DRS2016
Design + Research + Society
Future-Focused Thinking

2016 Design Research Society
50th Anniversary Conference
27-30 June 2016, Brighton, UK

Measuring design typicality – a comparison of objective and subjective approaches

Stefan Mayer[a], Jan R. Landwehr[a]

[a]Goethe University, Germany
*Corresponding author e-mail: smayer@wiwi.uni-frankfurt.de

Abstract: Design typicality plays a major role in consumers' reactions towards a product. Hence, assessing a product design's typicality is vital to predicting consumers' responses to a design. However, directly asking people for their subjective typicality experience may yield a biased measure as the rating arguably contains the overall aesthetic impression of the product. Against this background, we introduce four unbiased objective measures of design typicality (two based on feature points and two based on grids) and demonstrate their capability of capturing the subjective typicality experience. We validate the proposed measures in the context of automobile designs with ratings of aesthetic liking, processing fluency, and cumulative sales data by analysing 77 car models from four segments ranging from subcompact cars to SUVs. Our findings endorse the general notion that objective measures should be included in product design research; and the proposed objective approaches provide convenient means to easily assess design typicality.

Keywords: car design; aesthetic liking; design typicality; processing fluency

1. Introduction

Visual typicality is a general principle of product design aesthetics, and ample research suggests that humans prefer prototypical designs over atypical ones (e.g., Hekkert, Snelders, & van Wieringen, 2003; Veryzer & Hutchinson, 1998). The link between aesthetic preferences and typicality has been demonstrated in various domains such as faces (Langlois & Roggman, 1990), visual patterns (Martindale & Moore, 1988; Winkielman, Halberstadt, Fazendeiro, & Catty, 2006), paintings (Purcell, 1993), and automobile designs (Landwehr, Labroo, & Herrmann, 2011; Landwehr, Wentzel, & Herrmann, 2013). Typicality also plays an important role in the trendiness of a product (Blijlevens, Mugge, Ye, & Schoormans, 2013) and in the evaluation of fit between product and the context of product presentation (Blijlevens, Gemser, & Mugge, 2012).

However, studies establishing associations between typicality and aesthetic preferences have been criticized based on potential circularity in their relation (Boselie, 1995). In particular, if participants are asked to rate the most typical products they may simply choose the most beautiful ones (Hekkert et al., 2003). Hence, if one is interested in the true typicality of a product, directly asking people about their typicality perception may yield a biased measure because the answer is partly driven by the overall aesthetic impression, as it captures the response towards the stimulus instead of the typicality of the stimulus itself. An unbiased measure of design typicality, in contrast, would optimally capture solely the stimulus' typicality.

The purpose of this research is to introduce such objective, unbiased measures of design typicality and to compare them against subjective typicality ratings. We show that objective measures are able to capture the subjective typicality experience. Furthermore, we validate the proposed measures in the context of automobile designs from four segments, ranging from subcompact cars to SUVs, using ratings of aesthetic liking, processing fluency, and cumulative sales data. In doings so, we contribute to the field of product design research by providing algorithmic, objective approaches to assess design typicality that can be used from the very first step of the design process.

2. Theoretical background

When people are forming first impressions, they prefer prototypical stimuli. This "beauty-in-averageness" effect was initially reported for facial processing where Langlois and Roggman (1990) found that people prefer an average face that was created as a morph of all faces in a sample over any individual face in the sample. As an explanation for this preference served the reproductive fitness that is associated with prototypicality (Rhodes & Tremewan, 1996). However, subsequent studies demonstrated that this effect also holds true for stimuli that are not associated with a reproductive advantage. Halberstadt and Rhodes (2000), for example, showed that prototypical drawings of dogs and birds and photos of wristwatches were perceived as more attractive than less typical ones. This relationship between typicality and attractiveness is also reflected in ratings of aesthetic liking, defined as "the sensation that results from the perception of attractiveness (or unattractiveness) in products" (Crilly, Moultrie, & Clarkson, 2004, p. 552). Landwehr and colleagues (2013), for instance, found higher ratings of aesthetic liking for typical car designs at low level of exposure. They further showed that this positive effect of typicality also affects car sales and that the underlying psychological mechanism is processing fluency.

