

**Value Creation by Toolkits for User Innovation and Design:
The Case of the Watch Market**

Nikolaus Franke and Frank Piller

Working Paper

A later version of this paper is published in

Journal of Product Innovation Management, 2004, 21 (6): 401-415.

Author contacts:

Prof. Dr. Nikolaus Franke (*corresponding author*)
Vienna University of Business Administration and Economics
Department of Entrepreneurship
Augasse 2-6, 1090 Vienna, Austria
+43 / 1 / 31336-4582
E-mail: nikolaus.franke@wu-wien.ac.at

Dr. Frank T. Piller
Technische Universität München
Department of General and Industrial Management
Leopoldstr. 139, 80804 Munich, Germany
+49 / 89 / 289-24820
E-mail: piller@ws.tum.de

Value Creation by Toolkits for User Innovation and Design:

The Case of the Watch Market

Biographical notes:

Prof. Dr. Nikolaus Franke is a Professor of Entrepreneurship at the Vienna University of Economics and Business Administration and Director of the Research Center for Entrepreneurship and Innovation Management at the same institution. His research interests include innovation management, toolkits for user innovation, communities and entrepreneurship.

Dr. Frank Piller is a senior lecturer in the MBA program at the Technische Universitaet Muenchen (TUM), Germany, and Director of the TUM Research Group Mass Customization. His research interests include customer integration, mass customization and personalization, innovation and technology management as well as strategic management.

Value Creation by Toolkits for User Innovation and Design:

The Case of the Watch Market

Abstract

This study analyzes the value created by so-called "toolkits for user innovation and design", a new method of integrating customers into new product development and design. Toolkits allow customers to create their own product, which in turn is produced by the manufacturer. In our study, we asked (1) if customers actually make use of the solution space offered by toolkits, and if so, (2) how much value the self-design actually creates.

In our study, we used a relatively simple, design-focused toolkit for a set of four experiments with a total of 717 participants, 267 of whom actually created their own watches. The heterogeneity of the resulting design solutions was calculated using the entropy concept, and willingness to pay was measured by the contingent valuation method and Vickrey auctions.

Entropy coefficients showed that self-designed watches vary quite widely. On the other hand, significant patterns are still visible despite this high level of entropy, meaning that customer preferences are highly heterogeneous and diverse in style but not completely random.

We also found that consumers are willing to pay a considerable price premium. Their willingness to pay (WTP) for a self-designed watch exceeds the WTP for standard watches by far, even for the best-selling standard watches of the same technical quality. On average, we found a 100% value increment for watches designed by users with the help of the toolkit.

Taken together, these findings suggest that the toolkit's ability to allow customers to customize products to suit their individual preferences creates value for them in a B2C setting even when only a simple toolkit is employed. Alternative explanations, implications and necessary future research are discussed.

Introduction

The advent of the internet has facilitated new forms of producer-customer interaction in product development (Sharma and Sheth 2004). One promising new form of interaction is outlined in the concept of "toolkits for user innovation and design" (von Hippel 2001, Thomke and von Hippel 2002), or "user design" (Dahan and Hauser 2002). Both ideas are based on the proven ability of customers to design their own products (von Hippel 1988).

A toolkit is a design interface which enables trial-and error experimentation and gives simulated feedback on the outcome. In this way, users are enabled to learn their preferences iteratively until the optimum product design is achieved (von Hippel and Katz 2002). The manufacturer, in turn, produces the product to the customer's specifications. Toolkits exist in various fields, ranging from computer chips (ASICs) to individualized athletic shoes. Depending on the type of toolkit, the outcome is an individualized product (Park, Jun and MacInnis 2000) or even an innovation (Thomke and von Hippel 2002). The rationale underlying the toolkit, however, is the same: It allows the customer to take an active part in product development.

Any new concept must be analyzed thoroughly for its actual value. While several authors advocate the merits of the toolkit concept, others have recently discussed its limitations concerns (e.g., Agrawal, Kumaresh and Mercer 2001; Zipkin 2001). They argue that the broad solution space offered by toolkits is of limited value because for most users the cost of actively designing might exceed the benefits of getting an individualized product. Particularly in B2C applications, the value creation potential of

toolkits is questioned, and this notion has received some endorsement by the recent shut-down of Mattel's "MyDesign Barbie" and Levi's "Original Spin" site.¹

On the other hand, a considerable number of successful applications have been reported anecdotally. A recent literature review, however, shows that only very few academic studies have dealt with this new phenomenon. Among them, merely anecdotal case studies are the rule, and none have attempted a quantitative analysis of the value such toolkits deliver to customers (Franke and Piller 2003).

The purpose of this study is to advance our understanding of the new concept of toolkits for user innovation and design by analyzing their actual value creation from the customer's perspective. In a quantitative study of a design-focused watch toolkit in a B2C setting, we asked (1) if customers actually make use of the solution space offered by toolkits, and if so, (2) how much value the process of self-design actually creates.

Our article is organized as follows: Section 2 presents a literature review of the concept of toolkits for user innovation and design, and in Section 3 we proceed to describe the methods applied. Our findings are then summarized in Section 4. The final section discusses the implications of our study as well as alternative explanations and future research.

Literature Review

Toolkits for User Innovation and Design

Von Hippel (2001) defines toolkits for user innovation as a technology that (1) allows users to design a novel product by trial-and-error experimentation and (2) delivers immediate (simulated) feedback on the potential outcome of their design ideas.

This idea of outsourcing design-related tasks in product development to customers stands in sharp contrast to the traditional practice of new product market research. The traditional method of obtaining customer input is to gather data meticulously from representative customers in a chosen market sector and then to use this (need-related) information in order to create ideas for new products (e.g., Rangaswamy and Lilien 1997, Lonsdale, Noel and Stasch 1996). In order to reduce the risk of failure, need-related information from customers is integrated iteratively at many points in the new product development process (e.g., Cusumano and Selby 1995, Dahan and Hauser 2002, Holmes 1999). After many time-consuming iterations between the customer and manufacturer, a new or adapted product is found, usually at high cost. However, although enormous market research expenditure is the norm, the flop rates for new products are still relatively high (Cooper 1999, Crawford 1979, Griffin, Belliveau and Somermeyer 2002).

Toolkits offer potential advantages compared to the traditional method of new product development in that they enable an individual user to state or specify his/her preferences precisely. Moreover, the interaction between user and toolkit can be easier than the alternative of costly interaction between the user and manufacturer in the process of market research. Most notably, the information obtained with the help of a toolkit is located at the individual level: The manufacturer can then produce and deliver a designated product to suit the individual user. The resulting – potentially far closer – fit between user preferences and the product itself should yield a higher level of satisfaction with the new product and subsequently increase the customer's willingness to pay.

Obviously, there are variations in the types of available toolkits. Some very complex toolkits offer a large solution space and can not be employed without a precise technical understanding (e.g., toolkits for designing application-specific integrated circuits, as described by von Hippel and Katz 2002). They depend upon the customer taking on a very active role as designer and allow substantial innovations. Most of them are employed in B2B settings where the economic benefits of toolkits are apparent in many situations. Other toolkits, particularly in consumer markets, only offer a small solution space and only allow users to combine relatively few options (e.g., toolkits for designing eyeglasses, as described by von Hippel 2001). Although the underlying principle is the same, the latter toolkits focus on individuality and customization rather than on innovation. We therefore suggest using the enhanced term "toolkits for user innovation *and design*" as it describes this new concept's entire range of applications.

