Innovativeness in design investigated by eye movements and pupillometry

CLAUS-CHRISTIAN CARBON¹, FLORIAN HUTZLER² & MICHAEL MINGE²

Abstract

Innovative designs often break common visual habits by integrating more or less familiar parts into a new concept (Leder & Carbon, 2005). When design innovation is realized in an overly advanced way, the resulting designs tend to be rejected by perceivers at first glance, but seem to be favored when perceivers become more familiar with them (Carbon & Leder, 2005b). In the present study, we investigated the properties of innovative car interior designs by analyzing eye movements and the dilatation of the pupil when evaluating these designs. The analysis of eye movements indicates that innovative designs may be interpreted as more balanced in their conceptual structure. This was shown by an increased number of eye movements directed at the focus areas in car designs. Moreover, pupillometry data demonstrated that innovative designs are cognitively more demanding, thus evoking more interest. Such effects of innovativeness were particularly strong after participants had been exposed to and had elaborately dealt with the material during a phase of familiarization realized by the repeated evaluation technique.

Key words: design, innovation, innovativeness, eye tracking, pupillometry, adaptation, attractiveness, innovativeness, repeated evaluation technique (RET)

¹ Institute of Psychological Basic Research, Faculty of Psychology, University of Vienna, Austria.
² Institute of Cognitive and General Psychology, Freie Universität Berlin, Germany.
Innovativeness in design investigated by eye movements and pupillometry

Understanding innovation in design is an important challenge for designers who are looking for innovative but also widely accepted appearances, as well as for psychologically motivated researchers who want to understand the underlying constructs and processes that are related to innovation (Leder & Carbon, 2005). Furthermore, understanding innovation is also the bottom line for applied research and production companies that are interested in optimizing their products in terms of maximizing their return of investment: “Innovation is key to firms’ future” (van Heerde, Mela, & Manchanda, 2004, p.178).

Common measures

It is a common procedure that products or new design are evaluated before submission to the public market. This standard procedure is used to minimize design flops which are rejected by the consumers. Normally, this procedure is realized by inviting typical consumers or design experts to evaluate the attractiveness or liking by means of questionnaires or association tests. Moreover, it is a common practice to measure psychological concepts of attractiveness and innovativeness only once due to the assumption that a design is relatively stable in terms of its attractiveness. However, this assumption seems to be invalid, because effects of fashion, Zeitgeist, and culture strikingly demonstrate rather dynamic effects of attractiveness. Moreover, as demonstrated by Carbon and Leder (2005b), innovation also has a dynamic influence on attractiveness over time. Therefore, a single measurement is susceptible for invalid predictions of future liking and market success. Based on these findings, Carbon and Leder (2005b) have propagated a more adequate measurement which (a) measures attractiveness at least twice, and (b) simulates and challenges innovativeness to develop its dynamic nature. This is done by separating two test phases by an evaluation phase, in which participants are not only massively exposed to the experimental stimuli but are also asked to evaluate them by means of many different attributes. This so-called repeated evaluation technique (RET) provides the opportunity to measure innovativeness and the impact on attractiveness in a more dynamic and contextual sense.

Measurements of the eyes

The aim of the present study is to supplement findings of the dynamic aspect of innovativeness in design with additional measures, such as eye movements and pupillometry (measures on the dilatation of the pupil). Recent research has revealed that good designs are easier to look at in the sense that diverse exploration leads to global grasp of the visual composition without becoming entangled (Nodine & Krupinski, 2004). This is in accordance with Berlyne’s (1971) psychoaesthetic approach, which stated that for good compositions

---

3 In the present paper we will use the term design in terms of styling, thus referring to the appearance of the design how it is used in common speech English.
Innovativeness in design

Perceivers are more likely to reach a stage of “ergodic equilibrium”, a term introduced by Molnar (1968), in less amount of time. They are capable of detecting the meaning of the composition faster! Berlyne (1971) explained the higher liking value of such compositions by the assumption that by means of this given equilibrium, perceivers are allowed to scan for provided pleasure very soon. Moreover, Nodine and Krupinski (2004) proposed that a good composition is one that creates tension between different levels of information, for example what Nodine and Krupinski (2004) call global patterns and focal details. The ongoing tension between these different levels of information supports long-term interest. We think that innovative material has the potential to generate such tensions between different components; which in turn could be the reason that such material benefits from elaboration by higher attractiveness ratings.