One of the most influential explanations for a positive effect of typicality is based on the concept of processing fluency (Reber, Schwarz, & Winkielman, 2004; Winkielman et al., 2006). Fluency refers to the cognitive ease that people experience when processing a stimulus (Schwarz, 2004). A key prediction of processing fluency theory is that the ease of processing a stimulus determines the aesthetic response towards the stimulus (Reber et al., 2004) because higher fluency and hence easier processing is inherently positive and experienced as gut-level positive affect (Winkielman & Cacioppo, 2001). Provided that the positive feeling is not attributed to a different source, higher fluency will thus result in higher aesthetic liking (Winkielman, Schwarz, Fazendeiro, & Reber, 2003).

Importantly, research on fluency has shown that the experience of fluency can be determined by core characteristics of a stimulus that are nonspecific to its content (Reber et al., 2004). People can, for example, process stimuli that are high in symmetry, figure-ground contrast, and visual clarity more efficiently, resulting in feelings of fluency (Reber et al., 2004). Research linking typicality to fluency suggests that prototypical stimuli result in higher fluency because typical designs enable a quicker and more efficient visual categorization of the product due to the similarity to the visual traces stored in people's visual memory. This has been demonstrated, for example, by Winkielman and colleagues (2006) using random dot patterns. They found that prototypical patterns are classified more efficiently and recruit fewer neural resources.

Against this background, we examine the relationship among design typicality, processing fluency, and aesthetic liking of consumers using 3D models of cars. In accordance with previous studies on car models, we argue that consumers will experience greater fluency when processing typical rather than atypical designs and will interpret the fluency signal as their aesthetic liking.

3. Objective measurement of design typicality

In order to capture all possible aspects of typicality, we compared four objective measures of design typicality. In line with previous research on prototypicality (Langlois & Roggman, 1990), we consider a prototype as possessing the average values of the visual features of a specific category. Thus, all four measures follow the general idea that the prototype is a representation of the common characteristics of all designs within a category. Therefore, a car's design is typical if the distance to the prototype's design is low, and atypical if the distance to the prototype's design is high. The specific calculation of the prototype and the distance measure, however, differs between the four measures. Two are based on feature points, two are based on a grid that is placed over the image.

3.1 Feature point-based measures

The two feature point-based measures rely on the identification of characteristic design features. The first approach applies an established measure that has been used in prior studies on the typicality of car designs (Landwehr et al., 2011, 2013) and is based on techniques originally developed in the context of research on facial attractiveness (Langlois & Roggman, 1990). In particular, predefined characteristic feature points (e.g., vertex of headlights) are manually set for each image, followed by a visually averaged representation (i.e., a morph) of all products within a product category. The objective typicality measure is created based on the Euclidean distances between each of the feature points of a particular car and the corresponding feature points of the morph. We use this approach (subsequently referred to as *manually coded feature point measure*) as the first of four objective design typicality measures.

The second approach follows the same idea of characteristic feature points. However, in contrast to manually coding the feature points, we use a new, algorithmic approach based on perceptual image hashing (Monga & Evans, 2006). The advantage of this new approach is that feature points do not have to be set manually but are found automatically based on corners and high curvature points using wavelet based feature detection. This particular way of automatically finding feature points enables us to quickly analyse even large sets of product designs. Once the feature points are found, the distance between the identified sets of feature points of different stimuli is calculated using a Hausdorff-like distance measure (Huttenlocher, Klanderman, & Rucklidge, 1993). We subsequently call this approach *algorithmic coded feature point measure*.

3.2 Grid-based measures

The two grid-based measures are computationally very simple and do not rely on characteristic feature points, hence they are also applicable to less standardized shapes. In both approaches, we rely on the key idea of perceptual image hashing to reduce an image to its perceptually relevant parts. In particular, we first partition the image into grid cells and calculate the average grey value within each grid. We repeat this procedure for all images and create average grid cells over all images to construe a morphed grid prototype (i.e., a cross-blended image). The typicality of a particular image is then simply the correlation of the greyscale values of its grid with the greyscale values of the morph's grid.

We use two versions of the above grid approach. For the first version, we vary the grid size from a 2x2 grid up to the full pixel information to capture all levels of visual typicality (coarse structures, finer details). Typicality is calculated as the average across the typicality values over all grids (subsequently referred to as *grid measure*). For the second version, we only use the full pixel resolution and calculate typicality simply as the correlation of the greyscale image with the mean image (subsequently referred to as *grid measure at full resolution*).

4. Study setup

We apply the proposed measures to a database consisting of automobile designs from four segments ranging from subcompact cars to SUVs, and compare them to subjectively rated design typicality. As outcomes, we use ratings of aesthetic liking, processing fluency, and cumulative sales.

