings than lowly innovative stimuli.

Innovativeness Ratings

A repeated measures ANOVA revealed a significant main effect for innovativeness, $F(1, 26) = 85.85, p < .01, \eta_p^2 = 0.77$, but no other effect. Thus, innovativeness ratings differed along our preclassification and remained stable over time (again as in Carbon & Leder, 2005).

Experiment 2

In Experiment 2 we tested the critical condition: In order to explore whether innovativeness – when not made apparent – influences attractiveness and its dynamics, participants evaluated a set of stimuli that were homogeneous in innovativeness (either only lowly or only highly innovative stimuli).

Method

Participants

Fifty-four participants (41 female, 13 male) enrolled in various introductory psychology courses at the University of Vienna participated in the experiment for partial course credit. The mean age of the participants was 22.5 years (range: 18 to 38 years). Participants were randomly assigned to one of two conditions (Experiments 2a and 2b) with the restriction that approximately the same amount of men were in each condition (six in Experiment 2a and seven in Experiment 2b).

Stimuli and Apparatus

The stimuli and apparatus used were the same as in Experiment 1. In Experiment 2a, only the lowly innovative stimuli (9 stimuli) were presented to the participants. In Experiment 2b, only the highly innovative stimuli (9 stimuli) were presented to the participants.

Procedure

Again, Experiments 2a and 2b consisted of three consecutive phases (T1, evaluation phase, T2). All ratings were given on 7-point Likert scales and were self-paced.

Results and Discussion

Attractiveness ratings were clearly different from those in Experiment 1. Comparing the results of Experiment 2a and

2b, highly innovative stimuli were judged to be slightly more attractive than lowly innovative stimuli. Over time, attractiveness ratings increased consistently for both the highly and lowly innovative stimuli (see Table 1 and Figure 2). Also, innovativeness evaluations did not reveal differences between the two conditions. To analyze these effects, we conducted a 2 × 2 repeated measures ANOVA with innovativeness (high, low) as a between-subjects factor and time (T1, T2) as a within-subjects factor, separately for the attractiveness and innovativeness ratings. Given the sample size of 54 participants (27 evaluating only the highly innovative, 27 only the lowly innovative stimuli), we could expect to detect medium-sized effects with $f = .25$ (Cohen, 1988) with a probability of $1 - \alpha = .95$ (Faul et al., 2009).

Attractiveness Ratings

A repeated measures ANOVA for attractiveness ratings alone yielded a significant main effect of time, $F(1, 52) = 14.86$, $p < .01$, $\eta_p^2 = 0.22$. After the repeated evaluations, attractiveness ratings for both innovativeness levels increased in a concordant manner (see Figure 2). No specific effects for either level of innovativeness were found. Thus, the effects of innovativeness on attractiveness clearly depend on which set combination was evaluated. Differential effects were only found in a heterogeneous set (in Experiment 1) in that attractiveness ratings for highly innovative stimuli increased after repeated stimulus evaluations.

Innovativeness Ratings

A repeated measures ANOVA for innovativeness ratings revealed no significant effects. However, the factor time showed a trend toward significance: $F(1, 52) = 4.03$, $p = .051$, $\eta_p^2 = 0.07$. Thus, innovativeness of the stimuli did not significantly affect the corresponding innovativeness ratings. So, if the variation on the dimension innovativeness is not made explicit and thus awareness is not specifically raised by showing highly and lowly innovative stimuli together, then innovativeness does not show a specific effect.

Experiment 3

Experiments 1 and 2 differed in two respects. First, participants were asked to provide more ratings in Experiment 1 as compared to Experiment 2. This was due to the different numbers of stimuli (18 in Experiment 1 vs. 9 in Experiment 2). Second, Experiment 1 had a longer total experiment duration. In order to rule out that the different dynamics in attractiveness ratings with a selective gain for highly innovative stimuli in Experiment 1, but not in Experiment 2, was due to the different experimental procedures, Experiment 3 employed the same experimental design as Experiment 1, but used a subset of only four highly and four lowly innovative stimuli. This resulted in approximately the same number of evaluations and the same experiment duration as Experiment 2.

