

Mass or only "niche" customization? Why we should interpret configuration toolkits as learning instruments

Franke, Nikolaus; Hader, Christopher

Published in:
Journal of Product Innovation Management (JPIM)

DOI:
[10.1111/jpim.12137](https://doi.org/10.1111/jpim.12137)

Published: 01/01/2014

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Franke, N., & Hader, C. (2014). Mass or only "niche" customization? Why we should interpret configuration toolkits as learning instruments. *Journal of Product Innovation Management (JPIM)*, 31(6), 1214 - 1234.
<https://doi.org/10.1111/jpim.12137>

Mass or Only “Niche Customization”? Why We Should Interpret Configuration Toolkits as Learning Instruments*

Nikolaus Franke and Christopher Hader

In order to configure individual products according to their own preferences, customers are required to know what they want. While most research simply assumes that consumers have sufficient preference insight to do so, a number of psychologically oriented scholars have recently voiced serious concerns about this assumption. They argue that decades of consumer behavior research have shown that most consumers in most product categories lack this knowledge. Not knowing what one wants means being unable to specify what one wants—and therefore, they conclude, the majority of customers are unable to use configuration toolkits in a meaningful way. In essence, this would mean that mass customization should rather be termed “niche customization” as it will be doomed to remain a concept for a very small minority of customers only. This pessimism stands in sharp contrast to the optimism of those who herald the new possibilities enabled by advances in communication and production technologies as the dawn of a new era in new product development and business in general.

Which position is right? In order to answer this question, this research investigates the role of the configuration toolkit. Implicitly, the skeptic position assumes that the individual customers’ knowledge (or absence of knowledge) of what they want is an exogenous and constant term that does not change during the interaction with the toolkit. However, learning theories suggest that the customers’ trial-and-error interaction with the configuration toolkit and the feedback information they receive should increase their preference insight. If this was true and the effect size strong, it would mean that low a priori preference insight does not impede customers to derive value from mass customization.

Three experiments show that configuration toolkits should be interpreted as learning instruments that allow consumers to understand their preferences more clearly. Even short trial-and-error self-design processes with conventional toolkits bring about substantial and time-stable enhancements of preference insight. The value of this knowledge is remarkable. In the product category of self-designed watches, the 10-minute design process resulted in additional preference insight worth 43.13 euros on average or +66%, measured by incentive-compatible auctions. A moderator analysis in a representative sample shows that the learning effect is particularly strong among customers who initially exhibit low levels of preference insight.

These findings entail three contributions. First, it becomes evident that the interaction with mass customization toolkits not only triggers affective reactions among customers but also has cognitive effects—a response category not investigated before. Second, it suggests that the pessimism regarding the mass appeal of these toolkits is not justified—mass customization has the potential to truly deserve its name. The prerequisite for this, and this normative conclusion is the final contribution, is that the toolkit should not be interpreted as a mere interface for conveying preexisting preferences to the producer. Rather, it should be treated as a learning instrument. Several suggestions are made for how firms employing this innovative business model could design their toolkits towards this end.

Address correspondence to: Nikolaus Franke, WU Wien, Welthandelsplatz 1, 1020 Wien, Austria. E-mail: Nikolaus.franke@wu.ac.at. Tel: + 43 1 31336 4585.

* The authors would like to thank Frank Piller and Steffen Wiedemann for their contribution to the pilot study, the students of the WU course “E&I Research” of the summer semester 2008 who supported the authors during the data gathering of study 1, and Thomas Stöckl who greatly supported the data gathering and analysis of study 2. Finally, we gratefully acknowledge the funding of the Austrian Science Fund (Fonds zur Förderung der wissenschaftlichen Forschung FWF) that made this project possible.

Introduction

Is mass customization doomed to remain “niche customization,” a concept limited to only a small minority of consumers? The idea that tailoring products to the individual customers’ preferences has universal appeal has been challenged by some authors. “The promise . . . [has] been greatly exaggerated,” concludes the leading marketing scholar Itamar Simonson

(2005, pp. 42–43). “The value-added and impact of individually customized offers, as opposed to simple usage/benefit-based segmentation, will often be rather limited.” Bharadwaj, Naylor and ter Hofstede (2009, p. 225) propose that in extreme cases, it might be beneficial for companies to actively discourage those consumers who do not know exactly what they want from mass customization and “help [them] understand that the standardized [versus a customized] system is a more viable option.”

Their argument is that customers often lack insight into their own preferences, and tailoring products to their articulated preferences can therefore mean that customers will get products they eventually dislike (Syam, Krishnamurthy, and Hess, 2008). Indeed, a rich body of research into consumer decision-making indicates that most consumers essentially lack precise knowledge of what they want in most product categories (e.g., Bettman, Luce, and Payne, 1998; Loewenstein and Prelec, 1993; Tversky and Kahneman, 1981). It appears unreasonable to expect such consumers to be able to precisely define individual products that actually fit these unknown preferences (Bharadwaj et al., 2009; Kramer, 2007; Simonson, 2005; Syam et al., 2008). This argument is not mere speculation. Empirical studies by Franke, Keinz, and Steger (2009) and Bharadwaj et al. (2009) confirm that the value customers derive from customization is negatively affected by their level of preference insight. The conclusion that mass customization places excessive strain on most consumers is substantiated by the actual economic significance of mass customization. Although the seminal work on the mass customization concept by Pine (1999) is now two decades old, the market shares of customized consumer products are still atomic relative to traditional segmentation-based standard products (Gownder, 2011).

The skepticism of this line of argumentation stands in contrast to the optimism of those who herald the new possibilities enabled by advances in communication and production technologies as the dawn of a new era in business (Cook, 2008; O’Hern and Rindfleisch, 2009; Prahalad and Ramaswamy, 2004; Seybold, 2006; Sheth, Sisodia, and Sharma, 2000; von Hippel, 2005). Eventually it will be possible, so they contend, to put one of the most basic principles of marketing into effect, namely “giving customers what they want” (McKenna, 2002). Their optimism has been fueled by empirical research that has repeatedly demonstrated that customers are willing to pay a considerable premium for self-designed products (e.g., Franke & Piller, 2004; Franke et al., 2009; Schreier, 2006), and as a matter of fact, a growing number of firms are offering virtual configurators that allow consumers to self-customize T-shirts, sneakers, furniture, watches, business cards, cars, cereals, etc. Reflecting this optimism, the Marketing Science Institute has identified customer co-creation as a top research priority (Marketing Science Institute, 2008). However, these scholars have remained silent on the skeptics’ argument regarding insufficient preference insight.

Against this background, does mass customization have any chance of becoming the future paradigm of new product development and design? Or is it doomed to remain “niche customization,” a concept limited to the small minority of customers who have precise knowledge of what they want? In this paper, we attempt to bridge these contradicting positions. Indeed, tailoring a product to one’s own liking requires high preference insight—but this necessary preference insight may be generated during the process of self-designing. The core argument underlying this paper is that the knowledge of what one wants is not an exogenous and constant term. Learning theories suggest that trial-and-error interaction with the configuration toolkit and the feedback information the consumers receive should increase their preference insight. If this is true and the effect size strong, it would be short sighted to interpret configuration toolkits merely as an interface for conveying existing preferences to the producer. The toolkits’ potential would be better captured if we interpreted them as *learning instruments*. The research question investigated in this paper is therefore whether and to what extent self-designing a product with a configuration toolkit affects the customer’s preference insight.

The empirical base is three experiments with extant configuration toolkits for self-expressive goods (wrist watches and running shoes), where the focus is on esthetic properties of the product. The clear finding is that

BIOGRAPHICAL SKETCHES

Dr. Nikolaus Franke is professor of entrepreneurship and innovation at the Vienna University of Economics and Business (WU Wien) and leads of the Vienna User Innovation Research Initiative (<http://www.userinnovation.at>). He is interested in understanding the phenomenon of creative and innovative users and methods that help companies using this potential. Particularly, he researches the phenomena of lead users, toolkits for user innovation and design, and crowd sourcing.

Dr. Christopher Hader is a strategy consultant at Accenture. During his doctoral studies at WU Wien, he developed an interest in topics in the field of innovation and marketing. His research is primarily focused on user innovation phenomena like toolkits for user innovation. He continues to deepen his knowledge in this field during his consulting career.

customers *learn by self-designing*. Even short trial-and-error self-design processes with conventional toolkits bring about substantial and time-stable enhancements of preference insight, especially among customers who initially exhibit low levels of this insight. This allows the conclusion that at least in the category of self-expressive goods, self-designing with configuration toolkits does not require high preference insight a priori. The preference insight necessary to obtain value from mass customization can be enhanced *during* the self-design process *through* the activity of self-designing.

This offers theoretical and managerial contributions. On the theoretical side, it contributes to the emerging body of research devoted to customer interaction with configuration toolkits. A growing number of studies have addressed the *affective* customer reactions elicited (Dellaert and Stremersch, 2005; Franke and Schreier, 2010; Franke, Schreier, and Kaiser, 2010; Moreau and Herd, 2010; Valenzuela, Dhar, and Zettelmeyer, 2009). A better understanding of these reactions is also of practical interest for the design of toolkits. The same argument can be made for consumers' *cognitive* reactions. In contrast to consumers' affective reactions, cognitive reactions have hardly been the subject of academic research in this context. Therefore, the finding that toolkits strongly contribute to customer knowledge not only has theoretical and practical implications; it also opens up a promising and interesting trajectory for future research. Given that companies follow the interpretation of toolkits as learning instruments and design them in a way that maximizes learning effects, this warrants the optimistic prediction that mass customization will be attractive to many customers and thus holds the potential to truly deserve its name.

Two Different Interpretations of Configuration Toolkits

Toolkits as Interfaces for Conveying Explicitly Known Preferences

The customers' interaction with the configuration toolkit has seen little attention in research on mass customization. For example, this aspect is hardly even mentioned in Pine's pioneering book on mass customization (Pine, 1999), which has been cited over 2000 times. The functions of this interface are seen as (1) providing customers with information on the combinations possible within production capabilities and (2) allowing them to assemble and order the most preferred combination (Deng and Hutchinson, 2009; Kotha, 1995; Syam and Kumar, 2006).