Why and When Toolkits Make Sense

Two lines of argumentation have been brought forward to explain the potential benefits of toolkits for innovation and design: (1) the heterogeneity of customer preferences and (2) the problems associated with shifting preference information from the customer to the manufacturer.

It is common knowledge that customer preferences are heterogeneous and change quickly in many markets. The need for economies of scale, however, has forced manufacturers either to satisfy the general needs and preferences of a customer segment with a standard product (thus leaving many customers or potential customers dissatisfied, if only to a certain degree) or to offer a custom-made product at a very high price. Recently, new production technologies have dramatically reduced the fixed costs of tooling in manufacturing. These "mass customization" methods have enabled custom

goods to be produced with near mass production efficiency (Pine 1993, Tseng and Jiao 2001, Tseng and Piller 2003, Wind and Rangaswamy 2001).

To date, only few studies have attempted to quantify the heterogeneity of user preferences. In an empirical study on Apache's security software, Franke and von Hippel (2003a) show that users do in fact have very unique needs, leaving many displeased with standard products. Users even claimed that they were willing to pay a considerable premium for improvements which satisfy their individual needs. In a meta-analysis of published cluster analyses, Franke and Reisinger (2003) find evidence that this dissatisfaction is not an exception. Current practice in market segmentation generally leads to high levels of total variance left over as in-segment variation (approximately 50% on average). This means that a major group of customers remains somewhat dissatisfied with standard offerings, even in seemingly mature markets. Another indicator for the heterogeneity of user needs is the fact that many users take the time to innovate or modify existing products. Franke and von Hippel (2003b) present an overview of several studies and show that in the fields sampled to date, 10 to nearly 40% of users report having modified or developed a product for in-house use (in the case of industrial products) or for personal use (in the case of consumer products).

This would lead to the expectation that at least in some markets customers would value the opportunity to tailor a product to their specific needs and thus make use of the solution space offered by a toolkit.

The second line of argumentation focuses on the problem of shifting preference information from the customer to the manufacturer. Such information is known to be difficult to encode, transfer and decode (Cooper 1979, Dougherty 1990, Leonard-Barton 1995, Poolton and Barclay 1998). If a manufacturer conducted a market study on the

need for new products, the most frequent answer he would receive would probably be, "I want the same product, only better and cheaper." One theoretical explanation for this phenomenon is described by the concept of "information stickiness."

The stickiness of a given unit of information is defined as the incremental expenditure required to transfer that unit from one place to another in a form that can be accessed by a given information seeker. When this expenditure is low, information stickiness is low; when it is high, stickiness is high. The definition of "sticky" information is broader and also incorporates "tacit" knowledge (Polanyi 1958) as one of several possible causes of stickiness. Thus, information stickiness may be rooted in the characteristics inherent to the information itself (e.g., tacitness), and/or it may be due to the individual characteristics of an information seeker or provider and his style of interaction (von Hippel 1994).

Studies have shown that the stickiness of information can be very high in innovation-related matters (Ogawa 1998, von Hippel 1998). Many users are not truly aware of their needs when it comes to new products, and even if they are, they are often not able to formulate them explicitly. A toolkit can be a means of "unsticking" such information. It is often found that novel products are developed through "learning by doing" processes (von Hippel and Tyre 1995, Thomke, von Hippel and Franke 1998) or by "trial and error" (Ishii and Takaya 1992, Polley and Van de Ven 1996). In order to find a solution, the innovator needs to be informed about all of the possibilities at his disposal; he must try out various possibilities, learn from errors, compare different solutions, and thus engage in a time-consuming, step-by-step learning process. Toolkits provide just such a setting for "trial-and-error" learning. Obviously, they make sense when information stickiness is high and can be "unstuck" by trial-and-error learning.

Recently, some authors have emphasized the downside of toolkits for user innovation and design. Pine argues that the active learning role of the user-designer may lead to "mass confusion" instead of "mass customization" (Teresko 1994). Users might be overwhelmed by the number of possibilities at their disposal (Huffman and Kahn 1998, Kamali and Loker 2002, Stump, Athaide and Joshi 2002, Wind and Rangaswamy 2001, Zipkin 2001). Anyone who has been forced to choose from a very wide selection – for example, in a restaurant which offers 500 entrées – knows that equating a large number of possibilities with high customer satisfaction would be blind optimism. The human capacity to process information is limited (Miller 1956). The burden of having to choose from too many options may simply lead to information overload (Maes 1994, Neumann 1955). Consequently, users may turn away from the liberty to choose and decide for the standard (or starting) solution offered by a toolkit (Dellaert and Stremersch 2003, Hill 2003) – or they may even frown and turn their backs completely.

Empirical Studies on Toolkits

The number of firms operating with toolkits is growing steadily in industrial as well as consumer markets. Many examples can be found easily on the internet. A recent literature review revealed that empirical studies on toolkits, however, are scarce (Franke and Piller 2003). In short, the evolving literature on mass customization concentrates on technical and production aspects instead of the interface between user and producer, that is, the toolkit itself. The literature which directly addresses toolkits mostly supplies only anecdotal studies and describes toolkit cases in a narrative style. Furthermore, publications focus on firms implementing and using toolkits, not on users interacting

with them. We will review the most recent exceptions here (for an exhaustive overview, see Franke and Piller 2003).

Jeppesen (2002) analyzes 78 computer games and finds that toolkits, although well accepted by users, may increase the need for manufacturer support. This drawback is alleviated in many cases by user-to-user support systems. Franke and von Hippel (2003a) analyze the users of Apache security software, which is "open source" (i.e., it can be modified by skilled users). They find that users who introduce their own software modifications are significantly more satisfied than non-innovating users, thus they come to the conclusion that toolkits create value for users. Estimates reveal that the average (i.e., less skilled) Apache user is willing to pay a considerable amount (over \$5,000 per user) in order to ensure that their individual security needs in Apache are met to their full satisfaction. This indicates that the individual adjustment of a product to a user's needs potentially constitutes an enormous value increment. It has to be noted, however, that the respondents made those indications on the basis of hypothetical products (and not based on a concrete product or concept).

In an experimental study with 72 research subjects, Kamali and Loker (2002) examine the involvement of consumers who designed a T-shirt using a toolkit. The results point to an overall interest in designing as well as higher satisfaction with the toolkit as involvement increased. The higher level of interactivity also increases the customers' willingness to purchase. When asked about their willingness to pay more for truly customized T-shirts, participants responded affirmatively. However, the authors point out that future research should investigate this matter more thoroughly.

Dellaert and Stremersch (2003) study consumer interaction with a design toolkit for PCs. They find a trade-off between product utility (i.e., the utility of a customized

product better fitting a user's needs) and process complexity as perceived by the user. If perceived process complexity is high, perceived product utility decreases. The study also points to the fact that toolkits appeal more to expert consumers.

Park et al. (2000) and Levin et al. (2002) compare the effect of using either a subtractive or an additive option-framing method on the user perception of a customizable product. Both studies show that subtracting yields increased willingness to pay.

In summary, we can say that we are beginning to gain an understanding of the value and potential drawbacks of toolkits for user innovation and design. One piece that is still missing, however, is a quantification of such toolkit-generated value compared to offering standard, non-customized products. We do not know whether customers are willing and able to make use of the possibilities toolkits offer, and our research aims to contribute to answering these open questions.

Method

Research Object: a Watch Toolkit

We decided to focus on a single, relatively simple toolkit in a B2C setting that only allows design (and not innovation) activities. Compared to the possible research design of investigating the interactions with multiple toolkits, our approach has the advantage of providing deeper insight (i.e., high internal validity). A multiple-case approach always involves the risk of an apples-and-oranges solution. We have to admit, however, that the external validity of our approach is limited, that is, we always have to question the extent to which our results can be generalized. Certainly, they will not apply to sophisticated B2B toolkits which allow real innovation.