Methods

Participants

Twenty-seven participants (24 female, 3 male) from the University of Vienna participated in the experiment for partial course credit. The participants' mean age was 22.2 years (range: 19 to 45).

Stimuli and Apparatus

In Experiment 3, only a subset of the stimuli from Experiment 1 was used. Stimuli with medium levels of complexity and curvature were omitted. Thus, the set consisted of four highly and four lowly innovative stimuli. The apparatus was the same as in Experiments 1 and 2.

Procedure

The same experimental procedure as in Experiments 1 and 2 was used.

Results and Discussion

In Experiment 1, attractiveness ratings selectively increased for highly innovative stimuli, but not for lowly innovative stimuli. Experiment 3 was conducted to rule out the possibility that the differences between Experiments 1 and 2 in the results regarding attractiveness were due to different experimental procedures. Descriptively, the results replicated the findings of Experiment 1. At T1, lowly innovative stimuli were preferred over highly innovative stimuli (although the difference was not as large as in Experiment 1). Importantly, following repeated evaluations, attractiveness ratings for highly innovative stimuli increased more than for lowly innovative stimuli (see Table 1). As in Experiment 1, innovativeness ratings between highly and lowly innovative stimuli were clearly different (see Table 1). In order to analyze these effects, we calculated repeated measures ANOVAs with the factors time (T1, T2) and innovativeness (high, low) separately for attractiveness and innovativeness ratings. With a sample size of 27 participants, we could expect to detect middle to large effects of $f = .36$ (Cohen, 1988) with a probability of $1 - \alpha = .95$. Effect sizes for the dependent *t*-tests were calculated according to Formula 3 in Dunlap et al. (1996).

Attractiveness Ratings

The repeated measures ANOVA for the attractiveness ratings showed a significant main effect of time, $F(1, 26) = 23.72$, $p < .01$, $\eta_p^2 = 0.48$, and a significant interaction of time and innovativeness, $F(1, 26) = 6.03$, $p = .021$, $\eta_p^2 = 0.18$ (see Figure 2). Follow-up dependent *t*-tests showed that the attractiveness ratings for the highly innovative stimuli increased significantly from T1 to T2, $t(26) = 4.90$, $p < .01$, $d = 0.88$. In contrast, they remained rather stable for the lowly innovative stimuli, $t(26) = 2.00$, *ns*, $d = 0.41$. These results replicated the findings of Experiment 1 and showed that, in a heterogeneous stimulus set, highly as compared to lowly innovative stimuli showed a greater increase over time. Awareness of innovativeness, resulting from the evaluation of highly and lowly innovative stimuli within one set, seems to be critical for the appreciation of innovativeness over time.

Innovativeness Ratings

The repeated measures ANOVA for innovativeness ratings yielded a significant main effect of innovativeness, $F(1, 26) = 21.78$, $p < .01$, $\eta_p^2 = 0.46$, and a main effect of time, $F(1, 26) = 6.32$, $p = .018$, $\eta_p^2 = 0.20$, which was due to an increase in innovativeness (see Table 1). However, there was no interaction between the two factors. Highly and lowly innovative stimuli were clearly different with regard to their innovativeness ratings. This suggests that the innovativeness of the stimuli was apparent in this smaller set.