Compared to traditional standard product offerings, the most distinctive feature of toolkits is that they decompose the product into dimensions and attributes, similar to a conjoint approach. The large variety of combinations resulting from the combinations of these dimensions and attributes offers customers a far greater selection and thus potentially a much closer preference fit (Dellaert and Stremersch, 2005; Franke & Piller, 2003; Ghosh, Dutta, and Stremersch, 2006; Pine, 1999; Randall, Terwiesch, and Ulrich, 2007; von Hippel, 2001). The prototypical example of such a toolkit is the Dell computer configurator, which provides customers with numerous attributes within dimensions such as processor type, RAM, screen size, etc., but hardly any feedback information on the consequences of choices and combinations. The underlying assumption of such "shopping list" toolkits is that customers possess detailed knowledge of their preferences and are thus able to determine the idiosyncratic combination of attributes that matches their individual optimum most closely within the solution space (e.g., Bardakci and Whitelock, 2003; Kotha, 1995; Liechty, Ramaswamy, and Cohen, 2001; Pine, Peppers, and Rogers, 1995; Squire, Readman, Brown, and Bessant, 2006). Or, as Pine et al. (1995, p. 103) put it more explicitly, "customers . . . want exactly what they want—when, where, and how they want it." It is also interesting that much empirical research on mass customization use toolkits that provide many options but little or no feedback information (e.g., Bharadwaj et al., 2009; Dellaert and Stremersch, 2005; Franke et al., 2009; Huffman and Kahn, 1998; Liechty et al., 2001; Park, Jun, and MacInnis, 2000; Randall et al., 2007; Syam et al., 2008). Notably, most goods underlying these toolkits have a utilitarian character—they allow the configuration of individual PCs, laptops, hotel rooms, newspapers, automobiles, stereo systems, yellow page websites, treadmills, sofas, and home theater systems. It may be that scholars assume that regarding such goods, consumers' preference insight is sufficiently high and cannot be altered through the interaction with the toolkit. (This aspect will be deepened in the General Discussion section.)

Some scholars within this trajectory take learning processes into account. Their focus, however, is generally on the customers' learning about the toolkit's solution space with its myriad possibilities. For example, Huffman and Kahn (1998), Randall et al. (2007), and Valenzuela et al. (2009) investigate the presentation modes (e.g., attribute based or needs based) that best correspond to the customers' expertise and thus make it easier for them to minimize the distance between the combination and their preferences. The underlying assumption is still that

customers *have* an understanding of what they want and that this preference insight is hardly affected by interaction with the toolkit.

If we regard toolkits primarily as interfaces for conveying explicitly known preferences, the finding that preference insight is actually low among most consumers deals a significant blow to the viability of the mass customization concept (Bharadwaj et al., 2009; Kramer, 2007; Simonson, 2005). Customers who do not know precisely what they want will, of course, face difficulties specifying their individual ideal product. It is only consequent to suggest that in such cases, the producer should handle the task of analyzing individual customer needs instead of letting the customers self-design (Ghosh et al., 2006) or should offer traditional standard products instead (Bharadwaj et al., 2009; Simonson, 2005).

Toolkits as Learning Instruments

The conceptual work of von Hippel (2001), von Hippel and Katz (2002), and Wind and Rangaswamy (2001) relies on a different interpretation of toolkits. While they do not question the toolkit's function of transferring preference information from the customer to the producer, those scholars point out that a toolkit must first support customers in learning about their preferences: "Its focus is to help customers to better identify or define for themselves what they want" (Wind and Rangaswamy, 2001, p. 15). Moreau, Boney, and Herd (2011); Payne, Storbacka, and Frow (2008); as well as Randall, Terwiesch, and Ulrich (2005) also maintain that the toolkit should provide the consumers with information that enables them to develop a better understanding of their preferences. The primary means for such learning is presumed to be feedback information on the consequences of choosing specific attributes, dimensions, and in particular their interactions, when combined. The authors mentioned above emphasize this function as a second distinctive feature of toolkits: "It is crucial that toolkits . . . enable users to go through complete trial-and-error cycles as they create their designs" (von Hippel, 2001, p. 251). Randall et al. (2005) likewise suggest providing rich simulations of the self-designed product or interim solutions, explaining that such "prototypes are important even for professional designers; and they play an even bigger role for user design" (p. 80). In this interpretation of configuration toolkits, it is assumed that the individuals' preference insight can be increased by interacting with the toolkit. Implicitly, configuration toolkits are seen as learning instruments. Consequently, the finding from consumer research that most customers have

low preference insight does not actually question the value of self-design as a business concept for the mass market.

Which interpretation of configuration toolkits is more appropriate in guiding our theoretical thinking and the practical implementation of mass customization? Although Franke and Piller (2003) identified preference learning as a priority research issue in their literature review on mass customization, the authors of this paper are aware of no empirical study that has investigated the impact of interacting with a self-design toolkit on preference insight. Existing data from a different research project (Franke and Piller, 2004) hence were used for a pilot analysis.

Just Conveying or Learning Preferences? A Pilot Study

One hundred and sixty-two business students from a large European business school took part in "a short research experiment" (average age: 25.1 years; 56% females) in which they were asked about their preference insight with regard to the design of Swatch-type plastic wristwatches. In line with the findings from consumer research reported above, a self-assessment revealed that only a small minority of 20 students (12%) reported having high-preference insight, while most gave negative responses when asked whether they had "a clear idea of what [their] ideal watch should look like." They were asked to use a configuration toolkit which allowed them to self-design watches. They were informed that there would be a raffle in which they could win "their" watch. The toolkit used was the IDtown toolkit, which allows users to configure an individual watch within a design space of five generic watch types (sports, metal, ladies, etc.), with five design dimensions within each type (face, case, strap, hour and minute hands, seconds hand), and 30 to 150 design attributes within each dimension. There is no prescribed order of selection, and users can jump forward or backward whenever they wish. They also receive instant visual feedback showing the current design of their watch. During the participants' self-design processes, their behavior was tracked with a spy program (indiscernible to participants). Once they had finished, they were asked about their maximum willingness to pay (WTP) for their individual self-designed watch (using the contingent valuation method; Voelckner, 2006).

Which of the two interpretations of toolkits is supported? At first sight, it appears that the findings favor the "niche customization" prediction. Those participants who had low preference insight prior to starting the self-design

process derived less value from using the toolkit; their willingness to pay for their watches was lower than that of participants with high preference insight (WTP = 85.07 euros versus 95.35 euros), replicating findings from Bharadwaj et al. (2009) and Franke et al. (2009). However, if one looks more closely, a different view emerges. First, the difference between the two groups is not overly strong, nor is it significant ($p = .216$), suggesting that even participants with low initial preference insight somehow managed to come up with solutions that created value for them. How did they accomplish this without knowing exactly what they wanted? A plausible answer can be found in the participants' interaction patterns with the toolkit. It became obvious that participants did not just use the toolkit to "convey preexisting information." Such behavior would imply that participants with clear preference insight simply choose their favorite watch type, inspect the design elements, and then select the face, case, strap, hour and minute hands, and seconds hand that match their preferences most closely. However, not a single participant showed such linear interaction with the toolkit. They all engaged in considerable trial-and-error activities, that is, they tried out different combinations, jumped forward and backward, and iteratively progressed toward their eventual individual solution. However, a clear pattern emerged: Those participants with low initial preference insight were far more active than those who already had clear preference insight at the outset. The former changed watch types ($M = 3.9$ changes versus 2.0 changes, $p < .05$) and design elements ($M = 118.0$ changes versus 63.3 changes, $p < .000$) significantly more often than participants with clear preference insight, and they discarded their designs and started from scratch significantly more often ($M = 3.5$ new starts versus 1.9 new starts, $p < .05$). Generally, they spent significantly more time on the self-design process ($M = 13.6$ minutes versus 7.3 minutes, $p < .000$). Our interpretation is that this increased trial-and-error learning behavior enabled people with low initial preference insight to partly compensate for their relative disadvantage: They discovered what they wanted by interacting with the toolkit. Therefore, the value they eventually attributed to their self-designed watches is close to the value assigned by participants who knew what they wanted from the outset. In sum, this is a clear argument indicating that self-design toolkits are used as learning instruments by those customers who have low initial preference insight, and that this learning equips them with the necessary preference insight to define their individual product. Of course, this pilot study is only an initial exploration, as there are several alternative expla-

nations. Theory-guided controlled experiments appear necessary in order to rule them out.

Hypotheses on Toolkits and Preference Learning

What happens when consumers with imperfect preference insight attempt to design their own product with a configuration toolkit? Such toolkits allow the consumer to manipulate the product along several dimensions (such as color, shape, functions, etc.), each with a number of attributes, and they provide immediate informational feedback on the anticipated consequences of these trials. For example, they show the user–designer how certain color choices in the different parts of the product would look once combined to form a complete product. Feedback can, of course, also include functional feedback (von Hippel, 2001) or social feedback (Franke, Keinz, and Schreier, 2008). It appears plausible that these trials and the feedback information the user–designer receives will result in enhanced preference insight, i.e., a better understanding and knowledge of one's own preference structure.

This pattern is in line with connectionism learning theory (McClelland and Rumelhart, 1986; Smith, 1996). Connectionism learning theory constitutes a modern version of early learning models, especially the learning theory proposed by Thorndike (1911). It focuses on experience-based associative learning, which means that learning is portrayed as an incremental process of changing associations in a person's mind as a result of his or her own experiences (Smith, 1996). Self-designing products with a configuration toolkit allows simulated experiences, meaning that connectionism dovetails neatly with the research area of toolkits (Janiszewski and van Osselaer, 2000, generally compared different experience-based associative learning theories and concluded that connectionism is the theory best suited for explaining consumer learning). Congruent with the findings from our pilot study, connectionism suggests that the consumer does not assemble the product in a directed process but tries out and evaluates different alternatives iteratively, thus engaging in a learning process (Smith, 1996). What will be the outcome of such processes? Connectionism defines learning as a process of creating or modifying associations between different cues in one's mind (Smith, 1996). Cues enabled by the toolkit in this learning process include product attributes and attribute combinations as well as the user's like (or dislike) of the overall simulated product. The latter is derived from the individual's underlying "tacit" preferences (Bettman, Luce,

and Payne, 2008; Simonson, 2008). These preferences allow the individual to decide whether he or she likes a given (new) stimulus. The individual will then form associations regarding the value delivered by specific product dimensions, specific attributes, and specific combinations (see Janiszewski and van Osselaer, 2000, and Keller, 1993, for an analogous argument on how consumers learn brand meanings). It is important to emphasize that the preference learning with a toolkit is most likely not “blind” trial and error but an *adaptive* learning process. Step by step, consumers purposefully search for potential improvements to their current solution by identifying and evaluating promising alternatives, thereby building on and extending existing associations. They iteratively strengthen existing associations, create more associations, and also improve their coherence, thus undergoing a process of preference stabilization (Hoeffler and Ariely, 1999). The resulting set of associations corresponds to an individual’s *preference insight*: The more such associations exist, the stronger they are, and the more consistent they are, the higher the individual’s preference insight is. Thus, our fundamental hypothesis is as follows:

H1: Self-designing a product with a configuration toolkit that provides feedback information increases the individual’s preference insight in the respective product category.

The same argument can, of course, also be made for buying products in a shop. Connectionism learning theory predicts that a consumer also learns from inspecting standard products. The process is similar: The individual forms associations between the specific product attributes and his/her like (or dislike) of the overall product (McClelland and Rumelhart, 1986). The difference, however, is that adaptive learning receives far less support, and customers are not able to vary isolated attributes and inspect their effects like in a factorial experimental design (as they are when using a configuration toolkit). The individual is restricted to the given standard products as simulated experiences. Thus, even when very many and very different standard products are provided, the process of preference stabilization will be much less effective than with a configuration toolkit. Therefore, the learning effect regarding one’s own preferences may be much stronger if individuals engage in self-designing than if they inspect and compare standard products, even if they are provided with a large selection.