4.1 Car model database

We use standardized images from a database of greyscale 3D car models as stimulus material. Overall, the database includes 77 cars from 4 segments. In particular, the stimuli consisted of 17 images of subcompact car models, 26 images of compact car models, 17 images of mid-size car models, and 17 images of SUV models. The different objective measurements of typicality were applied to the images at their original resolution without any modifications.

The database further includes subjective ratings of aesthetic liking, subjective fluency, and subjective typicality of the car models based on a sample of 365 people. The ratings were accessible at the level of the car models (averaged over all individual ratings) as well as at the individual level of a person.

4.2 Sales data

Twelve months (January–December 2013) of officially recorded car sales registration data were obtained from the German Federal Motor Transport Authority (Kraftfahrt Bundesamt, KBA). Overall, 1,409,412 cars were sold in the subcompact car, compact car, mid-size car, and SUV segments in the German market in 2013. Of the 77 3D car models, 58 had the exact design of the cars sold in Germany in 2013. Sales data of these 58 models cover 75.93% of all sales within the four segments in 2013 (subcompact cars: 63.87%, compact cars: 83.89%, mid-size cars: 89.77%, SUVs: 54.19%).

5. Results

All measured variables are z-standardized per car segment to exclude between segment variance from further analyses, since we are only interested in the general relationship between visual typicality and measures of preferences instead of absolute differences between different car segments. To make the subjective typicality measure comparable to the objective typicality measures, we used the aggregated subjective evaluations per stimulus for all analyses (i.e., the image-wise means across all evaluations), unless noted otherwise.

5.1 Correlation of typicality measures

Figure 1 shows the correlation matrix heatmap of all five typicality measures for the 77 car models. Except from the manually coded feature point measure, all objective typicality measures correlate significantly with subjectively rated design typicality scores. The grid measure at full resolution has the highest correlation with the subjective typicality measure. Both the feature point measures as well as the grid measure are not correlated to each other and seem to capture different aspects of typicality.