Pupillometry is the study of the dilatation of the pupil. Previous research suggests that the pupil may be a measure as to how much attention is required for a certain task. For example, researchers have consistently found that people’s pupils are larger when they are attempting to solve more difficult problems, versus when they are dealing with less challenging problems. Thus, more demanding processes are susceptible for generating more dilated pupil dilatation (Beatty, 1982; Beatty & Kahneman, 1966; Ben Nun, 1986). Moreover, as pointed out by Granholm and Steinhauer (2004), the extent of pupil dilatation evoked by a stimulus reflects activation of specific modules that process the presented material. Interestingly, since the iris of the pupil operates autonomically, pupillometry is a powerful measure for autonomic responses with a meaningful temporal and fine-graded spatial resolution (see Nodine & Krupinski, 2004).

The present study

The aim of the present study was to investigate the nature of innovativeness in design, particularly as a predictor for attractiveness. Therefore, we used car interior designs which systematically varied in terms of innovativeness based on material that was created and evaluated in an earlier study (Leder & Carbon, 2005). In order to circumvent typical problems of the concept innovativeness, which seems to have a dynamic nature (in terms of differential effects on attractiveness depending on the degree of elaboration of the material), we used the repeated evaluation technique (RET) propagated by Carbon and Leder (2005b). Furthermore, we recorded eye movements and pupil dilatation during evaluating the designs.

---

4 An ‘ergodic equilibrium’ is defined by a concentration of fixations to compositional centers of interest to which the eye returns again and again in relatively short periods. Thus, this idea concentrates on the behavior of the beholder and not to physical attributes such as the balance of major regions of interest in the composition.

5 According to Nodine and Krupinski (2004), global patterns are accomplished by short gazes, interpreted as a kind of divergent reflection which gain an appreciation of composition and the overall character of the artwork; in contrast, focal details are grasped by long gazes, which reflect specific exploration in order to fit depicted objects into a meaningful visual concept.
Experiment

Method

Participants

Sixteen participants with an average age of 25.8 years took part in the experiment; ten of which were males. Participants were naïve to the experimental hypotheses and received a compensation of 6 Euro for participation. All participants had normal or corrected to normal vision and were tested individually.

Apparatus and stimuli

The stimulus material consisted of 8 drawings of car interiors (see Figure 1) that were presented on a 17-inch monitor with a size of 1000 x 750 pixels at a screen resolution of 1024 x 768 pixels. The car interiors varied according to the dimensions of Curvature (straight, curved), Innovativeness (low innovative, highly innovative), and Complexity (low complex, highly complex), which were fully balanced. The dimensions of the material are based on the stimuli used in Leder and Carbon (2005), which were evaluated regarding these three dimensions with 72 participants (see for details, Leder & Carbon, 2005) and that were found to be fairly independent from each other. The eye movements and the diameter of the left pupil were measured by an iView™ X RED-II-camera with a resolution of 50 Hz. The data was analyzed by the software program Begaze™.

Procedure

The procedure was very similar to that of the original study investigating effects of innovativeness on attractiveness of car interiors (Carbon & Leder, 2005b). In an initial phase, the participants rated eight stimuli separately on scales of attractiveness and innovativeness (test phase 1: T1). All ratings were indicated on a 7-point-Likert scale (from ‘1’: least significant, up to ‘7’: most significant). In this phase we also measured the eye movements and the pupil diameters of the participants while they were looking at the stimuli without any further instructions. We used the ratings and the data of eye tracking as base-rates. Subsequently, an extended rating phase followed. This phase, which consisted of 25 rating blocks6, is referred to as the RET (repeated evaluation technique) phase (Carbon & Leder, 2005b). Participants were instructed to rate the same stimuli as in T1 on several dimensions on “yes” or “no”

Innovativeness in design

<table>
<thead>
<tr>
<th></th>
<th>highly complex</th>
<th>low complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>straight</td>
<td>curved</td>
</tr>
<tr>
<td><strong>low innovative</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>highly innovative</strong></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Figure 1: Stimulus material used in the experiment. The top row depicts low innovative, the bottom row highly innovative designs. Both levels of design were fully balanced with the factors Complexity (high vs. low) and Curvature (straight vs. curved).
scales. After all RET ratings were given, there was a short break during which the participants were instructed to answer two final ratings as deliberately as possible, followed by the second rating phase for attractiveness and innovativeness (test phase 2: T2). In T2, eye movements and pupil diameters were measured a second time. The order of the stimuli was fully randomized for each rating block. Moreover, the order of the rating blocks with attractiveness ratings in the first place and innovativeness ratings in the second place was constant. All stimuli were presented for 8 s yielding approximately 30 minutes for each session.