H2: Self-designing a product with a configuration toolkit that provides feedback information increases the individual’s preference insight in the respective product category.

egory to a greater extent than inspecting standard products.

Learning effects will not be similar for all individuals (McClelland and Rumelhart, 1986). Of particular interest are those customers who exhibit low initial levels of preference insight, as mass-customization skeptics have argued that the lack of such insight constitutes a crucial barrier for the use of configuration toolkits. Thus, in order to investigate whether mass customization is only of interest to the small minority of customers who have a clear understanding of their own preferences or is in fact a concept of potentially much larger impact, the initial level of preference insight is taken as a moderator.

Is the absence of initial preference insight a problem or an advantage for preference learning by self-designing? An argument for the latter is that consumers who have low levels of preference insight may derive especially great benefits from trial-and-error learning interaction, as they simply have the most room for learning. Those who already have a clear understanding of what they want, in contrast, may reach a ceiling sooner. The underlying theoretical argument is the “power law of practice,” which states that cognitive learning increases rapidly in early stages and among individuals with low starting levels, whereas even minor improvements require considerable effort at later stages and among already knowledgeable individuals (Haider and Frensch, 2002; Newell and Rosenbloom, 1981). This pattern has been confirmed empirically in various fields and for very different learning tasks (see Johnson, Bellman, and Lohse, 2003; Ritter and Schooler, 2004), and it is considered an “empirical generalization” in cognitive sciences. However, there are also counter-arguments. Sometimes, learning curves are S-shaped, which means that individuals with low levels of prior knowledge are not (yet) effective in learning (Bechtel and Abrahamsen, 2002; Rumelhart and McClelland, 1986). When learning tasks are difficult, for example in a course of advanced econometrics, such S-curves can be quite heavy tailed: The learning effect is then marginal for all those subjects who fall below a relatively high level of prior knowledge (Bechtel and Abrahamsen, 2002). But typical mass customization configurators are quite intuitive and easy to handle (Franke & Piller, 2003; Moreau and Herd, 2010). Starting solutions, immediate feedback information, and the trial-and-error learning facilitated by the toolkit will enable individuals with particularly low initial preference insight to make fast progress in understanding what they want.

H3: The effect of self-designing on preference insight is moderated by the prior level of preference insight such

that the learning effect is stronger when individuals have a lower level of initial preference insight.

Study 1: The Effect of Self-Designing on Preference Insight

In order to test whether self-designing with a toolkit enhances preference insight, we devised a between-subject experiment. It started with a measurement of the participants' initial preference insight with regard to the esthetic design of running shoes. Then participants were provided with Nike's prototypical self-design toolkit, which allows users to self-design running shoes by selecting colors for various parts of the shoe, and which gives instant visual feedback on the resulting design. Once the participants had finished, their preference insight was again measured and analyzed whether it had changed. The learning increment was then compared with that of a second group who had the opportunity to inspect standard shoes on a website. In order to rule out a number of alternative explanations, two additional control groups were used.

Procedure, Sample, and Experimental Groups

One hundred and thirty-eight undergraduate and graduate business students from the authors' university (average age: 24.4 years; 42% females) took part in the experiment. The objective of the study was not revealed. Participants were compensated with the amount of 10 euros and a chance to win a product in a raffle. They were randomly assigned to four groups. Each participant was seated in a separate room, thus the setting ensured that there was no interaction between them. The instructions were standardized, and the instructors had been trained in a workshop prior to the study. Participants in all groups started with a short written questionnaire that contained control variables. The ensuing steps were different in the four groups.

Treatment group (pre/post self-design). After administering the written questionnaire with a number of control variables, participants were questioned verbally about their preference insight regarding running shoes (see Measurement section). Then they were instructed to use the Nike configuration toolkit to self-design the running shoe they liked best, a task which they managed to complete in 18.7 minutes on average ($SD = 9.9$). Prior to this task, they were informed that there would be a raffle in which they could win their self-designed shoe. In

this way, it was ensured that their self-design activities corresponded closely to those of real customers creating shoes they want to buy (which ensures reasonable external validity). Participants were able to use the full toolkit, and only ready-made shoe designs were disabled in order to force participants into a real trial-and-error learning process. After they had finished the self-design process, again their preference insight regarding running shoes was measured. If H1 is correct, a significant increase in preference insight before and after self-designing should become evident.

Control group 1 (pre/post shopping). After a measurement of preference insight similar to that performed in the Treatment Group, participants were introduced to the NIKEiD website, on which all self-design possibilities had been blocked. In this form, the website is equivalent to a large online shop. They were asked to inspect the approximately 3500 predesigned running shoes provided (with the same shoe types as in the former group) and to select the one design which they would want to receive if they won the raffle. In this way, a similar level of involvement regarding running shoe designs as in the Treatment Group was created. The inspection task took participants 13.6 minutes on average ($SD = 8.9$). As this period was shorter than that required by the Treatment Group, the time of exposure to running shoe designs (for both self-design and inspection activities) was included as a covariate in the later analysis. Finally, again their preference insight regarding running shoes was measured. If H2 is correct, then this group should exhibit a smaller increase in preference insight compared to the Treatment Group.

Control group 2 (post self-design). The objective with this control group was to eliminate the possible alternative explanation that it is not trial-and-error learning with the toolkit but the initial *measurement* of preference insight (and the cognitive processes triggered by this measurement) which cause the change in preference insight. Such mere measurement effects are sometimes reported in psychology (e.g., Fitzsimons and Morwitz, 1996), and experimental designs that control for them include the Solomon design (Solomon, 1949). It is necessary to rule out these effects in our setting because research has demonstrated that merely thinking about a product might alter the participants' preference structures (Xu and Wyer, 2007). The participants in this group were treated similarly to the Pre/Post Self-design (Treatment) Group, with the only difference being that preference insight was measured only after (and not before) the

self-design task, for which they needed 16.1 minutes on average ($SD = 8.5$). If H1 is correct and the alternative explanation of a mere measurement effect played no role in our experiment, we should observe no significant difference in the preference insight measurement between this group and the Treatment Group (pre/post self-design) after the task. There should be, however, a significant difference in the subjects' preference insight after the task between this group and Control Group 1 (pre/post shopping).

Control group 3 (pre/post alternative product self-design). The purpose of this group was to eliminate the possible alternative explanation that the effect measured might not be attributed to object-specific trial-and-error learning but to the *positive experience* of self-designing as such. Research has found that creating a product design with a toolkit triggers positive emotions of competence and autonomy (Franke and Schreier, 2010), a phenomenon that can be explained by self-determination theory (Ryan and Deci, 2000). Such high spirits might increase the participants' tendency toward acquiescence when asked about their preferences and therefore evoke a "mood effect" (Forgas, 1995). Hence, after the initial measurement of preference insight regarding running shoes, participants in this group were instructed to self-design a T-shirt with a toolkit they were provided with (from Shirtcity). This product category was selected because T-shirts are products of approximately similar interest to students. We chose this specific toolkit because its complexity and design freedom roughly correspond to that of the NIKEiD toolkit, and because a pilot study with $n = 7$ participants had revealed that the T-shirt toolkit evoked similar levels of positive emotions. As in the former groups, participants were informed that there would be a raffle in which they could win their self-designed T-shirt. Participants took an average of 11.1 minutes ($SD = 7.8$) to complete the task, which was shorter than the time required by participants in the other groups. However, as time was included as a covariate in the later analysis, this should not constitute a major problem. If H1 is correct and the alternative explanation of a mood effect played no role in the setting, this group should show a smaller increase in preference insight compared to the Treatment Group (pre/post self-design) and a lower level of post-treatment preference insight compared to Control Group 2 (post self-design).

Participants then were paid in all four groups, at the same time we informed them that it might be necessary to contact them again and requested their phone numbers.

Two weeks later, the researchers called them and measured their preference insight a third time. The purpose was to test whether the increase in preference insight remained stable over time, that is, whether the preferences learned had any strength (Song-Oh and Simonson, 2008). Studies have found that two weeks are likely to provoke a substantial forgetting effect with regard to the experiment and the answers given (Kwon, Cho, and Park, 2009). If there was an increase immediately after the stimulus task (in the pre/post self-design group) but no effect later, this would even question whether there was actually a learning effect at all, as most scholars agree that learning requires some time stability (Rodriguez, 2009). Altogether, the experiment took place in the following order: (1) questionnaire with control variables in all groups, (2) preference insight as a premeasurement (termed t_0 below) in the Treatment Group and Control Groups 1 and 3, (3) treatment (similar in Treatment Group and Control Group 2; different in Control Groups 1 and 3) at t_1 , (4) preference insight as a postmeasurement immediately after treatment (t_2) in all four groups, (5) and again two weeks after the experiment (t_2) in all groups.

Measurement

Preference insight. Measuring preference insight was the greatest challenge in this project. In the pilot study, the measurement was based on self-assessment (for a similar approach, see e.g., Bharadwaj et al., 2009), which raises validity issues. While people with high preference insight might be able to give a valid appraisal, those with low preference insight might encounter problems—after all, these individuals by definition lack insight. Research into the "Dunning–Kruger effect" has shown that ignorant people are often unaware of their ignorance, limiting the value of self-assessments (Kruger and Dunning, 1999). As preference insight is the core construct in this experiment, we thus refrained from self-assessments and employed a self-developed test instead. The idea underlying this test is that individuals with high levels of preference insight should be able to specify their individual ideal product more clearly when asked to do so than people with low preference insight (Hoeffler and Ariely, 1999; Simonson, 2005). A set of seven questions was used, such as "Which base color do you prefer for your running shoes?" covering different dimensions of the design of running shoes (see Appendix for the full set of questions). Prior to each measurement, participants were instructed to answer "I do not know" if they were unable to give a valid answer. Such responses are indicators of

low preference insight (value = 0), while any specific answers point to high preference insight (value = 1) in this dimension. The individual mean served as indicator for each participant's preference insight. One of the seven items (the one shown above) was used at both t_0 and t_2 in order to allow direct comparison. From the remaining six items, three items were randomly selected for each interview at t_0 (in random order). The other three items in each individual case were then used at t_2 , again in random order. Thus, with the exception of the first item mentioned above, the preference questions were different at t_0 and t_2 . The purpose of this procedure was to rule out memory effects. The findings reported below are based on all items; however, they remain stable if we (1) use only the one item asked both at t_0 and t_2 , and (2) only the items that were different at t_0 and t_2 . At t_2 , we used all seven items, once again in random order. We standardized item difficulty by subtracting the item-specific means across all participants at t_0 (when the measurement had not been impacted by the experimental tasks) and transformed the scale into an interval of [0; 1], where 0 corresponds to individuals who have no idea of the esthetic choices they would prefer, and 1 indicates persons who know precisely what they want.