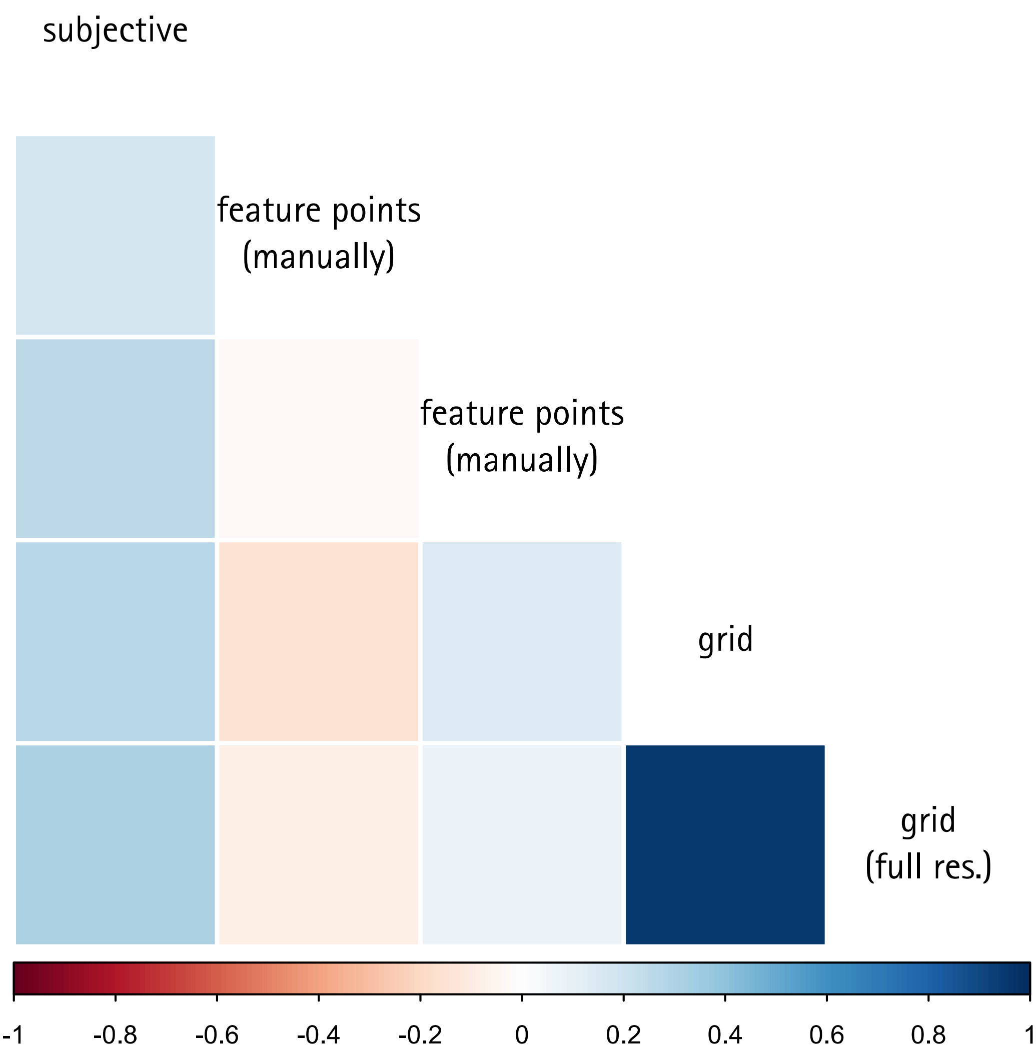

Figure 1 Correlation matrix heatmap for the typicality measures (see text for an explanation of the different measures).

5.2 Predicting sales, liking, and fluency with typicality

For the cumulative sales data, we ran five ordinary least squares (OLS) regressions at the individual image level, predicting sales from typicality. Each OLS regression tests a different typicality measure. Figure 2 visually summarizes the results. Subjectively rated typicality is significantly related to sales. Yet, the grid measure of typicality performs almost as good and is marginally related to sales. Both feature point measures and the full resolution grid measure are not significantly associated with sales.

To analyse the liking and fluency ratings of the car models, we chose a Linear Mixed-Model (LMM, Fitzmaurice, Laird, & Ware, 2004) approach, since the database contains both ratings also at the individual level of a person. To analyse the data, we relied on the lme()-function of the nlme library (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2013) of the software R (R Core Team, 2013). We again ran five models (one for every typicality measure) for both liking and fluency with random intercepts per participant, as the car models were nested within the participants.

For aesthetic liking, the models yield significant positive effects of typicality for all but the algorithmic feature point measure (Figure 3). In particular, the strongest effect can be found for the subjective typicality measure, followed by the manually coded feature point measure, the grid measure at full resolution, and the grid measure.

Fluency experience, on the other hand, is only associated with the subjective typicality measure and both grid measures (Figure 4). However, the predictive strength are higher for the subjectively rated typicality measure than the grid measure(s).