General Discussion

It was known from previous studies that highly innovative designs increase in attractiveness after a series of explicit evaluations, while lowly innovative designs decrease or remain constant (Carbon et al., 2006; Carbon & Leder, 2005). Whether innovativeness per se produces the effects or whether these effects depend on the set was not clear, so variation on this stimulus dimension was tested in the present study. We found that context, in terms of set homogeneity, strongly affected the appreciation of innovativeness. In a set of stimuli with heterogeneous innovativeness including highly and lowly innovative stimuli (Experiments 1 and 3), we replicated this known effect of innovativeness (see Figure 2). In contrast, when only one level of innovativeness – high or low innovativeness – was repeatedly evaluated (Experiment 2), then attractiveness ratings after repeated evaluations increased for both stimulus classes. Thus, innovativeness only affected attractiveness when both highly and lowly innovative stimuli were evaluated within one set. That is, when innovativeness was made apparent and was distinctive in the stimulus set. This was reflected in the innovativeness ratings and suggests that innovativeness was differentially apparent in the different stimulus sets. In Experiments 1 and 3, when heterogeneous sets were used, the innovativeness ratings between the two stimulus classes were clearly and significantly different. This was not the case in Experiment 2.

One might argue that the lack of differences in the attractiveness and innovativeness ratings for the differently innovative stimuli in Experiment 2 might have been due to participants using the scale differently: Participants might have used the whole scale for their ratings in Experiment 2 because they were judging the stimuli relative to each other (Helson, 1948; Parducci, 1995) or they might have felt obliged to use the whole scale for their judgments in order to provide information for the researcher (Schwarz, 1999). These explanations would be confirmed by higher standard deviations in Experiment 2 than in Experiments 1 and 3. However, as revealed in Table 1, the standard deviations were similar across experiments. Thus, we believe that the effects of innovativeness depend on the characteristics of the stimulus set being judged. Additionally, different dynamics with selectively stronger increases in attractiveness for highly innovative stimuli in Experiments 1 and 3 developed only in heterogeneous sets.

In all three experiments, innovativeness was also evaluated as a dimension during the first stimulus exposures. The differences in results therefore suggest that the effects of innovativeness on attractiveness require a distinctive variation in the stimulus set, not just the awareness that the dimension exists. The effects of innovativeness and attractiveness can also be seen in correlations between attractiveness and innovativeness for the highly innovative stimuli in T1. These (simple Pearson correlations) were $r = .48$ in Experiment 1 and $r = .46$ in Experiment 3, but only $r = .15$ in Experiment 2. These correlations are in accordance with the arousal theory (Berlyne, 1970a) as well as the two-factor theory (Berlyne, 1970b; Stang, 1974): Relative differences in arousal level (Berlyne, 1970a) or in boredom (Berlyne, 1970b; Stang, 1974) only affect attractiveness ratings and its dynamics when innovativeness is apparent. The attractiveness ratings for the lowly innovative stimuli at T2 in Experiments 1 and 3 are in accordance with an explanation based on boredom. The longer experiment duration and the higher number of ratings in Experiment 1 showed more boredom-like effects (O'Hanlon, 1981). In Experiment 3, attractiveness ratings for lowly innovative stimuli slightly increased toward T2 while in Experiment 1 their attractiveness slightly decreased. However, these interpretations should be further tested using psychophysiological measures that are sensitive to arousal and boredom, such as electro-dermal activity (Dawson, Schell, & Filion, 2000).

Conclusions

The present study demonstrated that the attractiveness and the dynamics of the attractiveness of innovativeness develop only when innovativeness is apparent through the stimulus set and distinct in the stimuli. A mere evaluation of innovativeness alone, as was explicitly asked for in all experiments in the first phase, did not trigger such specific changes in attractiveness. These findings are in accordance with theories emphasizing the relative nature of evaluations in general (e.g., Helson, 1948) and with theories emphasizing the relative nature of attractiveness evaluations (Berlyne, 1970a, 1970b; Stang 1974). From a basic research perspective, our findings emphasize the importance of explicitly considering stimulus dimensions, which affect evaluations. Moreover, they stress that evaluations are made in situations and context (Smith & Semin, 2004) and – at least as shown for innovativeness here – do not rely on internal and independent standards.