Control variables. The following control variables were included that may also influence preference insight: (1) involvement (seven items, $\alpha = .82$, adapted from Zaichkowsky, 1985), (2) innovativeness (six items, $\alpha = .82$, adapted from Jackson, 1983), (3) design

affinity (four items, e.g., "I like designing things," $\alpha = .80$), (4) experience with running shoes (three items; e.g., "I use running shoes very often," $\alpha = .81$), (5) experience with toolkits (four items; e.g., "I use toolkits often," $\alpha = .81$). All items were measured on seven-point scales ranging from "strongly disagree" to "strongly agree." Multi-item constructs were averaged, and participants were also asked to indicate their age and gender. The random assignment to groups ensured that there were no significant differences between the experimental groups in any of these variables. In addition, the time each participant took for the treatment task was measured.

Findings and Discussion

The first hypothesis stated that self-designing enhances preference insight. Findings clearly confirm this (Table 1). In the Treatment Group (pre/post self-design), a highly significant increase in preference insight between t_0 and t_2 can be observed (from $M = .62$ to $.76$, $p < .001$). At t_2 , participants in Control Group 2 (post self-design) showed levels of preference insight almost identical to those of participants in the Treatment Group ($M = .76$ versus $.75$, not significant [n.s.]), which suggests that the change observed in the Treatment Group was not caused by the measurement at t_0 . In both groups, the level of preference insight remained quite stable at t_2 ($M_{\text{Treatment Group}} = .74$ versus $.76$, n.s. and $M_{\text{Control Group 2}} = .76$ versus $.75$, n.s.), indicating that "real" and enduring

Table 1. Pre- and Post-Measurement of Preference Insight

Group	t_0	t_1 Treatment	t_2	t_3^a	t -test (t_0 - t_2)	t -test (t_0 - t_3)
	Preference insight M (SD)		Preference insight M (SD)	Preference insight M (SD)		
Treatment Group (TG, $n = 42$)	.62 (.22)	Self-design with running shoe toolkit	.76 (.09)	.74 (.08)	$p < .001$	$p < .001$
Control Group 1 (CG 2, $n = 37$)	.67 (.21)	Selection in running shoe shop	.70 (.23)	.65 (.21)	n.s.	n.s.
Control Group 2 (CG 1, $n = 31$)	not measured	Self-design with running shoe toolkit	.75 (.09)	.76 (.05)	n.a.	n.a.
Control Group 3 (CG 3, $n = 28$)	.63 (.27)	Self-design with T-shirt toolkit	.64 (.29)	.62 (.28)	n.s.	n.s.
ANOVA (Groups 1–4)	n.s. ^b		$p < .01^c$	$p < .01^c$		
Post hoc tests						
TG versus CG 1	n.s.		$p < .05$	$p < .05$		
TG versus CG 2	n.a.		n.s.	n.s.		
TG \cup CG 2 versus CG 1	n.a.		$p < .05$	$p < .01$		
TG versus CG 3	n.s.		$p < .001$	$p < .01$		
TG \cup CG 2 versus CG 3	n.a.		$p < .001$	$p < .01$		

^a Two weeks later.

^b Covariates: Involvement, innovativeness, design affinity, experience with running shoes, experience with toolkits; results are robust to the exclusion of covariates.

^c Additional covariates: Additional treatment time; results are robust to the exclusion of covariates.

n.a., not applicable; n.s., not significant.

learning had occurred. The differences observed were not caused by mood effects, as the toolkit interaction itself does not explain the increase in preference insight: In Control Group 3 (pre/post alternative product self-design), there is almost no difference between the participants' preference insight at t_0 , t_1 , and t_2 ($M = .63$ to $.64$, and $.62$). Altogether, the findings provide clear confirmation of H1. There is strong evidence that self-designing a product with a configuration toolkit generates a substantial and time-stable learning effect.

Also H2 is supported. Here, it was hypothesized that self-designing brings about a greater learning effect than inspecting standard products. In contrast to the Treatment Group, the small increase in preference insight in Control Group 1 (pre/post shopping) between t_0 and t_2 is insignificant (from $M = .67$ to $.70$, n.s.) and completely disappears in the measurement two weeks later at t_2 ($M = .65$), which suggests that a learning effect of inspecting standard items is not visible in our setting. More importantly, the difference in preference insight between the Treatment Group (pre/post self-design) and Control Group 1 (pre/post shopping) is significant at t_1 and t_2 ($M_{\text{Treatment Group}} = .76$ and $M_{\text{Control Group 1}} = .70$, $p < .05$ at t_1 and $M_{\text{Treatment Group}} = .74$ and $M_{\text{Control Group 1}} = .65$, $p < .05$ at t_2). Merging the Treatment Group (pre/post self-design) and Control Group 2 (post self-design) at t_1 and t_2 , the significance levels are pronounced. This allows the conclusion that self-designing with a configuration toolkit increases consumers' preference insight to a far greater extent than inspecting standard products in a shop does.

Two important questions remain, and both concern the construct of preference insight. The first relates to internal validity and asks whether our test-based measurement actually reflects preference insight. The second addresses external validity. Assuming that the measurement indeed captured preference insight: Does the increase in preference insight caused by self-design have any practical significance?

Study 2: Validation

In this between-subject experiment participants again self-designed a product with a configuration toolkit. We measured how this experience increases their preference insight in the respective product category (in this study: watches). The same measure as in Study 1 was used and compared with two alternative measures. This should allow a better interpretation of the learning effect size as it could be converted to a less abstract scale, namely the

increased monetary value this enhanced preference insight generates for the individual.

Validating the Measure of Preference Insight

Again, preference insight was measured by asking about the characteristics of each participant's subjective ideal product design. To this end, three questions in randomized order were used (see Appendix; item selection was again based on a pilot study with $n = 14$ watch customers). As in Study 1, the average number of answers other than "I don't know" was averaged and taken as measure of preference insight.

The first alternative measure refers to preference formation time, building on the finding that individuals with high preference insight are able to retrieve information from memory faster than participants with low preference insight, as preferences are more readily accessible in the former case (Fazio, Chen, McDonel, and Sherman, 1982). The time each participant needed to answer the three preference questions was thus precisely measured (indiscernible to participants). The individual mean response time is the second measure of preference insight.

The second alternative method refers to preference confidence. The underlying idea is that individuals with high preference insight will be more confident in their decisions than those with low preference insight (Chernev, Mick, and Johnson, 2003). The test resembles the classic "theater test" technique used in market research: The instructor told each participant that he would briefly show them two photos of watches. The participants would then have to decide quickly which one they would like to win in the raffle. The instructor then showed the photos for five seconds, the participants made their decisions, and the instructor wrote them down. After that, it was tested whether the participants would correct their decisions if the conditions were altered. In order not to reveal this objective to participants and to reduce the risk of a demand effect (Sawyer, 1975), a "trick" of sorts was employed: When noting the participants' decision, the instructor mumbled to himself: "Hmm, I thought so." He then added in an intimate tone, as if he were leaving his official role as instructor for a moment to give the participant a friendly tip: "You know, most of the participants so far have chosen the same watch type as you did. And because we have only one of each, the chance of winning the other model is about double. If you want, you can switch to the other one." Naturally, all participants received the same tip in a standardized tone and wording regardless

of the watch they had chosen. (The participants' reactions and comments clearly showed that the instructor played this "helpful" role very convincingly.) The instructor noted whether the participant changed his or her decision. The argument is that persons with low preference insight will have had difficulties devising a clear and stable preference order during the short inspection time (Hoeffler and Ariely, 1999). The probability that they would be willing to switch in exchange for a better chance of winning should be higher than among those participants who have clear preference insight and had thus developed a (more) firm preference order. The dummy variable "switch" or "no switch" was thus the third measure of preference insight. The watch designs and the stimulus in the course had been chosen based of two pilot studies ($n = 5$ and 12).

Assessing Value Generation by Increased Preference Insight

If the measurement of the test-based approach proves to be valid, one problem remains: It is difficult to assess if the increase in preference insight generated by self-designing with a configuration toolkit (as measured in Study 1) has any practical import. The properties of the test-based scale are unclear, and despite its statistical significance, it is unknown if the increase from .62 to .76 is a large one or not. This prompted the decision to translate it into a monetary measure.

The rationale is as follows: For a given individual, the opportunity to self-design a product should have a subjective value that is contingent upon the individual's specific preference insight (Bharadwaj et al., 2009; Franke et al., 2009). If this preference insight is low, it means that the subject does not know what he or she wants. Configuring a product according to one's own preferences is difficult, and the outcome is uncertain. If this individual was offered the possibility of buying a coupon for a product he or she can self-design, his or her willingness to pay (WTP) should be limited (in line with extant literature, the subjective value individuals attribute to a product is conceptualized as WTP, e.g., Sinha and Mandel, 2008). If, on the other hand, the individual's preference insight has increased, self-designing should be more attractive. Knowing what he or she wants, the subject can make better use of the opportunity to specify the product design. If offered a coupon for doing so, his or her WTP should in turn be higher. The difference in WTP in the two situations can be interpreted as the monetary value of gaining additional preference insight.

One hundred and six business students from the authors' university (average age: 23.58; 56.6% females) participated in this experiment. The incentives announced were 10 euros for every participant and the chance to win "products" in a raffle (as in Study 1, task and prizes were not revealed in order to avoid self-selection bias). The instructions were standardized and presented in the same way as in Study 1. The participants were randomly assigned to three experimental groups and seated in separate rooms throughout the entire experiment (including all measurements).

Treatment group (self-design). Participants in this group were instructed to self-design a watch with a configuration toolkit (from 121Time) which enables users to individualize wristwatches and offers very realistic visual design feedback as well as a large solution space. Participants managed to complete the self-design process in 9.6 minutes on average ($SD = 2.9$). As in the other studies, they had an incentive to take this task seriously, as they were informed that there would be a raffle in which they could win their respective self-designed watch. Once they had finished, their preference insight was measured using the three measures of (1) our preference test, (2) the precisely measured response time, and (3) switching behavior in the "theater test," a number of control variables, and the value they expected to derive from future watch self-design (WTP). For the last value, participants were offered a coupon for a watch configuration toolkit named "designAwatch" (valid for 12 months). They learned that the coupon would allow them to obtain a watch self-designed with this toolkit for free. The toolkit shown to them was, of course, very different from the one they had used before. Unlike 121Time, this toolkit included design options like dial backgrounds, dial markings, color options, material options, and functionality options. Had the toolkits been similar, there would have been an alternative explanation for the Treatment Group bidding higher than the following groups, namely that the increase in subjective valuation might be caused not by enhanced preference insight but by a higher level of familiarity with the *specific toolkit* and the lower risk perceived as a result. We therefore took great care in handling this aspect. Because existing toolkits were not satisfactory (i.e., too similar), a simulated toolkit was devised. Graphics software was used to create a "screenshot" of a fictitious toolkit with every reasonable design option one can imagine. In order to test whether the simulated toolkit was perceived as different from 121Time, a pilot study with $n = 33$ persons was conducted in which participants inspected both toolkits and

then were asked a number of questions. Findings confirm that the two toolkits are indeed perceived as quite different: The two items “The two toolkits appear similar” ($M = 2.21$, $SD = 1.19$, median = 2.00) and “If I designed a watch with toolkit A [121Time], it would give me an advantage regarding the functionality of toolkit B [designAwatch]” ($M = 2.94$, $SD = 1.75$, median = 2.00) met with fairly low agreement, while the item “The two toolkits are very different in structure” ($M = 6.00$, $SD = 1.03$, median = 6.00) saw a high level of agreement (all items 1 = “strongly disagree,” 7 = “strongly agree”). Participants then were informed that they could bid on the coupon and explained the principle of the BDM auction (Becker, Degroot, and Marschak, 1964) using an illustration. In such an auction, participants can submit binding bids for a product, after which they draw a card from an urn. If the individual participant’s bid is higher than or equal to the price on the card, they are required to purchase the product at the price indicated on the card. If the participant’s bid is lower, they cannot purchase the product. Test questions revealed that the participants understood the procedure well. Each participant was asked to write his or her bid down on a form and sign it. After the experiment, it was revealed that for technical reasons they could not get the coupons, and that they would be compensated with a raffle for a free watch from 121Time. All participants were content with this arrangement.