We agreed upon a product area that can be characterized by high heterogeneity of preferences, imagining that relatively simple "Swatch"-type watches in the 25 to 100-euro price range would be a promising selection. We based our investigations on 16 expert interviews with industry specialists (retailers, manufacturers, trade specialists) and found that at least 2,000 different models of these watches are offered on the (local) market where the study took place (Austria). Most producers frequently change their product range at least twice a year. In a five-year period, this would lead to more than 20,000 different "standard" designs. We interpreted this as a clear indication of high heterogeneity and dynamic preferences in this market.

The toolkit we chose is operated under the brand name "Idtown" by Global Customization Ltd., Hong Kong, a spin-off company founded by the Advanced Manufacturing Institute of the Hong Kong University of Science and Technology (HKUST). Idtown was one of the most established websites in the field and is referred to as being famous for business-to-consumer mass customization in many publications (e.g., Cairncross 2000, Khalid and Helander 2003, Piller 2003, Tseng and Jiao 2001). The website www.idtown.com operated continuously between October 1997 and March 2003. It was shut down due to management problems and for a major renovation in April 2003. The site is scheduled to re-open in 2004 and will then serve as a testing area for our future studies. A similar toolkit in operation is Factory1to1.com, which is operated by a Swiss watch maker.

The toolkit of Idtown is relatively simple. The problem-solving activities in which users engage consist only of the visual aspects of watch design (functional aspects of "what a watch does" are known to users and consistent across the design space). Users designing a watch can engage in learning by doing because they can

immediately look at a simulation that incorporates each design decision made. The toolkit thus allows for trial-and-error learning with an immediate feedback function (design). It contains a module library and opens up an immense solution space of at least 650 million different possible product designs. In our assessment, the toolkit is relatively easy to use. It offers a wide variety of design possibilities: Selecting and combining pre-defined options for the strap (80 alternatives), case (60 alternatives), face (150 alternatives), the hour/minute hands (30 alternatives) and the second hand (30 alternatives) of a watch. Truly innovative solutions, however, are not possible, and the role of the user merely consists in "designing" instead of "innovating." A screen shot of the website is shown in Figure 1.

<< insert figure 1 about here >>

Users start on the home page by choosing one of the basic product categories. They can follow a top-down approach and go through the different levels of the components, or they can choose the sequence of selecting options freely. It is always possible to go back one or more steps or to begin the design process over by returning to the home page. Moreover, users do not have to make a decision for every component but can choose a pre-configured option. During the entire process of configuration, the toolkit depicts the current selection with a full picture of the self-designed watch. Placing a customized watch in the shopping cart and proceeding to check out is similar to other online shopping websites.

Design of Experiment A

We sampled 165 users (see below for a more detailed discussion of the sample) and presented them with a token for a self-designed watch fabricated by Idtown. The design process was carried out independently on four remote PCs provided by the authors.

Empirical data were collected at different stages for each subject. First, we stored the individual design solution each participant came up with. Second, once they had finished the design process, the participants were asked to fill out a questionnaire which asked about their willingness to pay for the self-designed watch compared to selected standard (i.e., not user-designed) models.

Participants were recruited from among (graduate-level) management students who were on campus at the time of our studies. In our data collection design, a completely random sample was not entirely possible. Due to the exploratory nature of our research questions and the lack of sufficient data, however, this is not mandatory at this stage. Therefore, our data is biased in favor of young and adept persons who are familiar with the internet. According to the management of Idtown, it is worth noting that this group represented their major target group for actual sales. In order to obtain a sufficiently large sample of $n = 165$ participants, we asked 300 students "Would you be interested in taking part in a short research experiment? You will get a watch in return." The acceptance rate thus came to 55%. Most refusals were due to a lack of time because courses were beginning. We believe that the high acceptance rate is due to both the high incentive and probably also the fun such experiments entail.

Design of Supplementary Experiments B, C, and D

We conducted three more experiments in order to intensify our understanding of the findings from Experiment A (Table 1). In Experiment B, we asked another sample of

students (who had not designed a watch themselves) about their WTP for the user-designed watches (from Experiment A) and for their WTP for comparable standard types. The watches were displayed on two large posters showing the products in actual size. We did this in order to validate the WTP for the standards and to analyze whether the value increment of a self-designed watch is designer-specific or general. Both experiments were repeated with a different method of measuring WTP, that is, Vickrey auctions (see below). In all supplementary auctions, our sampling approach and acceptance rate was similar to Experiment A.

<< insert table 1 about here >>

Measuring Willingness to Pay

The core variable in our study is willingness to pay (WTP). Estimating a user's WTP is known to be a difficult task. Prior research offers several concepts for measuring WTP, ranging from actual transaction data to simulated auctions and survey data (Wertenbroich and Skiera 2001). In our study, we decided to use two different methods in order to cross-validate our results: The contingent valuation method and the Vickrey auction.

In the contingent valuation method (CVM), respondents are asked directly how much they are willing to pay for a product or service (Mitchell and Carson 1989). This approach is relatively easy to use but is said to overestimate actual WTP. Studies that compare CVM WTP with actual cash payments have shown actual spending behavior to be only 15-20% of expressed WTP (Franke and von Hippel 2003a).

Therefore, in order to validate our findings, we conducted two more experiments (C and D) by means of a Vickrey auction (VA). A VA is an auction in which

participants' bids are sealed and no bidder knows about the others. The item is sold to the highest bidder at a price equal to the second-highest bid, thus the winner pays less than the highest bid (Vickrey 1961).

It can be shown both empirically and using game theory that the dominant strategy of a bidder is to bid the same as the actual maximum WTP (e.g., Cox, Robertson and Smith 1982, Hoffman et al. 1993). In Experiments C and D, we conducted real auctions, that is, participants submitted real, binding bids for actual watches, and the respective winner actually purchased a watch. In contrast to CVM, our VA results are biased downward because respondents were not selected based on an actual desire to purchase a watch. Taken together, both measures (CVM and VA) might enhance our understanding of user WTP for individualized products in the watch market, although they are both biased.

Findings

Heterogeneity of Resulting User Designs

The purpose of toolkits is to address heterogeneous user preferences. If preferences are actually heterogeneous and the toolkit offers different design solutions, we should also expect heterogeneous user designs as the outcome of these processes. In this case, we would also expect a certain increase in WTP.

We display the heterogeneity of user designs in two ways. First, we show how many standard products would be necessary to meet user preferences (expressed in their individual designs). Second, we calculate entropy coefficients in order to express heterogeneity using a familiar ratio.

Necessary Standard Watches

In this part of our analysis, we investigate how many standard products would be necessary in order to meet customer preferences as well as the toolkit does. As the simple watch toolkit in our study only allows for configuration, an omniscient manufacturer could also (in theory) offer the corresponding set of standard watches. Table 2 shows the number of standard watches the manufacturer would have to offer in order to meet the preferences of our sample of 165 customers. The assumptions of our simulation are subsequently discussed.

<< *insert table 2 about here* >>

For the manufacturer, there are two decisions involved: (1) the intended proportion of customers he aims to reach with his standard watches and (2) the degree of satisfaction he believes is necessary.

Of course it is desirable for a manufacturer to offer standard products that are appealing to as many potential buyers as possible. Unfortunately, this also makes it necessary to offer many standard products. In our sample, the manufacturer would need 159 different standard watches to meet the preferences of our entire sample of 165 users. If the manufacturer settled for meeting the preferences of 80% of users (or, in our sample, 132 participants), the number of standard products would drop to 126.