**Results and discussion**

In the following, behavioral data (ratings of attractiveness and innovativeness), eye movement patterns and pupil dilatation data will be reported.

**Attractiveness ratings**

Mean ratings for attractiveness evaluations for each participant were submitted to a four-way repeated measurement ANOVA with Phase (T1, T2), Curvature (curved, straight), Complexity (low complex, highly complex), and Innovativeness (low innovative, highly innovative) as within-subjects factors (see raw data in Table 1).

The analysis revealed significant main effects of Curvature, $F(1, 15) = 22.23, p < .0001$, $\eta^2_p = .597$, Complexity, $F(1, 15) = 5.96, p < .05$, $\eta^2_p = .284$, and Innovativeness, $F(1, 15) = 4.59, p < .05$, $\eta^2_p = .234$. Moreover, and most interestingly, there was a two-way interaction between Phase and Innovativeness, $F(1, 15) = 5.11, p < .05$, $\eta^2_p = .254$ (see Figure 2), which indicated a modulating influence of innovativeness on attractiveness over time.

As in the original work by Carbon and Leder (2005b), the attractiveness of low innovative material decreased after exposure and extensive evaluation during the RET phase, an effect that was indicated by a main effect of Phase for the low innovative material, $M_{T1} = 4.28, M_{T2} = 3.59, F(1, 15) = 6.85, p < .02$, $\eta^2_p = .313$. However, the increase of attractiveness

**Table 1:**

Attractiveness ratings in test phase T1 and T2, displayed separately for the factors innovativeness, complexity and curvature (Attr = Attractiveness, Innov = Innovativeness).

<table>
<thead>
<tr>
<th></th>
<th>Straight Attr</th>
<th>Straight Innov</th>
<th>Curved Attr</th>
<th>Curved Innov</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Test phase T1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low innov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complex</td>
<td>3.13</td>
<td>1.36</td>
<td>4.31</td>
<td>1.49</td>
</tr>
<tr>
<td>Highly complex</td>
<td>4.63</td>
<td>1.26</td>
<td>5.06</td>
<td>1.34</td>
</tr>
<tr>
<td>Highly innov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complex</td>
<td>2.25</td>
<td>1.29</td>
<td>3.63</td>
<td>1.26</td>
</tr>
<tr>
<td>Highly complex</td>
<td>2.81</td>
<td>1.68</td>
<td>3.81</td>
<td>1.22</td>
</tr>
<tr>
<td><strong>Test phase T2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low innov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complex</td>
<td>2.94</td>
<td>1.34</td>
<td>3.81</td>
<td>1.60</td>
</tr>
<tr>
<td>Highly complex</td>
<td>3.13</td>
<td>1.78</td>
<td>4.50</td>
<td>1.32</td>
</tr>
<tr>
<td>Highly innov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complex</td>
<td>2.63</td>
<td>1.45</td>
<td>4.13</td>
<td>1.75</td>
</tr>
<tr>
<td>Highly complex</td>
<td>2.56</td>
<td>1.41</td>
<td>4.50</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Innovativeness in design

Mean ratings of attractiveness (left) and innovativeness (right) separately for Phase (test phase T1 and T2) and Innovativeness (low innovative, highly innovative).

for the highly innovative material was only demonstrable numerically but did not reach a level of significance, $M_{T1} = 3.13, M_{T2} = 3.45, F(1, 15) = 1.79, p = .20, \text{n.s.}$

Innovativeness ratings

Mean ratings for innovativeness evaluations for each participant were submitted to a four-way repeated measurement ANOVA with Phase (T1, T2), Curvature (curved, straight), Complexity (low complex, highly complex), and Innovativeness (low innovative, highly innovative) as within-subjects factors. The analysis revealed significant main effects of Curvature, $F(1, 15) = 6.41, p < .05, \eta^2_p = .299$, Complexity, $F(1, 15) = 6.95, p < .02, \eta^2_p = .317$, and Innovativeness, $F(1, 15) = 10.28, p < .01, \eta^2_p = .407$. Moreover, there was a two-way interaction between Phase and Curvature, $F(1, 15) = 12.87, p < .01, \eta^2_p = .462$, and a three-way interaction between Curvature, Complexity, and Innovativeness, $F(1, 15) = 5.75, p < .05, \eta^2_p = .277$.