Control group 1 (alternative product shopping). Participants in this group were given a distraction task: They had to select a T-shirt from about 1000 designs presented in the Threadless online shop, which took them 8.8 minutes on average ($SD = 4.6$). As in the former groups, participants were informed that they could win the T-shirt which best matched their preferences. The measurement procedure was the same as in the Treatment Group and included their preference insight regarding watches (again by the three measures), a number of control variables, and WTP for the “designAwatch” coupon. As this control group did not self-design a watch with a configuration toolkit before, their preference insight with regard to watches should correspond to that of the Treatment Group before self-designing (given valid randomization and no influence on the part of the distraction task). A significant difference in preference insight between the two groups would thus replicate Study 1 in a different product context. More importantly, the difference in WTP for the self-design coupon should reflect the monetary value of the increase in preference insight (caused by self-designing).

Control group 2 (alternative product customization). However, there is one important alternative explanation. Participants in the Treatment Group (Customization) may have not only improved their preference insight but also learned more about the *general possibilities* of self-designing products with configuration toolkits. It is important to bear in mind that, despite the popularity of mass customization in the media, most consumers have not yet self-designed a product in this way. As a result, it is necessary to rule out the alternative explanation that the better knowledge of configuration toolkits as such—and not the increased insight into one’s own preferences—is causal for the (potential) increase in subjective value. Therefore, this second control group was set up. Participants in this group were instructed and questioned in a way similar to participants in the Treatment Group, with the sole differences being the toolkit and product category. This group used a toolkit to self-design sneakers (Vans Custom Shoes) that had been chosen because a pilot study with $n = 10$ participants had revealed that the complexity, the size of the solution space, and the feedback information displayed were comparable to those of the toolkit used by the Treatment Group. If the alternative explanation does not apply, Control Group 2 should exhibit lower preference insight (their preference learning with regard to watches had not been stimulated) and in particular attribute a lower WTP to the self-design coupon than the Treatment Group. Most importantly, a mediator analysis should confirm that preference insight (and not knowledge about toolkits) explains the increase in WTP for the coupon in the Treatment Group relative to the other two groups.

Measurement

The following control variables that might also influence preference insight were measured (see Appendix): (1) design affinity (five items, $\alpha = .81$), (2) design experience (one item), (3) involvement (seven items, $\alpha = .89$, adapted from Zaichkowsky, 1985), (4) innovativeness (five items, $\alpha = .67$, adapted from Manning, Bearden, and Madden, 1995), (5) need for uniqueness (nine items, $\alpha = .89$, adapted from Tian, Bearden, and Hunter, 2001), (6) experience with watches (four items, $\alpha = .73$), (7) recently bought a watch/received a watch as a present (one item). Multi-item constructs were averaged. In order to control for the alternative explanation that participants simply developed higher knowledge or confidence regarding (watch) toolkits, five additional control variables covering different aspects of toolkit-related knowledge and confidence were

used (knowledge about toolkit handling, beliefs about user friendliness, technical knowledge assumed necessary, fun expected when using a toolkit, confidence in ability to use a toolkit successfully). All variables were measured on seven-point scales ranging from “strongly disagree” to “strongly agree.” Additionally, age, gender, and treatment time was measured. Finally, participants were asked to indicate their discretionary income, as it would obviously be an important predictor of WTP. There was no difference between the experimental groups in any of these control variables.

Findings

Validation of preference insight measurement. The first analysis relates to the extent to which the main method of asking test questions yielded results similar to those generated by the other two methods. The correlations are significant and point in the expected direction. The preference test is negatively correlated with response time ($r = -.17, p < .05$) and with switching behavior ($r = -.17, p < .05$). Both alternative measures are also significantly correlated ($r = .19, p < .05$). This allows the conclusion that the original measurement of preference insight possesses sufficient validity.

Value generation by increased preference insight. An analysis of means shows that the participants in the

Treatment Group (self-design) exhibit significantly higher preference insight in all three measures (see Table 2). This replicates the confirmation of H1 (Study 1) in a different product category, with a different toolkit, and with alternative measures. The increased preference insight translates into a higher valuation of the (future) opportunity to self-design a product: WTP for the coupon was significantly higher in the Treatment Group than in the other two groups (WTP = 107.77 euros versus 80.47, $p < .05$ and WTP = 64.81 euros, $p < .01$). The effect sizes are considerable. Compared to the participants who did not learn about their preferences by self-designing a watch, WTP increased by 66% relative to Control Group 1 (alternative product shopping) and 34% relative to Control Group 2 (alternative product self-design).

In addition, a mediator analysis was employed in which it was checked whether the increase in WTP is caused by additional preference insight (Dummy Variable 1: Treatment Group versus the other two groups). In order to control for the alternative explanation that enhanced knowledge about configuration toolkits might instead cause the increase in WTP, another dummy variable (Dummy Variable 2: Treatment Group and Control Group 2 versus Control Group 1), and all control variables that captured the participants' enhanced knowledge of toolkits were included. (The mediator analysis procedure is based on the method proposed by Baron and Kenny,

Table 2. Descriptive Data on Preference Insight/WTP/Knowledge and Beliefs about Toolkits—Measures

	Treatment Group (self-design) ($n = 35$) M (SD)	Control Group 1 (alternative product shopping) ($n = 36$) M (SD)	Control Group 2 (alternative product self-design) ($n = 34$) M (SD)	Post hoc tests (from an ANOVA/ANCOVA)	
				TG versus CG 1	TG versus CG 2
Preference insight:					
Preference test	.96 (.13)	.79 (.27)	.81 (.20)	$p < .001^a$	$p < .01^a$
M response time	1.17 (.96)	2.40 (1.27)	2.37 (1.56)	$p < .001^a$	$p < .001^a$
Switch of watch	.26 (.44)	.58 (.50)	.41 (.50)	$p < .01^a$	$p < .1^a$
Willingness to pay for self-design coupon (different watch toolkit)	107.77 (126.97)	64.14 (58.23)	80.47 (63.59)	$p < .01^b$	$p < .05^b$
Knowledge and beliefs about toolkits:					
Handling	2.11 (.87)	2.14 (1.15)	2.03 (.87)	n.s. ^c	n.s. ^c
User friendliness	2.21 (1.18)	2.47 (1.18)	2.21 (1.12)	n.s. ^c	n.s. ^c
Technical knowledge	5.46 (1.29)	4.81 (1.43)	5.00 (1.35)	$p < .05^3$	n.s. ^c
Fun	3.89 (.99)	3.39 (1.18)	3.79 (1.01)	$p < .05^3$	n.s. ^c
Confidence	4.43 (.82)	4.39 (.80)	4.38 (.60)	n.s. ^c	n.s. ^c

^a ANCOVA with covariates: Design experience, innovativeness, treatment time, age, gender, involvement, need for uniqueness, design affinity.

^b ANCOVA with additional covariates: Income, knowledge, and beliefs about toolkit aspects (handling, user friendliness, required technical knowledge, fun, confidence), recently bought a watch/received a watch as a present.

^c ANOVA without covariates.

n.s., not significant.

Table 3. Mediation of the Influence of Toolkit Use on WTP via Preference Insight

	Model 1 DV: Preference insight	Model 2 DV: WTP	Model 3 DV: WTP	Model 4 DV: WTP
Group variable (Treatment Group)	.335**	.181*		.113
Group variable (Treatment Group and Control Group 2)	.039	.067		.061
Preference insight			.265**	.218*
Control variables:				
Income	-.053	.326***	.344***	.343***
Design experience	-.024	.219*	.214*	.225*
Innovativeness	-.003	.209*	.208*	.211*
Involvement	-.049	.208*	.192*	.185*
Treatment time	-.041	.187*	.211*	.198*
R ²	.155	.317	.338	.355
Additional covariates (NS)	Age, gender, involvement, need for uniqueness, design affinity	Age, gender, involvement, need for uniqueness, design affinity, knowledge and beliefs about toolkits (handling, user friendliness, required technical knowledge, fun, confidence), recently bought a watch/received a watch as a present		

* $p < .05$.
 ** $p < .01$.
 *** $p < .001$.

1986, which has been used more recently by e.g., Goukens, Dewitte, Pandelaere, and Warlop, 2007.)

Model 1 (Table 3) is a regression model with preference insight (as measured by the test questions) as the dependent variable and the group variables as independent variables. It becomes evident that the group variable of customization (Treatment Group) is the only significant predictor of preference insight, which mirrors the descriptive findings above. Model 2 also uses both group variables as independent variables, but it uses WTP for the toolkit coupon as the dependent variable. Using a watch-specific toolkit (group variable, Treatment Group) significantly increases WTP. In order to test whether this increase is caused by enhanced preference insight, Models 3 and 4 are examined. In Model 3, preference insight is the predictor of WTP, although additional control variables that may influence WTP, including knowledge and beliefs about toolkits, had been incorporated. As predicted, the results show that preference insight has a positive influence on WTP. Model 4 includes both group variables. While preference insight remains significant in this model, both group variables are now insignificant. This can be seen as clear evidence that preference insight does mediate the positive relationship between having self-designed with a toolkit and WTP for self-design in the same product category. In fact, it is the only significant mediator. A Sobel test for this mediation was significant ($p < .05$). Again, results remain robust when covariates are excluded.

In summary, the findings from Study 1 were replicated and validated in three major ways: (1) There is clear evidence that the test-based measure is valid, (2) the effect of self-design on preference insight could be replicated successfully with another toolkit and in another product category, and (3) it becomes apparent that the preference learning effect is quite substantial in terms of the subjective monetary value generated for the individual.

Study 3: Does Preference Learning by Self-Designing Follow a “Power Law of Practice” Pattern?