The manufacturer could also give up the ambition of meeting every customer's preferences. For example, the manufacturer might think it would be enough to meet customers' preferences at the level of 80% (i.e., the customer would get what he wants in exactly 4 of the 5 design dimensions, while his preferences in the fifth dimension

would not be met). In this case, the number of standards necessary would be 134 for 100% market coverage and 101 if the manufacturer settled for attaining this 80% satisfaction level for only 80% of the customers.

The results show that a relatively large number of standard designs are necessary for our small sample. Only in cases where manufacturers accept a relatively small fraction of customers as a target group and believe that low levels of individual customer satisfaction are sufficient will the resulting number of necessary standard designs drop to manageable levels. One must not forget that in reality the market does not consist of 165 customers but of several million potential watch buyers. Although it is not possible to simply extrapolate the numbers, it becomes obvious that indeed a *huge* number of standard watches would be necessary to meet the preferences of a sufficiently large target group at a satisfactory level.

Finally, the assumptions underlying our analysis are not to be overlooked: (1) We assume that a manufacturer has perfect knowledge of customer preferences and thus of the "optimum" standard designs. Of course, this is a huge overestimate of the capabilities of marketing research. (2) We assume that any deviation in such a standard watch from the "ideal" watch (i.e., the one the individual created) is equally negative. This is a simplification, as some differences may not matter very much to a user, while others might be perceived as a huge setback. (3) We assume that the toolkit-designed watch satisfies the user/designer completely, that is, the self-designed watch is treated as the "ideal" product (which again is a simplification, as it is not likely that the simple configuration toolkit will allow the manufacturer to meet every customer's needs in their entirety).

Entropy of user designs

Originally derived from the physical sciences (thermodynamics), the concept of entropy was introduced in information theory by Claude Shannon (Shannon and Weaver 1949).

Entropy is a measure of the degree of disorder, uncertainty, or randomness of a probabilistic system. In management sciences, it has been used to measure diversification (Vachani 1991), individual decision-making strategies (Gensch and Soofi 1995), and brand purchasing behavior (Herniter 1973), among other things.

The entropy coefficient E of a system consisting of n possible states, with p_i being the probability that the system will be in state i , can be calculated as follows:

$$(1) \quad E = -\sum_i^n p_i \log_2(p_i)$$

The base of the algorithm is arbitrary, thus the relative rather than the absolute entropy value is important. For example, if one threw a die six times and got the numbers (1, 2, 3, 4, 5, 6), the entropy of the system – the die – would be 2.58. Essentially, no concentration is visible, thus the entropy coefficient is at its maximum. If we used a bogus die and got the numbers (1, 3, 1, 5, 1, 1), the entropy would be 1.25. The relative entropy coefficient ($E_{\text{emp}}/E_{\text{max}}$) for the second die would thus be 48.4%, indicating some pattern in the system.

In our sample of user-designed watches, we observed relatively high entropy (Table 3). The relative entropy coefficients are fairly close to the maximum.

<< insert table 3 about here >>

The univariate model shows that only the "strap" dimension shows a somewhat lower entropy (73.7% of maximum entropy). Here, only 30 of the 80 design alternatives were chosen by the 165 users, with one single strap accounting for 64 designs chosen (38.8% of the sample). In the other dimensions, the entropy coefficient is higher. Naturally, the concentration declines as more dimensions are analyzed, meaning that entropy increases.

In the case of two dimensions, some combinations appear to be "natural" complements. The highest frequency could be observed with a specific strap / second hand combination, which alone accounted for 34 user designs. When all five dimensions are analyzed simultaneously (i.e., the "complete" watch design is analyzed), entropy is 99.6% of its maximum possible value. The system is very close to maximum disorder, meaning that preferences at this level are quite heterogeneous.

On the other hand, the observable concentration pattern is clearly significant in all cases analyzed.² Therefore, we can conclude with some certainty that although the heterogeneity of preferences is very high in our sample, there is still some tendency to cluster beyond pure chance. Preferences are not completely heterogeneous but follow some weak patterns.

Value of User Design

So far, we have only *assumed* that deviations from the ideal design are relevant for users. Theoretically, it could be the case that the high observed heterogeneity of design solutions actually is random because users simply do *not care* about the design of watches. In this section, we therefore analyze the value increment for self-designed watches by checking whether people really care about their unique designs. Would they

actually pay more to have their preferences met? As outlined in Section 3.4, we measure the value of the user design by two means: The contingent valuation method (CVM) and Vickrey auctions.

Contingent Valuation Method

We chose two standard models of the same price and of equal quality as benchmarks for self-designed watches. We had identified them as much-sold standard watches in a "quick-and-dirty" interrogation of two retailers. They are referred to as Standard Watches 1 and 2 below (for a visual impression, see Figure 2). Another benchmark we used in Experiment A was an imagined "ideal" watch in the same segment which was user-designed by an imaginary "perfect" toolkit without any restrictions in the solution space (but still of equal and constant technical quality).

<< *insert figure 2 about here* >>

The results were remarkable (see Table 4). We found that each user/designer's WTP for his self-designed watch was 48.5 euros, more than twice the WTP for the two standard types 1 and 2 with the same technical quality (21.5/21.5 euros). Differences are highly significant at $p < .000$ (paired t-tests). 87% and 85% (respectively) of all participants were willing to pay more for the self-designed watch than for Standard Watches 1 and 2. Obviously, the toolkit facilitated a high value increment for most of the respective user/designers

Having inquired about the WTP for the imagined "ideal" watch designed with the imaginary "ideal" toolkit, the WTP jumped to 92.0 euros – again an almost 100% increase (although the WTP for *imaginary* products should certainly not be taken

literally). The differences concerning the WTP for the self-designed watch are highly significant at $p < .000$ (paired t-test).

In Experiment B, we asked another sample of subjects (who had not designed a watch themselves) about (1) their WTP for the user-designed watches (from Experiment A), (2) their WTP for the two standard watches (1 and 2) used in Experiment A, and (3) their WTP for the four best-selling watches of the same quality. In order to identify these bestsellers, we had thoroughly interviewed 16 retailers, manufacturers, and industry experts. The bestsellers in the particular product and quality category all came from the Swiss brand "Swatch" (referred to as Standard Watches 3 to 6 below). In the experiments, the "Swatch" label was concealed in order to isolate the *design* aspect.

The results clearly show the reliability of the WTP for Standard Watches 1 and 2 (Experiments A and B). The differences in both experiments were small (21.5 euros vs. 22.4; 21.5 vs. 23.1) and not significant. The differences between the WTP for the self-designed watch (Experiment A) and the best-selling Standard Watches 3 to 6 (Experiment B), however, are substantial (48.5 euros vs. 18.5/20.0/25.8/32.7) and highly significant for all four pairs ($p < .000$). This confirms the value increment created by self-design, even when compared to the best-selling watches on the market.

We also found that the mean WTP for the self-designed watches of *other* users (i.e., the WTP of non-designers in Experiment B for user-designed watches from Experiment A) is notably different from that of the user/designer himself. The mean WTP for one and the same watch decreases from 48.5 to 23.1 euros when non-designers are asked about their WTP. The difference is highly significant at $p < .000$.