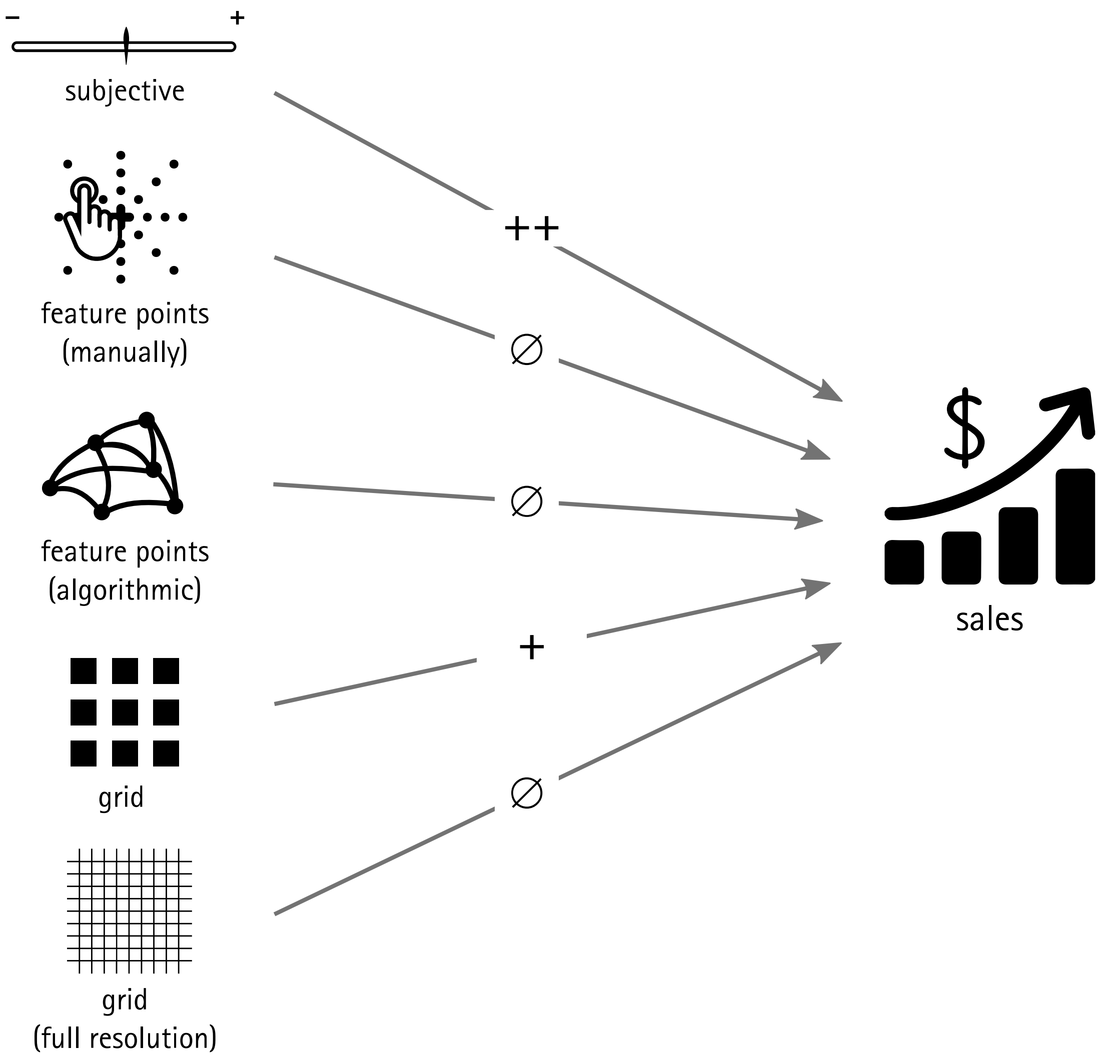

Figure 2 Results (visually simplified) for the prediction of cumulative sales from the five typicality measures. Each arrow represents the result of a distinct OLS regression with the respective typicality measure as predictor. A higher number of plus signs denotes a higher predictive power.