Study 3 is a within-subject experiment designed to test H3. By using a truly representative sample from an online panel, it also allows us to overcome a limitation of Studies 1 and 2 not yet mentioned, namely the use of a student sample. For this experiment, the 121Time toolkit was used, which allows users to self-design wristwatches.

Procedure, Sample, and Experimental Groups

The data were obtained from a self-administered online questionnaire. A random sample was drawn from the leading national online panel, which is nationally representative of Austrian residents with an e-mail account. Sample size was increased because the online setting is less controlled, which inevitably results in more “noise”

Table 4. Moderators of Preference Insight Learning

	Model 1 DV: Delta preference insight	Model 2 DV: Delta preference insight
Perceived user friendliness of toolkit	-.120*	-.122*
Design time	.108*	.111*
General learning skills	.111*	.109*
Prior level of preference insight (Prior level of preference insight) X (Prior level of preference insight)	-.458***	-.393***
R ²	.248	.251

* $p < .05$.*** $p < .001$.

and thus reduced effect sizes (Dandurand, Shultz, and Onishi, 2008). A total of 920 panel participants were contacted and asked to fill out the questionnaire; 310 participants answered the questionnaire completely, which represents a response rate of 34%. Comparisons of early and late respondents showed no significant differences, indicating the absence of response bias (Armstrong and Overton, 1977). The mean age of the participants was 38.3 years ($SD = 12.7$); 50% were female. Participants had the opportunity to enter a raffle for “valuable products” which were not specified any further. The questionnaire started with an initial measurement of preference insight regarding watches. Participants were then informed that they could win a watch to be self-designed in no more than 20 minutes (the toolkit was embedded in the questionnaire and included a timer). The mean design time was 7.54 minutes ($SD = 6.47$). After the self-design process, preference insight (and a number of additional variables) was measured again.

Measurement

Preference insight. Preference insight was measured using the method applied in Studies 1 and 2. Participants were asked about the ideal watch design based on a number of dimensions and took the number of answers other than “I don’t know” as an indicator of preference insight. Since the questions were not asked by an instructor but presented on a computer screen, a number of closed answers were offered to choose from, including “I don’t know.” The questions from Studies 1 and 2 were adapted in this way, and five questions were added, meaning that a total of eight preference insight questions was used in this study. In order to avoid memory effects, the same scheme as in Study 1 was employed, meaning that the questions during pre- and post measurement were randomized. Each participant received a random selection of four (out of the eight) items for premea-

surement and the remaining four items for post-measurement. The items were corrected in the same way as in Study 1.

Alternative explanations for preference learning. In order to avoid an omitted variable bias, three additional variables that are considered to be established predictors of preference learning were included (Ackerman, 1987; Smith, 1996; von Hippel and Katz, 2002): (1) perceived user friendliness of the toolkit, measured by the item “How difficult did you perceive the design process to be with the 121Time toolkit?” (1 = “very easy,” 7 = “very complex”); (2) general learning skills, measured using five self-developed items (e.g., “I was a good student,” 1 = “strongly disagree,” 7 = “strongly agree,” Cronbach’s $\alpha = .82$) and combined in an averaged index; and (3) design time based on log data.

Findings and Discussion

H3 stated that individuals with low initial preference insight would exhibit stronger learning effects than those who already had some understanding of what they wanted. An OLS regression model was used for the test (see Table 4). The independent variables were the prior level of preference insight and the control variables of perceived difficulty, learning skills, and design time. The dependent variable was the individual delta of preference insight before and after treatment (i.e., the learning effect). A Goldfeld–Quandt test showed evenly distributed residuals ($F(153,153) = 3.26$, NS), thus homoscedasticity could be assumed.

Findings clearly confirm H3: The lower the prior level of preference insight is, the greater the learning effect of self-designing with a toolkit becomes ($b = -.458$, $p < .001$). The learning effect is stronger in effect size and exists independently of the other factors that also have a significant impact on preference learning (perceived

user friendliness $b = .120$, $p < .05$, general learning skills $b = .111$, $p < .05$, design time $b = .108$, $p < .005$, $R^2 = .248$). There is no evidence of an S-shaped learning curve, as the squared term in Model 2 is insignificant.

General Discussion

Most research on mass customization so far has viewed toolkits merely as the technical interface via which preexisting individual product design preferences are conveyed to the producer. While much extant research simply assumes that consumers have clear enough preferences to do so, more psychologically oriented scholars have recently warned that decades of consumer research have revealed rather low levels of preference insight in most consumers (Kramer, 2007; Simonson, 2005).

Configuration Toolkits as Learning Instruments

The findings reported in this paper suggest that configuration toolkits should be interpreted as learning instruments that help consumers discover their own preferences. They can be more than just technical interfaces for conveying individual preference information. The effects of single self-design processes lasting only a few minutes with toolkits that were not even specifically designed for learning purposes are remarkable: Enhanced preference insight levels remain stable over two weeks and generate substantial value for the customer, measured in hard currency. These findings complement recent research on customers' affective reactions evoked by interaction with a toolkit, such as product (Moreau and Herd, 2010; Valenzuela et al., 2009) and process satisfaction (Dellaert and Stremersch, 2005), pride and feelings of accomplishment (Franke et al., 2010), and enjoyment (Franke and Schreier, 2010). Patterns found suggest that interacting with a toolkit also elicits *cognitive* reactions, namely the knowledge of what one wants. It would be intriguing to study both affective and cognitive reactions simultaneously, i.e., to measure the extent to which interaction with a toolkit not only increases preference insight, but also changes preferences, and which affective factors motivate and result from both preference insight and change.

However, given that preference learning effects exist and that they are substantial, the obvious next research step would be to investigate what exactly triggers them and how they can be enhanced. Literature on learning psychology offers a rich source of principles and techniques to make learning easier and more effective (e.g., Jarvis, 2005). From the many possibilities, two appear of

paramount importance. First, it is known from many studies that *feedback* is essential to learning (see e.g., Butler and Winne, 1995), and hence the better the feedback provided by the toolkit is, the better the customer's preference learning will be. But what is good feedback in the case of self-design toolkits? Most existing toolkits restrict feedback to instant 2-D visual feedback and price information. Would 3-D visualization or a simulation of the product in the customer's individual use environment facilitate better preference learning? Would "expert" feedback help? Such feedback could refer to chosen options and combinations, esthetic, and functional aspects of the self-designed product, and its fit to the person (Randall et al., 2005). This kind of feedback could be generated by trained company employees, similar to feedback by good salespersons at the point of sale. It could also be generated automatically on the basis of artificial intelligence. The great progress in this field has opened up new and innovative possibilities that deserve further theoretical and empirical analyses. Another particularly promising method would be to include a function that allows self-designers to submit their (interim) design solutions for rapid "social" feedback from other users who are online. Combining the concept of configuration toolkits with social networks—both existing ones and communities purposefully initiated by the firm—appears to be a powerful yet hardly explored idea. Initial research suggests that such peer feedback can be very helpful for self-design (Franke et al., 2008). Gaining a better understanding of which kind of feedback information will have the greatest learning effects on which type of customer in a given situation and product category constitutes a fascinating trajectory for further research.

The second imperative from learning psychology is to ensure that the participant's *capabilities* and the *difficulty* of the task are well-matched (Bandura, 1993; Winne, 1997). If a task is too complex and difficult, the individual will react with frustration and is not likely to learn much. If, on the other hand, the task is too easy, the individual will get bored, and the learning effect might be equally poor (Pekrun, Goetz, Titz, and Perry, 2002). "Task difficulty" in the context of a toolkit refers to how many decisions have to be made and how complex those decisions are. They encompass e.g., the number of design or preference dimensions, attributes and attribute levels offered, and (beyond choice decisions) the extent to which customers are enabled or required to provide creative input themselves, as in the process of creating a design with design tools. Studies by Huffman and Kahn (1998), Valenzuela et al. (2009), Dellaert and Stremersch

(2005), and Randall et al. (2007) provide important initial insights on these issues, although those authors use product utility and process satisfaction—not learning effects—as dependent variables. As prior levels of preference insight and learning capabilities will vary among consumers, would a “graded” configuration toolkit—in which customers decide for themselves how many design options, feedback details, etc., they want to tackle—be a solution? As Randall et al. (2005) note, it is somewhat perplexing in any case that although customization strategies build on the insight that people have very different preferences and may therefore attach value to very different products, most firms deliver only one standardized self-design toolkit.

Mass, Not Niche Customization

The findings suggest that the segment of consumers for whom mass customization potentially creates value is quite large. It comprises not only those rare consumers with clear insight in their preferences. Also those individuals who do not yet know what they want can benefit from self-designing as they may learn and improve their preference insight. Thus, if producers interpret configuration toolkits as learning instruments, this warrants a more optimistic prediction regarding the future of customer cocreation, namely that it holds the potential to actually become *mass* customization in the strict sense of the term.

To this end, two preconditions must be fulfilled: First, mass customization providers need to build configuration toolkits in such a way that they actually support preference learning as much as possible. How can this be achieved? While inspecting existing toolkits on the web, one easily gains the impression that the learning objective has largely been neglected. However, as Hoch and Deighton (1989) put it, “Learning must be accounted for not as something independent of marketing action, but as a process that marketing has the power to leverage” (p. 16), and this surely holds for toolkits as well. This is where research interests and managerial interests meet, and managers are encouraged to take up the measures suggested in the former section. Second, toolkit providers must overcome a Catch-22 problem. Findings show that customers with a low level of preference insight initially expect to derive little value from using configuration toolkits and might therefore hesitate to start such self-design processes. This value assessment can be increased considerably but not until customers have actually begun a self-design process, as this would increase their preference insight. So, how can consumers with a low level of

preference insight be motivated to start a self-design process if the level of preference insight achieved by the self-design process is actually a prerequisite for starting it? Certainly not by advertising toolkits with the message “Create your personal dream product,” as the standard argument seems to be. In fact, such a communication strategy risks evoking the unspoken response, “So it’s not for me—I don’t know what my dream product should look like.” When advertising toolkits, firms should therefore communicate clearly that self-designers do not necessarily need to know exactly what they want from the outset, but will instead be enabled to discover their preferences without difficulty in the course of the self-design process. Thus, messages like “Find out which product is really best for you” seem far more promising. Another idea would be to work on lowering the entry barriers in order to encourage consumers to try self-designing even if they have never considered it before. For example, this could be achieved by combining the toolkit with an online shop. Standard product descriptions could include a “further customize the product” option, which might stimulate customers to try modifying a product, essentially using the standard product as a starting solution for their own (initially unplanned) self-design activities. A third possibility for firms would be to emphasize the fun aspect of self-designing rather than the potential benefit of obtaining a product perfectly tailored to one’s preferences, as research has shown that this is an important side effect of toolkit use (Franke and Schreier, 2010). To the authors’ knowledge, there is no extant research that systematically analyzes the factors which prompt consumers to consider trying out self-design toolkits or prevent them from doing so.