This denotes that the user-designed watches are not primarily designed *better* than standard watches but appear to be *better adapted* to the personal preferences of the

user/designer. It may also point to another value-creating effect of self-design (see discussion below). It is interesting, however, that the designs which amateurs make relying on a simple toolkit in a sketchy 13-minute design process bring about an even higher mean WTP in two cases (23.1 vs. 18.5 euros for the best-selling Standard Watch 3; and 23.1 vs. 20.0 euros for the best-selling Standard Watch 4; differences not significant). In two other cases, WTP comes relatively close to the bestsellers (23.1 vs. 25.8 for Standard Watch 5; 32.7 euros for Standard Watch 6; the latter difference being significant at $p < .000$). Thus, when treated as potential standard watches, the amateurs' designs were on average attributed approximately the same value by the market as the best-selling standard models created by professional designers.

Vickrey Auctions

Experiments C and D were carried out in order to validate the findings regarding the relative differences in WTP for self-designed and standard watches with a more sophisticated method (Vickrey auctions). As discussed in our methods section, CVM is likely to overestimate actual WTP, while the Vickrey auctions (VA) are biased downward.

Our results show that WTP measured by VA is, in fact, different from the CVM results. The "overestimation ratio," however, is reasonably constant for most watches, ranging from 23.2 to 30.7%, with the sole exception being Standard Watch 6. This discovery generally confirms our analyses based on the CVM presented above, although it must be noted that the WTP for Standard Watch 6 comes very close to that of the self-designed watch (the difference not being significant). Thus, our analysis shows that on the watch market even a simple toolkit for user innovation and design yields a very high value increment compared to most best-selling standard products.

<< insert table 4 about here >>

Discussion

In our project, we found that participating users attribute a high value increment to their own design activities, even in a B2C setting where the economic benefit of a customized solution is at first glance not as apparent as in B2B markets. The WTP (willingness to pay) for a self-designed watch is almost twice as high as for the best-selling standard model available on the market. Also, the product designs are indeed *very* heterogeneous. Thus, it would appear that offering individualized products by means of toolkits for user innovation and design is a promising way to exploit seemingly mature markets even further, although, of course, increased costs must also be taken into account.

In addition, we also found that *other* potential customers (i.e., non-designers) liked user-designed products. The other potential customers were not informed about the source of the design of those products, yet the mean WTP for "toolkit watches" is equal to the WTP for the bestsellers made by professional designers. We can only surmise how attractive user-designed watches would be if *lead users* (instead of average users; von Hippel 1986) generated new design solutions with a toolkit which allows *real creative input* (instead of the relatively simple toolkit in our study). Therefore, in addition to employing toolkits as a means of individualizing products, manufacturers should consider using toolkits as a new market research method in order to introduce promising new *standard* products or product designs.

A question worth pursuing is *why* users are willing to pay such a high price premium for their self-designed products? Literature on toolkits to date emphasizes the *functional* benefit, that is, adapting a product to suit an individual preference or need. Our finding that product designs are very heterogeneous gives rise to the interpretation that this factor indeed plays an important role in our study. It is still merely an interpretation, and qualitative indicators from the experiments (such as statements from participants) lead us to conjecture that other factors also have an impact on subjective value creation for user-designers. Specifically, the self-designed product not only has a *well-adapted* design, it is also an *individual* design.

Thus, there might be something like "pride of authorship" (Dabholkar and Bagozzi 2002, Dittmar 1992, Lea and Webley 1997). The active role of designing the product oneself is likely to constitute a psychological benefit to users. Everyday examples of an analogous effect can be found in people who hang up 5,000-piece jigsaw puzzles they have completed themselves instead of hanging up pictures, although objectively jigsaw puzzles look less attractive than simple (and much cheaper) posters.

The self-designed product is also *unique*, and it has been found that people attribute greater value to products that are unique than to ones that are common, all other things – particularly the objective value – being equal (Brock 1968; Fournier 1991).

In addition to the output, the *process* itself might also be a source of subjective value. It is likely that users enjoy the design process due to a "flow" experience (Csikszentmihalyi 1996) and the joy of performing an artistic and creative act. This would have strong implications for toolkit development (e.g., creating an entertaining process with larger solution spaces). Many examples can be found on the do-it-yourself

market, where many activities would be incomprehensible if no specific benefits associated with either the process or the outcome outweighed the direct and opportunity costs (Banks 1998, Toffler 1989).

Finally, it must be noted that the sunk costs of time spent on designing, some notion of fairness (custom must be more difficult, so it is fair to pay more) or simple expectations (prior life experience tells us that individual products are more expensive), and other psychological explanations might also play a role. Future research should separate these effects both theoretically (see Schreier 2003 for a very recent attempt) and empirically. Knowledge about the sources of the WTP premium is crucial to our understanding of the phenomenon of customer integration with toolkits for user innovation and design, as these sources constitute success factors for the toolkits themselves. Studies that aim to address this topic should measure WTP for both self-designed and standard products at the individual level user's level. Otherwise it will not be possible to explain the increment using independent variables (such as the ones mentioned above) using a regression or structural equation model. As such methods are sensitive to outliers, we propose the use of Vickrey auctions to measure WTP.

Other opportunities for further research can be derived from certain limitations of our study. For example, we use students as subjects for research. While this is a common method (see Cooper et al. 1999; Höst, Regnell and Wohlin 2000), the overall population of watch buyers is much larger and far more diverse. Thus, it should be rewarding to replicate our study using a different sampling frame.

Furthermore, we only analyzed the dyad of "user and toolkit." However, Jeppesen (2003), Franke and Shah (2003), and Piller et al. (2003) discuss the benefits of user communities which cooperate in design and innovation activities. Thus, it would be

interesting to examine the effects of collaborative design by users with regard to the final product, process satisfaction, heterogeneity, and WTP.

Our empirical analysis focused on users designing watches. Though watches are a very common product and their market is characterized by high heterogeneity of demand, it would be worth investigating whether the findings drawn from our study also apply to other industries, such as automobiles, computers, clothing, footwear, or even self-service applications (Dabholkar and Bagozzi 2002, Dellaert and Stremersch 2003). Toolkits in these product areas are not limited to aesthetic variability but also allow an individualized fit (measurements) and functions. Some allow true innovation. The WTP increment is likely to be even higher than in the case of our simple toolkit.

Acknowledgements

The authors would like to thank Steffen Wiedemann, Martin Schreier, Marion Pötz and the students of the "E&I Research" course (2002/2003) for their valuable assistance.

We are also indebted to Helmut Strasser (Chair of Experimental Mathematics and Statistics, Vienna University of Economics and BA) for his contributions regarding the concept of entropy. Also, we would like to express our gratitude to Abbie Griffin and the two anonymous reviewers, whose comments helped us to improve the article significantly. Finally, the authors would like to thank the "*Jubilaefonds der Oesterreichischen Nationalbank*" for funding this research project.