The innovativeness ratings also replicated the original findings by Carbon and Leder (2005b): The massive evaluation phase between T1 and T2 did not lower the general innovativeness rating of the material, neither of the low innovative nor of the highly innovative version.

Eye movements

For the analysis of eye movement patterns, four regions of interest were defined: Steering wheel, console, mirror and right (see Figure 3). These regions were empirically determined, as they corresponded with more than 90% of all fixations.
In a first step, the number of fixations exhibited in the four regions of interest was analyzed. As shown in Table 2, a high number of fixations were only observed for two regions (i.e., steering wheel and console). Only these two main focus regions were included in further analyses.

First, the overall number of fixations and the mean fixation duration exhibited on these main focus regions were analyzed, subsequently, it was examined, which region of interest was fixated first. All eye tracking measures were collapsed over the attractiveness and innovativeness blocks.

The overall number of fixations was analyzed by a three-way repeated measurement ANOVA with Phase (T1, T2), Area (Console, Steering wheel) and Innovativeness (low innovative, highly innovative) as within-subjects factors. The analysis revealed no significant main effects, but a two-way interaction between Area and Innovativeness, $F(1, 14) = 6.47, p < .05, \eta_p^2 = .316$ (see Figure 4): for the highly innovative material, there was a greater balance of the number of fixations between the console area and the steering wheel area. The mean fixation duration was analyzed by an ANOVA structured in the same way as that performed for the number of fixations. For this analysis none of the main effects and no interactions yielded significant results, $F$s($1, 15) < 1.70, p > .21, n.s.

Table 2:

Overall number of fixations in the pre-defined regions of interest at test phase T1 and T2 displayed separately for highly and low innovative designs.

<table>
<thead>
<tr>
<th></th>
<th>Steering wheel</th>
<th>Console</th>
<th>Mirror</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>Low Innovative</td>
<td>119</td>
<td>125</td>
<td>121</td>
<td>118</td>
</tr>
<tr>
<td>Highly Innovative</td>
<td>116</td>
<td>125</td>
<td>118</td>
<td>118</td>
</tr>
</tbody>
</table>
Finally, we analyzed the probability of first fixations directed to the focused areas. Not surprisingly, the regions console and steering wheel received most of the first fixations (84.38 % of all possible regions). A three-way repeated measurement ANOVA with Phase (T1 vs. T2), Area (Console vs. Steering wheel) and Innovativeness (low vs. highly innovative) as within-subject factors was conducted. The analysis revealed significant main effects of Phase, $F(1, 15) = 5.00, p < .05, \eta^2 = .250,$ and Innovativeness, $F(1, 15) = 5.00, p < .05, \eta^2 = .250.$ Moreover, a trend for an interaction between Area and Innovativeness was observed, $F(1, 15) = 3.46, p = .083, \eta^2 = .188, n.s.$ Highly innovative material had a higher probability of a first fixation than low innovative material ($M = 43.8\%$ vs. $40.6\%).$ Moreover the first looks at the focused areas increased from T1 to T2 ($M = 39.1\%$ vs. $45.3\%).$ Concerning the trend of an interaction between Area and Innovativeness, highly innovative stimuli once again elicited more balanced eye movement patterns: Both of the main focus areas (steering wheel and console) had approximately the same probability of receiving a first fixation, a trend that could not be observed for low innovative material (see Figure 5).
For illustration purposes, two typical (individual) scanpaths (for low and highly innovative car designs, recorded during T2) are depicted in Figure 6.

Figure 6:
Two typical (individual) scanpaths for a low innovative design (upper picture) and a highly innovative design (lower picture): the scanpaths for the low innovative design seem to be less balanced with more frequent fixations to peripheral spots, whereas the scanpaths for the highly innovative design demonstrates more balanced and dense eye movements between regions of interest.
Innovativeness in design

Pupillometry

Pupillometry (i.e., the dilatation of the pupil) was assessed by the averaged horizontal diameters of the pupil of the left and the right eye (in 1/1000 cm). The data was analyzed by a four-way repeated measurement ANOVA with Phase (T1, T2), Curvature (curved, straight), Complexity (low complex, highly complex), and Innovativeness (low innovative, highly innovative) as within-subjects factors. The analysis revealed a significant main effect of Complexity, $F(1, 15) = 9.70, p < .01, \eta_p^2 = .393$, and a two-way interaction between Phase and Complexity, $F(1, 15) = 10.30, p < .01, \eta_p^2 = .407$, and a trend for an interaction between Phase and Innovativeness, $F(1, 15) = 3.87, p = .068, \eta_p^2 = .205, n.s$. As shown in Figure 7, a modulating effect of time was observed for the complexity and, as a trend, for the innovativeness factor: The more complex/innovative a design was, the greater was the increase of pupil diameter over time.