Are Our Conclusions Limited to Self-Expressive Goods?

This research project is restricted to configuration toolkits that allow the esthetic design of individual watches and running shoes. These products are self-expressive goods, and it appears plausible to generalize findings to mass customization configurators for other products in this category, such as T-shirts, skis, cell phones, furniture, or jewelry. From a strict logical perspective it is not possible to infer that also configuration toolkits for individual *utilitarian* goods (e.g., laptops or stereo systems) hold the potential to bring about similar learning effects among consumers. It appears, however, that this is quite likely.

First, consumers might also lack preference insight regarding utilitarian products (Simonson, 2005). If consumers buy e.g., a laptop, a mattress, or a refrigerator

for the first time, they often do not know which functionalities and criteria matter in general and which are important particularly for their purposes (Mourali, Laroche, and Pons, 2005). In other words, the preference insight of many customers is low. In such situations, learning what one wants might be helpful as it obviously improves the quality of the decision. Feedback that would confront the customer with simulated consequences of choices and combinations might induce such learning processes. Von Hippel and Katz (2002) provide anecdotal evidence of such effects. They document that toolkits for individualized telephone-answering systems and computer chips (that both have a clear utilitarian character) facilitate a learning-by-doing process enabling customers to find out what they want. Of course the feedback in such toolkits must be quite different from the mainly visual representation of the self-designed product in esthetic self-design of self-expressive goods. If the good is typified mainly by its functionality (and not by its esthetics), then feedback of course must refer to this. For example, a gardening toolkit that allows creating one's own garden may give some sort of alarm when the customer positions a garden pond too close to a broadleaf tree: "in fall leaves might fall on the pond—they might quickly silt up your pond, in summer this might provide too much shade and so stop aquatic plants from growing." It is clear that sometimes such feedback is difficult or in some instances even impossible to implement in the configurator with today's technology. For example, systems that allow customers to configure their individual hamburgers, cereals, or pizza simply cannot come up with the immediate simulated taste of any interim combination considered. However, in principle the possibility of trials and feedback information upon these trials will result in enhanced preference insight also in utilitarian product categories. In many instances also today's technology will allow users to incorporate such feedback in configuration toolkits or simulate it using "social" feedback from other users. Again, it must be noted that this line of argumentation is not backed by empirical evidence. Thus, research that studies learning effects in utilitarian product categories and investigations in how far configuration toolkits can also serve as preference learning instruments also beyond self-expressive goods would be of great value.

Other Limitations

This study is not free of other limitations, which constitute additional opportunities for research. The most critical point might still be the measurement of preference insight, the focal construct of this paper. By using a

test-based approach validated with two independent methods, the limitations of self-assessment approaches were overcome. As further research in this area will depend heavily on the valid measurement of preference insight, however, refining these measurement techniques might constitute an attractive research question in its own right. Advances in neuroscience may also offer promising new possibilities (Lee, Amir, and Ariely, 2009).

References

- Ackerman, P. L. 1987. Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin* 102 (1): 3–27.
- Armstrong, J. S., and T. S. Overton. 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research* 14 (August): 396–402.
- Bandura, A. 1993. Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist* 28 (2): 117–48.
- Bardakci, A., and J. Whitelock. 2003. Mass-customisation in marketing: The consumer perspective. *Journal of Consumer Marketing* 20 (5): 463–79.
- Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–82.
- Bechtel, W., and A. Abrahamsen. 2002. *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks*. Malden, MA: Blackwell Publishing.
- Becker, G. M., M. H. Degroot, and J. Marschak. 1964. Measuring utility by a single-response sequential method. *Behavioral Science* 9 (3): 226–32.
- Bettman, J. R., M. F. Luce, and J. W. Payne. 1998. Constructive consumer choice processes. *Journal of Consumer Research* 25 (3): 187–217.
- Bettman, J. R., M. F. Luce, and J. W. Payne. 2008. Preference construction and preference stability: Putting the pillow to rest. *Journal of Consumer Psychology* 18 (3): 170–74.
- Bharadwaj, N., R. W. Naylor, and F. ter Hofstede. 2009. Consumer response to and choice of customized versus standardized systems. *International Journal of Research in Marketing* 26 (3): 216–27.
- Butler, D. L., and P. H. Winne. 1995. Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research* 65 (3): 245–81.
- Chernev, A., D. G. Mick, and M. D. Johnson. 2003. When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of Consumer Research* 30 (2): 170–83.
- Cook, S. 2008. The contribution revolution. *Harvard Business Review* 86 (10): 60–69.
- Dandurand, F., T. R. Shultz, and K. H. Onishi. 2008. Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods* 40 (2): 428–34.
- Dellaert, B. G. C., and S. Stremersch. 2005. Marketing mass-customized products: Striking a balance between utility and complexity. *Journal of Marketing Research* 42 (2): 219–27.
- Deng, X. J., and J. W. Hutchinson. 2009. Aesthetic self-design: Just do it yourself. Working Paper. Columbus: Ohio State University, Fisher College of Business.
- Fazio, R. H., J.-M. Chen, E. C. McDonel, and S. J. Sherman. 1982. Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology* 18 (4): 339–57.
- Fitzsimons, G. J., and V. G. Morwitz. 1996. The effect of measuring intent on brand-level purchase behavior. *Journal of Consumer Research* 23 (1): 1–11.

- Forgas, J. P. 1995. Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin* 117 (1): 39–66.
- Franke, N., P. Keinz, and M. Schreier. 2008. Complementing mass customization toolkits with user communities: How peer input improves customer self-design. *Journal of Product Innovation Management* 25 (6): 546–59.
- Franke, N., P. Keinz, and C. J. Steger. 2009. Testing the value of customization: When do customers really prefer products tailored to their preferences? *Journal of Marketing* 73 (5): 103–21.
- Franke, N., and F. T. Piller. 2003. Key research issues in user interaction with user toolkits in a mass customization system. *International Journal of Technology Management* 26 (5-6): 578–99.
- Franke, N., and F. T. Piller. 2004. Value creation by toolkits for user innovation and design: The case of the watch market. *Journal of Product Innovation Management* 21 (6): 401–15.
- Franke, N., and M. Schreier. 2010. Why customers value mass-customized products: The importance of process effort and enjoyment. *Journal of Product Innovation Management* 27: 1020–31.
- Franke, N., M. Schreier, and U. Kaiser. 2010. The “I designed it myself” effect in mass customization. *Management Science* 65 (January): 125–40.
- Ghosh, M., S. Dutta, and S. Stremersch. 2006. Customizing complex products: When should the vendor take control? *Journal of Marketing Research* 43 (4): 664–79.
- Goukens, C., S. Dewitte, M. Pandelaere, and L. U. K. Warlop. 2007. Wanting a bit(e) of everything: Extending the valuation effect to variety seeking. *Journal of Consumer Research* 34 (3): 386–94.
- Gownder, J. P. 2011. Mass customization is (finally) the future of products. Forrester Report.
- Haider, H., and P. A. Frensch. 2002. Why aggregated learning follows the power law of practice when individual learning does not: Comment on Rickard (1997; 1999), Delaney et al. (1998), and Palmeri (1999). *Journal of Experimental Psychology: Learning, Memory & Cognition* 28 (2): 392–406.
- Hoch, S. J., and J. Deighton. 1989. Managing what consumers learn from experience. *Journal of Marketing* 53 (2): 1–20.
- Hoeffler, S., and D. Ariely. 1999. Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology* 8 (2): 113–39.
- Huffman, C., and B. E. Kahn. 1998. Variety for sale: Mass customization or mass confusion? *Journal of Retailing* 74 (4): 491–513.
- Jackson, D. N. 1983. *Jackson personality inventory—Revised*. Port Huron, MI: Sigma Assessment Systems, Inc.
- Janiszewski, C., and S. M. J. van Osselaer. 2000. A connectionist model of brand-quality associations. *Journal of Marketing Research* 37 (3): 331–50.
- Jarvis, M. 2005. *The psychology of effective learning and teaching*. Cheltenham, UK: Nelson Thornes.
- Johnson, E. J., S. Bellman, and G. L. Lohse. 2003. Cognitive lock-in and the power law of practice. *Journal of Marketing* 67 (2): 62–75.
- Keller, K. L. 1993. Conceptualizing, measuring, managing customer-based brand equity. *Journal of Marketing* 57 (1): 1–22.
- Kotha, S. 1995. Mass customization: Implementing the emerging paradigm for competitive advantage. *Strategic Management Journal* 16 (Summer): 21–42.
- Kramer, T. 2007. The effect of measurement task transparency on preference construction and evaluations of personalized recommendations. *Journal of Marketing Research* 44 (2): 224–33.
- Kruger, J., and D. Dunning. 1999. Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality & Social Psychology* 77 (6): 1121–34.
- Kwon, K., J. Cho, and Y. Park. 2009. Influences of customer preference development on the effectiveness of recommendation strategies. *Electronic Commerce Research & Applications* 8 (5): 263–75.
- Lee, L., O. N. Amir, and D. Ariely. 2009. In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research* 36 (2): 173–87.
- Liechty, J., V. Ramaswamy, and S. H. Cohen. 2001. Choice menus for mass customization: An experimental approach for analyzing customer demand with an application to a web-based information service. *Journal of Marketing Research* 38 (2): 183–96.
- Loewenstein, G. F., and D. Prelec. 1993. Preferences for sequences of outcomes. *Psychological Review* 100 (1): 91–108.
- Manning, K. C., W. O. Bearden, and T. J. Madden. 1995. Consumer innovativeness and the adoption process. *Journal of Consumer Psychology* 4 (4): 329–45.
- Marketing Science Institute. 2008. *Research priorities*. Cambridge, MA: Marketing Science Institute.
- McClelland, J. L., and D. E. Rumelhart. 1986. *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2*. Cambridge, MA: MIT Press.
- McKenna, R. 2002. *Giving customers what they want in an anytime, anywhere world*. Boston, MA: Harvard Business School Press.
- Moreau, C. P., L. Bonney, and K. B. Herd. 2011. It’s the thought (and the effort) that counts: How customizing for others differs from customizing for oneself. *Journal of Marketing* 75 (5): 120–33.
- Moreau, P. C., and K. B. Herd. 2010. To each his own? How comparisons with others influence consumers’ evaluations of their self-designed products. *Journal of Consumer Research* 36 (5): 806–19.
- Mourali, M., M. Laroche, and F. Pons. 2005. Antecedents of consumer relative preference for interpersonal information sources in pre-purchase search. *Journal of Consumer Behaviour* 4 (5): 307–18.
- Newell, A., and P. S. Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. In *Cognitive skills and their acquisition*, ed. J. R. Anderson, 1–55. Hillsdale, NJ: Lawrence Erlbaum Associates.
- O’Hern, M., and A. Rindfleisch. 2009. Customer co-creation: A typology and research agenda. *Review of Marketing Research* 6: 84–106.
- Park, C. W., S. Y. Jun, and D. J. MacInnis. 2000. Choosing what I want versus rejecting what I do not want: An application of decision framing to product option choice decisions. *Journal of Marketing Research* 37 (2): 187–202.
- Payne, A. F., K. Storbacka, and P. Frow. 2008. Managing the co-creation of value. *Journal of the Academy of Marketing Science* 36 (1): 83–96.
- Pekrun, R., T. Goetz, W. Titz, and R. P. Perry. 2002. Academic emotions in students’ self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist* 37 (2): 91–106.
- Pine, J. B. 1999. *Mass customization—The new frontier in business competition*. Boston, MA: Harvard Business School Press.
- Pine, J. B., D. Peppers, and M. Rogers. 1995. Do you want to keep your customers forever? *Harvard Business Review* 73 (2): 103–14.
- Prahalad, C. K., and V. Ramaswamy. 2004. *The future of competition: Co-creating unique value with customers*. Boston, MA: Harvard Business School Press.
- Randall, T., C. Terwiesch, and K. T. Ulrich. 2005. Principles for user design of customized products. *California Management Journal* 47 (4): 68–85.
- Randall, T., C. Terwiesch, and K. T. Ulrich. 2007. User design of customized products. *Marketing Science* 26 (2): 268–80.
- Ritter, F. E., and L. J. Schooler. 2004. The learning curve. In *International encyclopedia of the social & behavioral sciences*, Vol. 13, ed. N. J. Smelser and P. B. Baltes, 8602–5. Amsterdam: Elsevier.
- Rodriguez, T. J. 2009. *Understanding human behavior*. Manila: Rex Book Store.
- Rumelhart, D. E., and J. L. McClelland. 1986. *On learning the past tenses of English verbs*. Cambridge, MA: MIT Press.
- Ryan, R. M., and E. L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55 (1): 68–78.