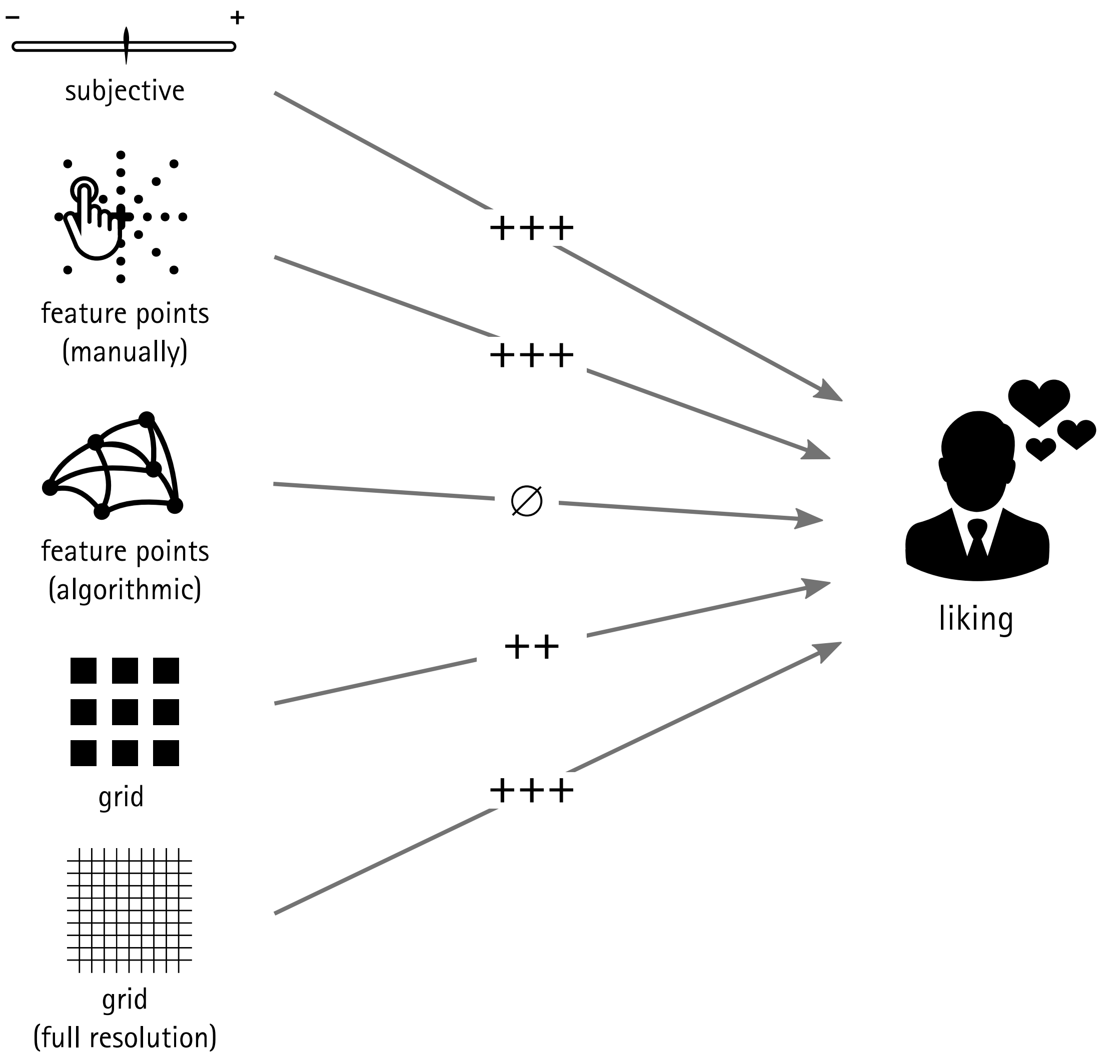

Figure 3 *Results (visually simplified) for the prediction of aesthetic liking from the five typicality measures. Each arrow represents the result of a distinct LMM regression with the respective typicality measure as predictor. A higher number of plus signs denotes a higher predictive power.*