References

- Banks, Erik (1998). The Prosumer and the Productivity Paradox. *Social Policy* 8 (2): 10-14 (Summer).
- Bettis, Richard A. (1991). Strategic Management and the Straightjacket. *Organization Science* 2 (3): 315-319.
- Brock, T.C. (1968). Implications of Commodity Theory for Value Change. In: A.G. Greenwald, T.C. Brock, and T.M. Ostrom (Eds.), *Psychological Foundations of Attitudes* (pp. 243-275). New York: Academic Press.
- Cairncross, Frances. (2000). All Yours: The Dream of Mass Customization. *The Economist* April 1st.: 59-61.
- Cooper, David J., Kagel, John H., Lo, Wei, and Gu, Qing Liang. (1999). Gaming against Managers in Incentive Systems: Experiments with Chinese Managers and Chinese Students. *American Economic Review* 89: 781-804.
- Cooper, Robert G. (1999). *Winning at New Products: Accelerating the Process from Idea to Launch*. 2nd ed., Reading, MA: Perseus.
- Cooper, Robert G. (1979). The Dimensions of Industrial New Product Success and Failure. *Journal of Marketing* 43: 93-103.
- Cox, James C., Roberson, Bruce, and Smith, Vernon L. (1982). Theory and Behavior of Single Object Auctions, In: *Research in Experimental Economics*, Vol. 2, Vernon L. Smith (ed.). Greenwich: JAI Press, 375-388.
- Crawford, C. Merle. (1979). New Product Failure Rates – Facts and Fallacies, *Research Management* 22: 9-13 (September).
- Cusumano, Michael A. and Selby, Richard W. (1995). *Microsoft Secrets*. New York, The Free Press.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York: HarperCollins Publishers.
- Dabholkar, Pratibha A. and Bagozzi, Richard. (2002). An Attitudinal Model of Technology-Based Self Service: Moderating Effects of Consumer Traits and Situational Factors. *Journal of the Academy of Marketing Science* 30 (3): 184-201.
- Dahan, Ely and Hauser, John. (2002). The Virtual Customer. *Journal of Product Innovation Management* 19: 332-353.
- Dellaert, Benedict G.C. and Stremersch, Stefan. (2003). Modeling the Consumer Decision to Mass Customize. *Proceedings of the 2003 World Congress on Mass Customization and Personalization (MCPC 2003)*. Munich: TUM.
- Dittmar, Helga. (1992). *The Social Psychology of Material Possessions: To have is to be*. New York: St. Martin's Press.
- Dougherty, Deborah. (1990). Understanding New Markets for New Products. *Strategic Management Journal* 11 (1): 59-78.

- Fournier, S. (1991). A meaning-based framework for the study of consumer-object relations. In: R.H. Holman and M.R. Solomon (Eds.) *Advances in Consumer Research*. Provo: Association for Consumer Research (Vol. 18, pp. 736-742).
- Franke, Nikolaus and Piller, Frank T. (2003). Key Research Issues in User Interaction with User Toolkits in a Mass Customization System. *The International Journal of Technology Management* 26 (5/6): 578-599.
- Franke, Nikolaus and Reisinger, Heribert. (2003). Remaining Within-Cluster Variance: a Meta-Analysis of the "Dark Side of Cluster Analysis. *Working Paper, Vienna University of Economics and Business Administration*.
- Franke, Nikolaus and von Hippel, Eric. (2003a). Satisfying Heterogeneous User Needs via Innovation Toolkits: The Case of Apache Security Software. *Research Policy* 32: 1199-1215.
- Franke, Nikolaus and von Hippel, Eric. (2003b). Finding commercially attractive user innovations: An exploration and test of lead user theory, *MIT Working Paper*.
- Griffin, A. Belliveau, P. Somermeyer, S. (2002). *The PDMA Toolbook for New Product Development*. New York: Wiley Publishing.
- Gensch, Dennis H, Soofi, Ehsan S. 1995. Information-theoretic estimation of individual consideration set. *International Journal of Research in Marketing* 12 (1) 25-39.
- Herniter, Jerome D. (1973). An entropy model of brand purchase behavior. *Journal of Marketing Research* 10 (4): 361-376.
- Hill, Kimberly. (2003). Customers love/hate customization. *CRM-Daily.com* (online e-zine), April 10, 2003.
- Hoffman, Elizabeth, Menkhaus, Dale J., Chakravarti, Dipankar, Field, Ray A., and Whipple, Glen D. (1993). Using Laboratory Experimental Auctions in Marketing Research: A Case Study of Packaging for Fresh Beef. *Marketing Science* 12 (3): 318-338.
- Holmes, Maurice. (1999), Product Development in the New Millennium – a CIPD Version. In: *Proceedings of the Product Development Management Association Conference 1999*. Marco Island: PDMA (October).
- Höst, Martin, Regnell, Björn, and Wohlin, Claes. (2000). Using Students As Subjects: A Comparative Study of Students and Professionals in Lead-Time Impact Assessment. *Empirical Software Engineering* 5 (2): 201-214.
- Huffman, Chynthia and Kahn, Barbara. (1998). Variety for Sale: Mass Customization or Mass Confusion. *Journal of Retailing* 74: 491-513.
- Ishii, Kazuyo and Takaya, Ichimura. (1992). A Process Design Approach Based on the Fusion Model. *Technovation* 12: 499-508.
- Jeppesen, Lars B. (2002). The Implications of "User Toolkits for Innovation", *Working Paper of the Copenhagen Business School*.
- Kamali, Narges and Loker, Suzanne. (2002). Mass Customization: On-line consumer involvement in product design. *Journal of Computer-Mediated Communication* 7 (4) (July).

- Khalid, Halimahtun M. and Helander, Martin G. (2003). Web-Based Do-It-Yourself Product Design, In: *The Customer Centric Enterprise: Advances in Mass Customization & Personalization*, Mitchell M. Tseng and Frank T. Piller (eds.). New York/Berlin: Springer, 447-266.
- Lea, Stephen E.G. and Webley, Paul. (1997). Pride in Economic Psychology. *Journal of Economic Psychology* 18: 323-340.
- Leonard-Barton, Dorothy. (1995). *Wellsprings of Knowledge*. Boston: Harvard Business School Press.
- Lonsdale, Ronald T., Noel, Noel M., and Stasch, Stanley F. (1996). Classification of Sources of New Product Ideas, In: *PDMA Handbook of New Product Development*, Milton D. Rosenau (ed.). New York: Wiley, 179-194.
- Maes, Pattie. (1994). Agents that Reduce Work and Information Overload. *Communications of the ACM* 37 (7): 31-40, 146 (July).
- Miller, George A. (1956). The Magic Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *The Psychological Review* 63: 81-97.
- Mitchell, Robert C. and Carson, Richard T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington: Resources for the Future.
- Neumann, John. (1955). Can We Survive Technology? *Fortune* 91 (6): 106 (June).
- Ogawa, Susumu. (1998). Does Sticky Information Affect the Locus of Innovation? Evidence from Japanese Convenience-Store Industry. *Research Policy* 26: 777-790.
- Park, C. Whan, Jun, Sung Youl, MacInnis, Deborah J. (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.
- Piller, Frank T. (2003). *Mass Customization*. 3rd edition, Wiesbaden: Gabler.
- Piller, Frank (2004). Analysis: Why Levi Strauss Finally Closed it's "Original Spin" MC Operations. Mass Customization News, online newsletter. 7(1) (February).
- Piller, Frank T., Koch, Michael, Moeslein, Kathrin, and Schubert, Petra. (2003). Managing High Variety: Using Communities and Personalization to Overcome the Mass Confusion Phenomenon, In: *Proceedings of the EURAM 2003 Conference*. Milan: European Academy of Management.
- Pine, Buddy J. (1993). *Mass Customization*. Boston: Harvard Business School Press.
- Polanyi, Michael. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: University of Chicago Press.
- Polley, Douglas and Van de Ven, Andrew H. (1995). Learning by Discovery during Innovation Development. *International Journal of Technology Management* 11: 871-883.
- Poolton, Jennifer and Barclay, Ian. (1998). New Product Development From Past Research to Future Applications. *Industrial Marketing Management* 27: 197-212.
- Rangaswamy, Arvind and Lilien, Gary L. (1997). Software Tools for New Product Development. *Journal of Marketing Research* 34 (1): 177-184 (February).