- Sawyer, A. G. 1975. Demand artifacts in laboratory experiments in consumer research. *Journal of Consumer Research* 1 (4): 20–30.
- Schreier, M. 2006. The value increment of mass-customized products: An empirical assessment. *Journal of Consumer Behaviour* 5 (4): 317–27.
- Seybold, P. 2006. *Outside innovation: How your customers will co-design your company's future*. New York, NY: Collins.
- Sheth, J. N., R. S. Sisodia, and A. Sharma. 2000. The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science* 28 (1): 55–66.
- Simonson, I. 2005. Determinants of customers' responses to customized offers: Conceptual framework and research propositions. *Journal of Marketing* 69 (1): 32–45.
- Simonson, I. 2008. Will I like a “medium” pillow? Another look at constructed and inherent preferences. *Journal of Consumer Psychology* 18 (3): 155–69.
- Sinha, R. K., and N. Mandel. 2008. Preventing digital music piracy: The carrot or the stick? *Journal of Marketing* 72 (1): 1–15.
- Smith, E. R. 1996. What do connectionism and social psychology offer each other? *Journal of Personality & Social Psychology* 70 (5): 893–912.
- Solomon, R. L. 1949. An extension of control group design. *Psychological Bulletin* 46 (2): 137–50.
- Song-Oh, Y., and I. Simonson. 2008. Choice set configuration as a determinant of preference attribution and strength. *Journal of Consumer Research* 35 (2): 324–36.
- Squire, B., J. Readman, S. Brown, and J. Bessant. 2006. The impact of mass customisation on manufacturing trade-offs. *Production & Operations Management* 15 (1): 10–21.
- Syam, N., P. Krishnamurthy, and J. D. Hess. 2008. That's what I thought I wanted? Miswanting and regret for a standard good in a mass-customized world. *Marketing Science* 27 (3): 379–97.
- Syam, N. B., and N. Kumar. 2006. On customized goods, standard goods, and competition. *Marketing Science* 25 (5): 525–37.
- Thorndike, E. L. 1911. *Animal intelligence*. New York: Macmillan.
- Tian, K. T., W. O. Bearden, and G. L. Hunter. 2001. Consumers' need for uniqueness: Scale development and validation. *Journal of Consumer Research* 28 (1): 50–66.
- Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (4481): 453–58.
- Valenzuela, A., R. Dhar, and F. Zettelmeyer. 2009. Contingent response to self-customization procedures: Implications for decision satisfaction and choice. *Journal of Marketing Research* 46 (6): 754–63.
- Voelckner, F. 2006. An empirical comparison of methods for measuring consumers' willingness to pay. *Marketing Letters* 17 (April): 137–49.
- von Hippel, E. 2001. Perspective: User toolkits for innovation. *Journal of Product Innovation Management* 18 (4): 247–57.
- von Hippel, E. 2005. *Democratizing innovation*. Cambridge, MA: MIT Press.
- von Hippel, E., and R. Katz. 2002. Shifting innovation to users via toolkits. *Management Science* 48 (7): 821–33.
- Wind, J., and A. Rangaswamy. 2001. Customerization: The next revolution in mass customization. *Journal of Interactive Marketing* 15 (1): 13–32.
- Winne, P. H. 1997. Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology* 89 (3): 397–410.
- Xu, A. J., and R. S. Wyer Jr. 2007. The effect of mind-sets on consumer decision strategies. *Journal of Consumer Research* 34 (4): 556–66.
- Zaichkowsky, J. L. 1985. Measuring the involvement construct. *Journal of Consumer Research* 12 (3): 341–52.

Appendix: Measurement Scales

Preference insight (Study 1): Which color should the sole of your ideal running shoe be? Which color should

the laces of your ideal running shoe be? Which color should the lining of your ideal running shoe be? Which color should the stitches of your ideal running shoe be? Do you prefer visible air cushions on your running shoes? Do you prefer a running shoe with just a few large color areas or with many small color areas? Which base color do you prefer on your running shoes? Measured as “specific answer” (1) or “I don't know” (0).

Involvement (Study 1): Important–unimportant; meaningful–meaningless; useful–useless; interesting–boring; exciting–unexciting; relevant–irrelevant; valuable–valueless. Measured as a semantic differential with seven-point scales.

Innovativeness (Study 1): I prefer tasks which require original thinking. I always search for new ways to look at things. People often ask me for help with creative tasks. I often surprise people with novel ideas. I often try to find new applications for everyday things. I do not have an especially vivid imagination (reversed). Measured on seven-point scales (1 = strongly disagree; 7 = strongly agree).

Design affinity (Study 1): I like designing things. I would call myself a designer. I have a grasp of artistic creation and design. I like painting and drawing. (1 = strongly disagree; 7 = strongly agree).

Experience with running shoes (Study 1): I use running shoes very often. I generally spend a lot of time searching for new running shoes. Compared to the average person, I know a lot about running shoes. (1 = strongly disagree; 7 = strongly agree).

Experience with toolkits (Study 1): I often work with toolkits. I often work with the Nike toolkit. I am familiar with the functionality of toolkits. I am familiar with the functionality of the Nike toolkit. (1 = strongly disagree; 7 = strongly agree).

Preference insight (Study 2): Which color should the bezel (the metal ring around the dial) of your ideal watch be? On the dial of your ideal watch, should the numbers be written as numerals or represented by lines? Which color should the casing of your ideal watch be? Measured as “specific answer” (1) or “I don't know” (0).

Design affinity (Study 2): I like designing things. I would call myself a designer. I have a grasp of artistic creation and design. I like painting and drawing. I am good at working with graphics programs (like CorelDraw or Photoshop). (1 = strongly disagree; 7 = strongly agree); averaged index.

Design experience (Study 2): I have already designed a product in the past. (1 = strongly disagree; 7 = strongly agree.)

Involvement (Study 2): Important–unimportant; valuable–valueless; useful–useless; meaningful–meaningless; boring–fascinating; unessential–essential; necessary–unnecessary. Measured as a semantic differential with seven-point scales; averaged index.

Innovativeness (Study 2): Before buying a new brand, I like to consult a friend who has experience with that brand. I seldom ask friends about their experiences with a new product before I buy it myself. I often search for new information about new products and brands. I often search for new products and services. I like being in situations where I encounter new information about products. (1 = strongly disagree; 7 = strongly agree.)

Need for uniqueness (Study 2): When I buy a product, it is important for me to find something which communicates my uniqueness. I have bought unusual products or brands to create an unusual personal image. I often try to get a more interesting version of a standard product because I want it to be inventive. I often dress unconventionally, even when it is probable that others will be bothered by it. Concerning the products I buy or the situations in which I use them, I have often broken accepted customs. I enjoy challenging the taste of people I know by buying something which they will not accept at first. When products or brands I like become very popular, I lose interest in them. I often try to avoid products or brands when I know that the average population buys them. When a product I own becomes popular among the general population, I use it less often. (1 = strongly disagree; 7 = strongly agree.)

Recently bought a watch/received a watch as a present (Study 2): I recently bought a watch or received a watch as a present. (1 = strongly disagree; 7 = strongly agree.)

Knowledge and beliefs about toolkit aspects (Study 2): Toolkits which allow the user to design individual watches: . . . are generally very easy to handle. . . . are generally very customer-friendly. . . . generally require high technical knowledge. . . . are generally fun. I am confident that I would understand a watch toolkit (which allows the user to design individual watches and with which I am not familiar) very quickly. (1 = strongly disagree; 7 = strongly agree.)

Discretionary income (Study 2): How high is your discretionary income per month (the amount which remains available to you after you have covered your fixed costs like rent and insurance)? (“<€100,” “€100–199,” “€200–299,” “€300–399,” “€400–499,” “>= €500”).

Preference insight (Study 3): How should the bezel (the metal ring around the dial) of your ideal watch be decorated? [I don’t know/I’m not sure; none, different metal than the body; gems; orientation marks (rotary); time marks (rotary); other] Which shape should the body of your ideal watch have? [I don’t know/I’m not sure; rectangular; round; oval] Which design should the seconds hand of your ideal watch have? [I don’t know/I’m not sure; longer and different color than the other hands; longer and the same color as the other hands; shorter and different color than the other hands; shorter and the same color as the other hands; other] Which design should the digits on your ideal watch have? [I don’t know/I’m not sure; normal digits; lines only] Which color should the digits on your ideal watch be? [I don’t know/I’m not sure; black; white; red; blue; other] Which color should the face of your ideal watch be? [I don’t know/I’m not sure; black; white; red; blue; other] Which color should the body of your ideal watch be? [I don’t know/I’m not sure; silver/metal; black; white; other] Which color should the wristband of your ideal watch be? [I don’t know/I’m not sure; black; white; brown; silver/metal; other] Measured as “specific answer” (1) or “I don’t know” (0).

Perceived user friendliness of the toolkit (Study 3): How difficult did you perceive the design process to be with the 121Time toolkit? (1 = very easy, 7 = very complex.)

General learning skills (Study 3): I am a person who learns how to handle new products very quickly. It is easy for me to get into new topics. I was a good student. It is no problem for me to acquire new knowledge. I find manuals for new technical equipment (mobile phones, computers, etc.) easy to understand. (1 = strongly disagree, 7 = strongly agree.)