These findings also have implications for applied contexts, such as for testing the appreciation of innovative product designs that are to be introduced into the market. Our results suggest (1) that in order to create awareness of the appreciable aspects of innovativeness, innovative designs should be tested together with lowly innovative designs; and (2) that innovativeness profits when evaluated after a phase of repeated evaluations. Testing only once, as often done in one-shot marketing studies, may not capture the possible dynamics of attractiveness. This procedure of presenting heterogeneous stimulus sets and using repeated evaluations seems to be a good approximation of processes that occur under real-life exposure conditions. For example, when one sees an innovative car, one might automatically judge the car in relation to other cars on the streets. Moreover, one might repeatedly see the car in print and TV advertisements, and in person, and one might talk to friends about it. So, if you want an innovative product to be found attractive for its innovativeness, present it with its less innovative competitors!

Acknowledgments

This research was supported by a grant to Helmut Leder and Claus-Christian Carbon by the FWF "Fonds zur Förderung der wissenschaftlichen Forschung" (National Austrian Scientific Fund; P18910). We particularly thank the action editor, Fred Mast, and two anonymous reviewers for their helpful comments for improving this manuscript. We thank Christine Talker for designing the stimuli and Michael Forster for conducting parts of the experiments.

References

Berlyne, D. E. (1970a). *Esthetics and psychobiology*. New York: Appleton-Century-Crofts.

Berlyne, D. E. (1970b). Novelty, complexity, and hedonic value. *Perception and Psychophysics, 8*, 279–286.

Bornstein, R. F., & D'Agostini, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional account. *Social Cognition, 12*, 103–128.

Bornstein, R. F., Kale, A. R., & Cornell, K. R. (1990). Boredom as a limiting condition on the mere exposure effect. *Journal of Personality and Social Psychology, 58*, 791–800. doi: 10.1037/0022-3514.58.5.791

Cammann, R. (1990). Is there no MMN in the visual modality? *Behavioral and Brain Sciences, 13*, 234–234.

Carbon, C. C. (2010). The cycle of preference: Long-term dynamics of esthetic appreciation. *Acta Psychologica, 134*, 233–244. doi: 10.1016/j.actpsy.2010.02.004

Carbon, C. C., Grüter, T., Grüter, M., Weber, J. E., & Lüschow, A. (2010). Dissociation of facial attractiveness and distinctiveness processing in congenital prosopagnosia. *Visual Cognition, 18*, 641–654. doi: 10.1080/13506280903462471

Carbon, C. C., Hutzler, F., & Minge, M. (2006). Innovativeness in design investigated by eye movements and pupillometry. *Psychology Science, 48*, 173–186.

Carbon, C. C., & Leder, H. (2005). The repeated evaluation technique (RET): A method to capture dynamic effects of innovativeness and attractiveness. *Applied Cognitive Psychology, 19*, 587–601. doi: 10.1002/acp.1098

Carbon, C. C., Michael, L., & Leder, H. (2008). Design evaluation by combination of repeated evaluation technique and measurement of electrodermal activity. *Research in Engineering Design, 19*, 143–149. doi: 10.1007/s00163-008-0045-2

Carbon, C. C., & Talker, C. (2006). *Evaluating design perception of 36 car interiors*. Unpublished research report, University of Vienna, Vienna, Austria.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271.

Cooper, R. G. (2001). *Winning at new products*. Reading, MA: Addison-Wesley.

Cox, G. (2005). *The Cox review of creativity in business, HM Treasury*. Retrieved from http://www.hm-treasury.gov.uk/coxreview_index.htm

Dawson, M. E., Schell, M. S., & Filion, D. L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 200–222). Cambridge, UK: Cambridge University Press.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*, 170–177. doi: 10.1037/1082-989X.1.2.170

Faerber, S. J., Leder, H., Gerger, G., & Carbon, C. C. (2010). Priming semantic concepts affects the dynamics of esthetic appreciation. *Acta Psychologica, 135,* 191–200. doi: 10.1016/j.actpsy.2010.06.006

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. doi: 10.3758/BRM.41.4.1149

Halberstadt, J. (2006). The generality and ultimate origins of the attractiveness of prototypes. *Personality and Social Psychology Review, 10*, 166–183. doi:10.1207/s15327957pspr1002_5

Halberstadt, J., & Rhodes, G. (2003). It's not just average faces that are attractive: Computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychonomic Bulletin and Review, 10*(1), 149–156.