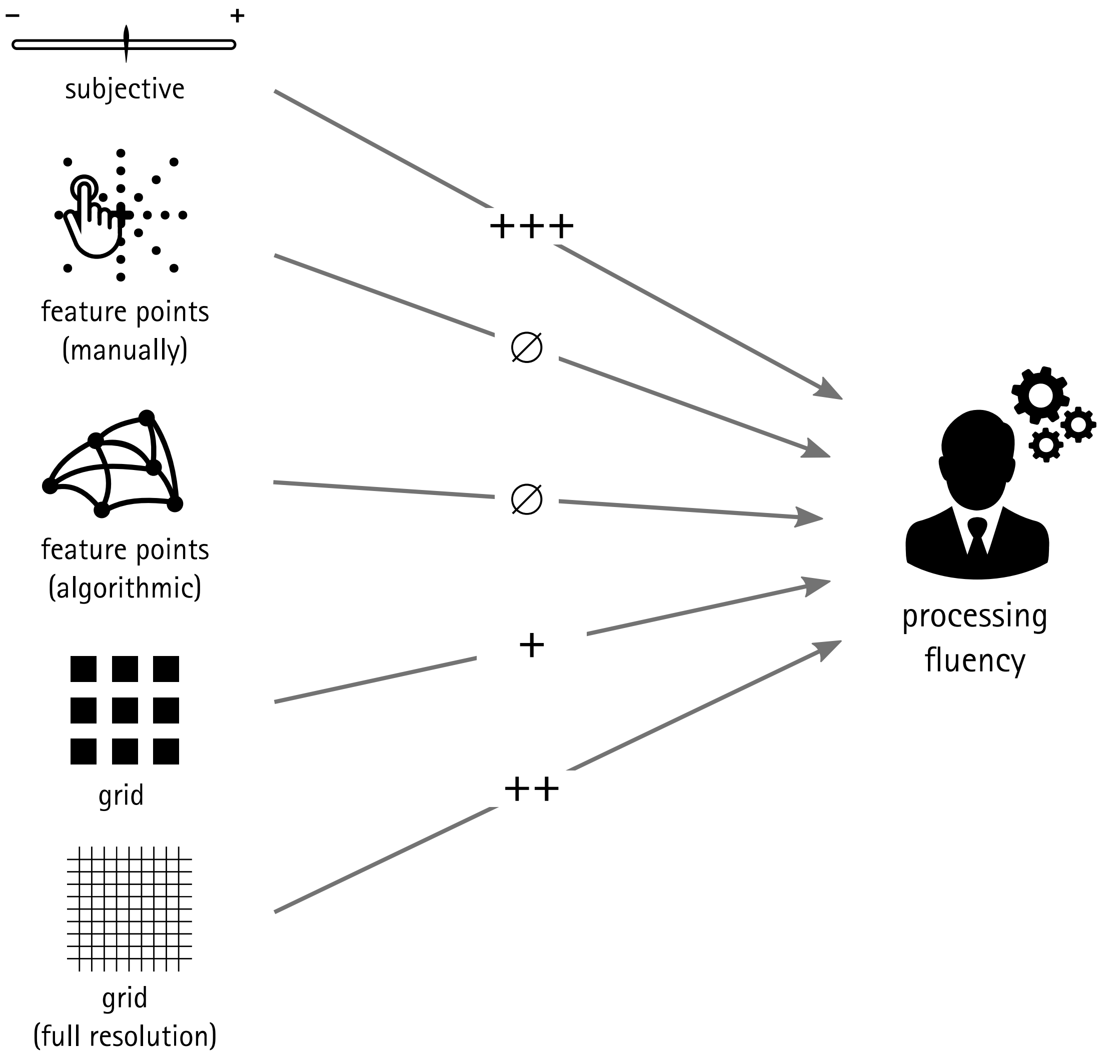

Figure 4 Results (visually simplified) for the prediction of processing fluency from the five typicality measures. Each arrow represents the result of a distinct LMM regression with the respective typicality measure as predictor. A higher number of plus signs denotes a higher predictive power.

6. Discussion

Design typicality has strong effects on aesthetic evaluations of consumers. Hence, assessing a product design's typicality plays a key role in predicting consumers' responses to a design. However, directly asking people for their subjective typicality experience may yield a biased measure as the rating arguably contains the overall aesthetic impression of the product. Although this can be useful for predictions, this approach may not be used as an unbiased measure of a design's typicality.

Therefore, the aim of this research was to examine whether unbiased objective measures of design typicality can capture a subjective typicality experience and predict outcomes of typicality. In doing so, we are the first to provide several objective measures of design typicality, and to compare them against subjective typicality as well as to provide evidence of their predictive strength regarding processing fluency, aesthetic liking, and sales. To this end, we first looked at the correlations of the objective measures with the subjectively reported typicality experience. Results show that the proposed objective typicality measures substantially correlate with the subjective measure. Admittedly, the correlations are not tremendously high. However, we used the arguably biased subjectively measure to benchmark the proposed unbiased measures. In consequence, the fact that the subjective measure possibly contains more than typicality could explain the rather modest correlation with an unbiased measure.

To overcome this objection and further assess the suitability of the objective measures, we further estimated a series of regression models predicting sales, aesthetic liking, and processing fluency with each of the four objective measures as well as the subjective measure. We find that besides the subjective measure, especially the grid measure significantly predicts car sales, aesthetic liking, and processing fluency, albeit having partly only marginal significance. However, as we tested only 77 car models, the measures could have performed better if the study design had sufficient power; this modification is advisable for future research. The found positive effects of typicality is well in line with established theory and prior research (Landwehr et al., 2011; Reber et al., 2004). Thus, as a recommendation, if your goal is to econometrically model the effect of (pure) typicality, it seems adequate to use the unbiased grid measure, thereby avoiding the reversed causality problem.

Our research assumed that findings from face research can be transferred to product design (Langlois & Roggman, 1990). In particular, we anticipated that typicality can be conceived as proximity to an average. Yet, this presumption doesn't necessarily have to be true, as measures that compute the representation of a prototype as the average of a given set of designs heavily depend on the specifics of the designs that were selected for the set. Subjectively perceived typicality, on the other hand, could be influenced by designs that are not part of the selected set of designs, changing the point of reference. Our findings, however, give reason to embrace the notion of averageness. Since we operationalized four different measures in this way with conclusive results regarding subjectively rated typicality, aesthetic liking, and processing fluency, it seems to be legitimate to compute objective typicality as the proximity to an average over a given set of designs

Still, some limitations have to be discussed. First of all, the question arises of how well the objective measures are suited in products where features and shapes are diverse and do not follow common grounds. It is to be expected that especially the two feature point-based measures are prone to errors in such situations. Further research also should address whether the visibility of explicit brand identifiers as well as the perspective used for the stimuli do have an impact on typicality measurements.

Overall, our findings demonstrate the possibility to use algorithmic, objective measures of design typicality to predict consumers' aesthetic preferences. The proposed measures can help designers either to assess the typicality of a newly created design, or to predict consumers' responses to a given design. Fortunately, the proposed grid measure that performs best is easy to understand and implement, allowing for an application in a wide range of contexts. Our research thereby contributes to the product design community by providing algorithmic, objective approaches to assess design typicality that can be used from the very first step of the design process. Finally, our findings endorse the general notion that objective measures should be included in product design research due to their robust capability of quantifying aesthetics.

7. References

Blijlevens, J., Gemser, G., & Mugge, R. (2012) The importance of being "well-placed": The influence of context on perceived typicality and esthetic appraisal of product appearance. *Acta Psychologica*, *139*(1), 178–86. http://doi.org/10.1016/j.actpsy.2011.11.004

Blijlevens, J., Mugge, R., Ye, P., & Schoormans, J. P. L. (2013) The influence of product exposure on trendiness and aesthetic appraisal. *International Journal of Design*, *7*(1), 55–67.