- Schoormans, Jan P.L., Ortt, Roland J., and de Bont, Cees J.P. (1995). Enhancing Concept Test Validity by Using Expert Consumers. *Journal of Product Innovation Management* 12: 153-162.
- Schreier, Martin. (2003). Value Creation by Toolkits for User Innovation and Design. *Working Paper, Vienna University of Economics and Business Administration.*
- Shanklin, William L. and Ryans Jr., John K. (1984). *Marketing High Technology.* Lexington, MA: Lexington Books.
- Shannon, Claude and Weaver, W. (1949). *The Mathematical Theory of Communication.* Evanston: University of Illinois Press.
- Sharma, Arun and Sheth, Jagdish N. (2004). Web-based Marketing. The Coming Revolution in Marketing Thought and Strategy. *Journal of Business Research* 57: 696-702.
- Stump, Rodney, Athaide, Gerard, and Joshi, Ashwin W. (2002). Managing Seller-Buyer New Product Development Relationships for Customized Products. *Journal of Product Innovation Management* 19: 439-454.
- Teas, Kenneth R. (1994). Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment. *Journal of Marketing* 58 (1): 132-139.
- Teresko, John. (1994). Mass Customization or Mass Confusion. *Industry Week* 243 (12): 45-48.
- Thomke, Stefan and von Hippel, Eric. (2002). Customers as Innovators: A New Way to Create Value. *Harvard Business Review* 80 (2): 74-81 (April)
- Thomke, Stefan, von Hippel, Eric, and Franke, Roland. (1998). Modes of Experimentation: An Innovation Process - and Competitive - Variable. *Research Policy* 27: 315-332.
- Toffler, Alvin. (1989). *Power Shift.* New York: Bantam Books.
- Tseng, Mitchell M. and Jiao, Jianxin. (2001). Mass Customization, In: *Handbook of Industrial Engineering*, 3rd edition, Gaviel Salvendy (ed.). New York: Wiley, 684-709.
- Tseng, Mitchell M. and Piller, Frank T. (2003). *The Customer Centric Enterprise: Advances in Mass Customization & Personalization*, New York/Berlin: Springer.
- Vachani, Sushil (1991). Distinguishing Between Related and Unrelated International Geographic Diversification: A Comprehensive Measure of Global Diversification. *Journal of International Business Studies* 22 (2): 307-323.
- Vickrey, William S. (1961). Counter Speculation: Auctions and Competitive Sealed Tenders. *Journal of Finance* 16 (1): 8-37 (March).
- Von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science* 32: 791-806.
- Von Hippel, E. 1988. *The Sources of Innovation.* New York: Oxford University Press.
- Von Hippel, Eric. (1994). Sticky Information and the Locus of Problem Solving: Implications for Innovation. *Management Science* 40 (4): 429-439 (April).

- Von Hippel, Eric. (1998). Economics of Product Development by Users: The Impact of “Sticky” Local Information. *Management Science* 44 (5): 629-644 (May).
- Von Hippel, Eric. (2001). Perspective: User Toolkits for Innovation. *The Journal of Product Innovation Management* 18: 247-257.
- Von Hippel, Eric and Katz, Ralph. (2002). Shifting Innovation to Users via Toolkits. *Management Science* 48 (7): 821-834 (June).
- Von Hippel, Eric and Tyre, Marcie. (1995). How Learning is Done: Problem Identification in Novel Process Equipment. *Research Policy* 24: 1-12.
- Wertenbroch, Klaus and Skiera, Bernd. (2002). Measuring Consumer Willingness-to-Pay at the Point of Purchase. *Journal of Marketing Research* 39: 228-241.
- Wind, Jerry and Rangaswamy, Arvind. (2001). Customerization: the next revolution in mass customization. *Journal of Interactive Marketing* 15 (1): 13-32 (Winter).
- Zipkin, Paul. (2001). The Limits of Mass Customization. *Sloan Management Review* 42 (3), 81-87 (Spring).

Tables

Table 1: Overview of Experiments

	<i>Sample</i>	<i>Acceptance Rate</i>	<i>Action</i>	<i>Method of measuring WTP</i>
Experiment A	n=165	55%	Design of individual watch	CVM
Experiment B	n=248	50%	Inspection of user-designed watches (and standards) displayed on a poster	CVM
Experiment C	n=102	45%	Design of individual watch	Vickrey auction
Experiment D	n=202	50%	Inspection of user-designed watches (and standards) displayed on a poster	Vickrey auction

Table 2: Number of Standard Watches Necessary

<i>Decision 2: Satisfaction level of individual customer ^b</i>	<i>Decision 1: Share of customers ^a</i>					
	<i>0%</i>	<i>20%</i>	<i>40%</i>	<i>60%</i>	<i>80%</i>	<i>100%</i>
<i>0%</i>	1	1	1	1	1	1
<i>20%</i>	1	2	4	7	12	22
<i>40%</i>	1	2	8	17	31	58
<i>60%</i>	1	4	16	33	58	91
<i>80%</i>	1	10	35	68	101	134
<i>100%</i>	1	27	60	93	126	159

^a Percentage of customers whose preferences are meant to be met (n=165)

^b Minimum satisfaction for the individual customer (100% = customer gets exactly the face, strap, case, hour/minute hand and second hand he wants, 80% = customer gets a watch that meets his preferences in only 4 out of the 5 dimensions, 60% = customer gets a watch that meets his preferences in only 3 out of the 5 dimensions, etc.

Cells: Number of standard watches necessary to achieve both objectives.

Source: Experiment A

Table 3: Entropy of User-Designed Watches

<i>One dimension</i>			<i>Two dimensions</i>			<i>Three dimensions</i>		<i>Four dimensions</i>		<i>Five dimensions</i>	
object	% of max. entropy ^a	p ^b	object	% of max. entropy ^a	p ^b	object	% of max. entropy ^a	object	% of max. entropy ^a	object	% of max. entropy ^a
F	88.8	<.001	F+C	97.0	<.001	F+C+S	98.4	F+C+S+H	99.4	F+C+S+H+Sec.	99.6
C	89.2	<.001	F+S	94.3	<.001	F+C+H	99.2	F+C+S+Sec.	99.1		
S	73.7	<.001	F+H	96.6	<.001	F+C+Sec	98.8	F+C+H+Sec.	99.6		
H	83.3	<.001	F+Sec.	95.7	<.001	F+S+H	98.1	F+S+H+Sec.	98.9		
Sec.	87.4	<.001	C+S	86.9	<.001	F+S+Sec.	97.1	C+S+H+Sec.	96.9		
			C+H	92.8	<.001	F+H+Sec.	98.5				
			C+Sec.	91.6	<.001	C+S+H	95.2				
			S+H	87.7	<.001	C+S+Sec.	93.3				
			S+Sec.	87.6	<.001	S+H+Sec.	96.2				
			H+Sec.	91.1	<.001	S+H+Sec.	93.7				

F = face, C = case, S = strap, H = hour/minute hand, Sec. = second hand

^a Empirical entropy coefficient divided by maximum entropy of system

^b Test of null hypothesis: "there is no concentration in the data."