Hekkert, P., & Leder, H. (2008). Product esthetics. In H. Schifferstein & P. Hekkert (Eds.), *Product experience* (pp. 259–286). Amsterdam, Netherlands: Elsevier.

Hekkert, P., Snelders, D., & Van Wieringen, P. C. W. (2003). Most advanced, yet acceptable: Typicality and novelty as joint predictors of esthetic preference in industrial design. *British Journal of Psychology, 94*, 111–124. doi: 10.1348/000712603762842147

Helson, H. (1948). Adaption level as a basis for a quantitative theory of frames of reference. *Psychological Review, 55*, 297–313.

Leder, H., Augustin, D., & Belke, B. (2005). Art and cognition! Consequences for experimental esthetics. *Bulletin of Psychology and the Arts, 5*, 11–20.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of esthetic appreciation and esthetic judgments. *British Journal of Psychology, 95*, 489–508. doi: 10.1348/0007126042369811

Leder, H., & Carbon, C. C. (2005). Dimensions in appreciation of car interior design. *Applied Cognitive Psychology, 19*, 603–618. doi: 10.1002/acp.1088

Lee, A. Y. (2001). The mere exposure effect: An uncertainty reduction explanation revisited. *Personality and Social Psychology Bulletin, 27*, 1255–1266. doi: 10.1177/01461672012710002

Liu, Y. (2003). Engineering esthetics and esthetic ergonomics: Theoretical foundations and a dual-process research methodology. *Ergonomics, 46*, 1273–1292. doi: 10.1080/00140130310001610829

Mairesse, J., & Mohnen, P. (2002). Accounting for innovativeness and measuring innovativeness: An illustrative framework and an application. *American Economic Review, 92*, 226–230. doi: 10.1257/000282802320189302

Moulson, T., & Sproles, G. (2000). Styling strategy. *Business Horizons, 43*, 45–52. doi: 10.1016/S0007-6813(00)80008-7

O'Hanlon, J. F. (1981). Boredom: Practical consequences and a theory. *Acta Psychologica, 49*, 53–82. doi: 10.1016/0001-6918(81)90033-0

Parducci, A. (1995). *Happiness, pleasure, and judgment*. Hillsdale, NJ: Erlbaum.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and esthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*, 364–382. doi: 10.1207/s15327957pspr0804_3

Rhodes, G., Jeffery, L., Watson, T., Clifford, C., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science, 14*, 558–566. doi: 10.1046/j.0956-7976.2003.psci_1465.x

Robinson, B. M., & Elias, L. J. (2005). Novel stimuli are negative stimuli: Evidence that negative affect is reduced in the mere exposure effect. *Perceptual and Motor Skills, 100*, 365–372. doi: 10.2466/PMS.100.2.365

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*, 93–105. doi: 10.1037//0003-066X.54.2.93

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25*, 638–656. doi: 10.1521/soco.2007.25.5.638

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology, 36*, 53–117. doi: 10.1016/S0065-2601(04)36002-8

Stang, D. J. (1974). Methodological factors in mere exposure research. *Psychological Bulletin, 81*, 1014–1025. doi: 10.1037/h0037419

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113*, 345–361. doi: 10.1037/0033-2909.113.2.345

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by preattentive sensory memory. *Nature, 372*, 90–92. doi:10.1038/372090a0

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychology Science, 10*, 224–228. doi: 10.1111/1467-8721.00154

Gernot Gerger

University of Vienna
Faculty of Psychology
Liebiggasse 5
1010 Vienna
Austria
gernot.gerger@univie.ac.at