General discussion

We investigated the role of innovativeness in design with low innovative and highly innovative car design interiors over time. Based on the idea that innovative designs break common visual habits and, therefore, are initially rejected by the perceivers – but will be favored after a period of familiarization and elaboration – we used the repeated evaluation technique (RET, Carbon & Leder, 2005b). By means of the RET, we were able to demonstrate typical effects of innovative material with an interaction between Phase and Innovativeness, with a trend for low attractiveness scores in an initial rating and high scores in the final test phase. In contrast, a reverse trend (a decrease of initially high attractiveness ratings) was found for the low innovative material (see Carbon & Leder, 2005b).
Moreover, the number of fixations exhibited on main focus areas aided in analyzing eye movement patterns and revealed a more balanced pattern of fixations for highly innovative material. When highly innovative designs were presented, participants fixated the console nearly as often as the steering wheel. This pattern of results held also true, when only the initial fixations of a presentation were analyzed. Then, again, a greater balance of fixations was found between the two main focus areas for the highly innovative material. This might be an indication that innovative car interior designs (represented by shade line drawings in the present paper) had a higher degree of visual rightness (Locher, 2003) and thus created more tension and a longer termed interest for the design (Nodine & Krupinski, 2004).

Furthermore, when analyzing the participants’ behavior by means of pupillometry, a greater increase of the pupil was found when participants were exposed to highly innovative or highly complex material. As the dilatation of the pupil is positively related to task demands, attention and activation, it seems plausible that highly innovative material generated more interest and cognitive demands. This could be the reason why highly innovative material was initially rejected, but benefited from deeper elaboration in terms of attractiveness. In the initial phase, such designs seem to break common visual habits: They have a rather unfamiliar appearance and thus were relatively untypical for car interiors. Participants tend to reject these designs, because they are too advanced and therefore not acceptable (Hekkert, Snelders, & van Wieringen, 2003). After prolonged exposure to and evaluation of these designs during the RET phase, innovative designs did not only become more familiar, but also cognitively fluent (Leder, 2003). We assume that under these conditions, participants were now able to explore the innovative structure of the material. A design that demands cognitively sophisticated processing appears to elicit more interest than a low innovative design that is indeed rather prototypical and familiar, and at the same time boring and less demanding. Therefore, (highly) innovative designs are likely candidates to become desirable and attractive designs in the future. However, the relation between innovativeness and market success is by far deterministic, but certainly follows other important factors, too, such as: market strategies, the Zeitgeist, and fashion trends.

The differential results from several measures like ratings, eye movements and pupil dilatation obtained in the present study for different levels of innovativeness encourage further investigations of art and design appreciation by such measures. It also seems worthwhile for future studies to analyze the eyetracks between the first and the second test phase according to the scanpath theory by Noton and Stark (e.g., Noton & Stark, 1971a, 1971b, 1971c) that postulates fixed scanpaths during initial viewing of a pattern which were found to be highly similar in a second, recognition phase. Moreover, we think that the specific combination of these measures with alternative methods such as adaptation measures (Carbon & Leder, 2005a), electro-dermal measures (Roy, Bouchein, Fowles, & Gruzelier, 1993), EEG/MEG (Cela-Conde et al., 2004) or fMRI (Vartanian & Goel, 2004) is most promising for future research. Moreover, further studies might also combine these technically more advanced methods with traditional measures, assessing the perceivers’ thoughts during the exploration of artworks (Locher, 2003). Most likely, only such a combination of multivariate methods of analysis is suited for capturing the complex nature of innovativeness in design and aesthetics.

To summarize, effects of innovativeness in design are not only demonstrable in initially low ratings of attractiveness that increase after extensive (repeated) exposure and evaluation
– we were also able to trace effects of innovativeness in eye movement patterns and pupil dilatation.

References

Author note

We thank Andrea Lyman and Melissa Võ for proof-reading this manuscript. Moreover, we are glad to thank Helmut Leder and an anonymous reviewer for valuable comments on a previous version of this manuscript. Correspondence concerning this manuscript should be addressed to CCC at the University of Vienna, Faculty of Psychology, Department of Psychological Basic Research, Liebiggasse 5, A-1010 Vienna, Austria, phone: +43-(0)1-4277-47921, email: ccc@experimental-psychology.com