Source: Experiment A

Table 4: Results of the Willingness-to-Pay Analysis

WTP of ...	WTP (in euros) for							
	<i>User-designed watch</i>	<i>Standard Watch 1^a</i>	<i>Standard Watch 2^a</i>	<i>"Ideal" Watch^b</i>	<i>(Best-selling) Standard 3^c</i>	<i>(Best-selling) Standard 4^c</i>	<i>(Best-selling) Standard 5^c</i>	<i>(Best-selling) Standard 6^c</i>
Experiment A (n=165 user-designers, CVM) ^d	<i>mean</i>	48.5	21.5	21.5	92.0			
	<i>(std. dev.)</i>	(50.0)	(13.3)	(22.3)	(105.3)			
	<i>median</i>	40.0	20.0	20.0	75.0			
Experiment B (n=248 other subjects, CVM) ^e	<i>mean</i>	23.1	22.4	23.1	18.5	20.0	25.8	32.7
	<i>(std. dev.)</i>	(11.6)	(15.7)	(17.7)	(13.4)	(16.3)	(20.8)	(19.1)
	<i>Median</i>	21.8	20.0	20.0	15.0	20.0	25.0	30.0
Experiment C (n=102 user-designers in Vickrey auction) ^f	<i>mean</i>	15.5						
	<i>(std. dev.)</i>	(18.9)						
	<i>median</i>	7.0						
Experiment D (n=202 other subjects in Vickrey auction) ^g	<i>mean</i>		5.5	5.9	4.3	5.1	7.0	14.9
	<i>(std. dev.)</i>		(9.7)	(10.2)	(7.8)	(8.5)	(11.3)	(15.5)
	<i>median</i>		1.0	2.0	0.0	0.5	2.0	10.0
Ratio of WTP measured by Vickrey auction and CVM	30.7%	25.6% (Exp. A + D)	27.4% (Exp. A + D)		23.2% (Exp. B + D)	25.5% (Exp. B + D)	27.1% (Exp. B + D)	45.6% (Exp. B + D)

^a Often-bought standard watches in same price and quality segment as user-designed watches.

^b "The ideal watch you can imagine designing with the perfect toolkit" (imaginary option to freely change form of cases, faces, colors etc. at the same technical standard and quality of product).

^c Most often sold standard products ("bestsellers") within same price and quality segment.

^d All user-designers were asked about their WTP for their self-designed watch, for the standard watches (all displayed on the computer monitor), and for the imaginary "ideal" watch.

^e All subjects were asked about their WTP for 30 watches (displayed in full color and size on a poster). Beyond the standard watches (which were of course constant for each subject), subjects were asked about their WTP for a random selection from the 165 user-designed watches. Subjects were not aware of the different sources of the watch designs. The order of stimuli was random.

^f Each user-designer was asked for his bid (WTP) for the self-designed watch (displayed on the computer monitor) and knew that only the top 10% of the bids would actually get the watch.

^g All subjects were asked for their bids (WTP) for the standard watches (displayed on a poster), and knew that only the top 10% of the bids would actually get the watch.

Sources: Experiments A, B, C, and D

Figures

Figure 1: The Idtown.com Toolkit

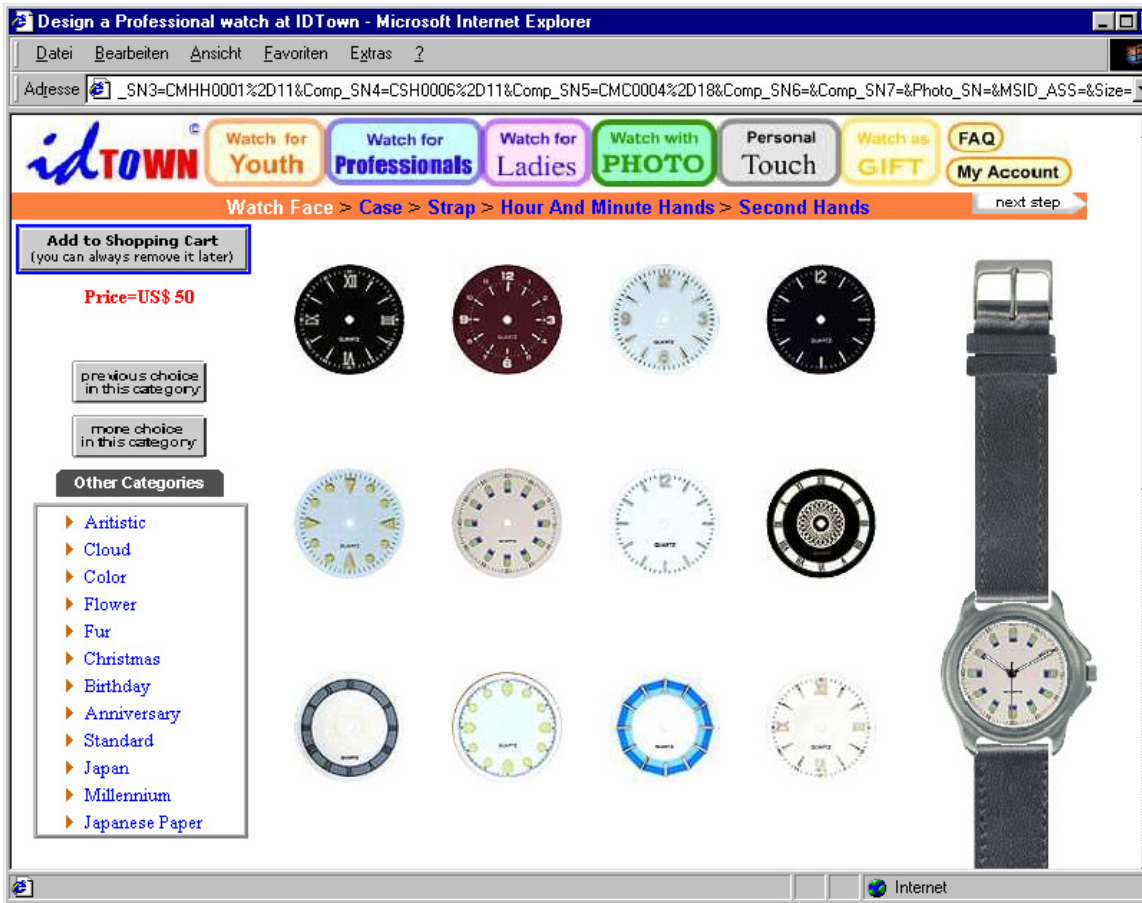


Figure 2: Standard Watches Used in Experiments



¹ Levi's closed its "Original Spin" (mass customization) operations in October 2003, after being in this business for almost ten years. Customers received the program quite well and happily paid the premium of about ten to twenty percent compared to the standard products. Analyses show that Levi's lacked, among other problems, a functioning toolkit to help users to capture the entire value for themselves, and also for the company. Thus, costly and error-prone interactions between the company and its customers were rather the norm than the exception, leading to a rather unstable business model.

Mattel, another pioneering company in the field, abandoned its customized "MyDesign Barbie" as well, though the company had a rather sophisticated toolkit for children users on the internet. In an interview conducted by the authors, one manager told us that the reason for stopping the program was indeed too much user feedback. The 39-dollar customized doll (a premium of about 100 percent) attracted so many orders that the supply chain and fulfillment system was not able to handle all orders in the promised time, leading to dissatisfied customers due to long delivery times. The company was not prepared to capture the customer value it was creating in its manufacturing system. Given the relatively small volume of sales of the customized dolls compared to overall sales volumes, Mattel decided not to invest in its manufacturing and logistics capabilities but merely to keep the toolkit online with the order button. Today, users can still configure and reconfigure dolls, but just for the fun of doing it.

² Probabilities were calculated using Monte-Carlo simulations with 10,000 iterations because the number of empty cells exceeded acceptable sizes and thus conventional chi-square tests are misleading. Corresponding simulations for multivariate concentrations (based on hierarchical log linear models) would involve so much programming (because of empty cells and the chi-square problem mentioned above) that we decided not to calculate them.