Boselie, F. (1995) Against Prototypicality as a Central Concept in Aesthetics. *Empirical Studies of the Arts*, *9*(1), 65–73. http://doi.org/10.2190/ERDR-FN28-PUEE-EU7F

Crilly, N., Moultrie, J., & Clarkson, P. J. (2004) Seeing Things: Consumer Response to the Visual Domain in Product Design. *Design Studies*, *25*(6), 547–577. http://doi.org/10.1016/j.destud.2004.03.001

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004) *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons.

Halberstadt, J., & Rhodes, G. (2000) The Attractiveness of Nonface Averages: Implications for an Evolutionary Explanation of the Attractiveness of Average Faces. *Psychological Science*, *11*(4), 285–289. http://doi.org/10.1111/1467-9280.00257

Hekkert, P., Snelders, D., & van Wieringen, P. C. W. (2003) Most Advanced, Yet Acceptable: Typicality and Novelty as Joint Predictors of Aesthetic Preference in Industrial Design. *British Journal of Psychology*, *94*(1), 111–124. http://doi.org/10.1348/000712603762842147

Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993) Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(9), 850–863. http://doi.org/10.1109/34.232073

Landwehr, J. R., Labroo, A. A., & Herrmann, A. (2011) Gut Liking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecasts. *Marketing Science*, *30*(3), 416–429. http://doi.org/10.1287/mksc.1110.0633

Landwehr, J. R., Wentzel, D., & Herrmann, A. (2013) Product Design for the Long Run: Consumer Responses to Typical and Atypical Designs at Different Stages of Exposure. *Journal of Marketing*, *77*(5), 92–107. http://doi.org/10.1509/jm.11.0286

Langlois, J. H., & Roggman, L. A. (1990) Attractive Faces Are Only Average. *Psychological Science*, *1*(2), 115–121. http://doi.org/10.1111/j.1467-9280.1990.tb00079.x

Martindale, C., & Moore, K. (1988) Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 661–670. http://doi.org/10.1037//0096-1523.14.4.661

Monga, V., & Evans, B. L. (2006) Perceptual Image Hashing Via Feature Points: Performance Evaluation and Trade-Offs. *IEEE Transactions on Image Processing*, *15*(11), 3452–3465. http://doi.org/10.1109/TIP.2006.881948

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2013) nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-111.

Purcell, A. T. (1993) Relations between Preference and Typicality in the Experience of Paintings. *Leonardo*, *26*(3), 235. http://doi.org/10.2307/1575817

R Core Team. (2013) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.r-project.org

Reber, R., Schwarz, N., & Winkielman, P. (2004) Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Pschology Review*, *8*(4), 364–82. http://doi.org/10.1207/s15327957pspr0804_3

Rhodes, G., & Tremewan, T. (1996) Averageness, exaggeration, and facial attractiveness. *Psychological Science*, *7*(2), 105–110. http://doi.org/10.1111/j.1467-9280.1996.tb00338.x

Schwarz, N. (2004) Metacognitive Experiences in Consumer Judgment and Decision Making. *Journal of Consumer Psychology*, *14*(4), 332–348. http://doi.org/10.1207/s15327663jcp1404_2

Veryzer, R. W., & Hutchinson, J. W. (1998) The Influence of Unity and Prototypicality on Aesthetic Responses to New Product Designs. *Journal of Consumer Research*, *24*(4), 374–385. http://doi.org/10.1086/209516

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*(6), 989–1000. http://doi.org/10.1037//0022-3514.81.6.989

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006) Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*(9), 799–806. http://doi.org/10.1111/j.1467-9280.2006.01785.x

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003) The Hedonic Marking of Processing Fluency: Implications for Evaluative Judgment. In J. Musch & K. C. Klauer (Eds.), *The Psychology of Evaluation: Affective Processes in Cognition and Emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum.

About the Authors:

Stefan Mayer is currently a Ph.D. candidate at Goethe University Frankfurt. His research interests cover various aspects of aesthetic product design, especially how the relevance of a product's design can be quantified and how visual aesthetics can be objectively measured.

Jan R. Landwehr is a Professor of Marketing at Goethe University Frankfurt and holds the Chair for Product Management and Marketing Communications. His research focuses on empirical aesthetics, product design, symbolic communication, and genesis of emotional preferences.