Preference Inconsistency in Multidisciplinary Design Decision Making

Erin F. MacDonald
Sloan School of Management,
Massachusetts Institute of Technology,
Cambridge, MA 02142
e-mail: erinmacd@mit.edu

Richard Gonzalez
Department of Psychology,
University of Michigan,
Ann Arbor, MI 48109-1043
e-mail: gonzo@umich.edu

Panos Y. Papalambros
Department of Mechanical Engineering,
University of Michigan,
Ann Arbor, MI 48104-2125
e-mail: pyp@umich.edu

A common implicit assumption in engineering design is that user preferences exist a priori. However, research from behavioral psychology and experimental economics suggests that individuals construct preferences on a case-by-case basis when called to make a decision rather than referring to an existing preference structure. Thus, across different contexts, preference elicitation methods used in design decision making can lead to preference inconsistencies. This paper offers a framework for understanding preference inconsistencies, giving three examples of preference inconsistencies that demonstrate the implications of unnoticed inconsistencies, and also discusses the design benefits of testing for inconsistencies. Three common engineering and marketing design methods are discussed: discrete choice analysis, modeling stated versus revealed preferences, and the Kano method. In these examples, we discuss perceived relationships between product attributes, identify market opportunities for a "green" product, and show how people find it is easier to imagine delight rather than necessity of product attributes. Understanding preference inconsistencies offers new insights into the relationship between user and product design. [DOI: 10.1115/1.3066526]

Keywords: customer preference, preference construction, context effect, utility theory, multidisciplinary design, decision making, Kano, discrete choice analysis, conjoint analysis, preference model

1 Introduction

Research in behavioral psychology over the past 30 years has challenged the view that user preferences exist a priori, suggesting instead that people construct preferences as needed on a case-by-case basis in response to the decision at hand. This phenomenon is referred to as preference construction [1]. Such research has shown violation of the utility theory assumptions, such as independence, and that preferences can change in response to question phrasing such as the framing of the decision. In particular, changes in question phrasing can trigger shifts in preference from one option to the other, in what is termed a preference reversal [1].

Preference construction theory has yet to be incorporated into preference elicitation models in engineering design; these models do not currently account for the fact that preference construction can lead to inconsistency in preference measurements. Inconsistency is symptomatic of two or more different preference constructions and may be identified by performing two or more carefully constructed measurements of preference. When collecting preferences from a group of individuals using a common preference elicitation tool, such as a survey, the tool plays a large role in preference outcome. This has implications because many important decisions, such as those of juries and doctors, are prone to inconsistencies based on the context in which their preferences are elicited [2,3]. As multidisciplinary design increasingly integrates engineering models with models from other disciplines, including marketing, sensitivity to the limitations of preference models is increasingly warranted, so that better design insights can be gained. Indeed, as we show in this paper, preference inconsistencies are not necessarily problematic but can be exploited to provide a deeper understanding of the decision process used by people in their product choices.

We begin with a brief literature review, a demonstration of the impact of construction of preference on current design methods, and a summary of preference construction research that has been conducted in the mechanical engineering design community without being previously identified as such. We then frame extant research in preference construction as it applies to engineering design with a set of new conceptual terms that we believe highlight the underlying processes.

Next, we give three examples of inconsistencies important in engineering design that serve to demonstrate the research potential of incorporating preference construction theory into engineering design research: (i) discrete choice survey analysis, with small manipulations in survey formulation causing large inconsistencies in preference; (ii) a combination approach, where discrete choice survey analysis, buy/not buy scenarios, and past purchase information are used to demonstrate inconsistency; and (iii) Kano category classification and capturing the voice of the customer in quality function deployment (QFD). The design of paper towels is used in (i) and (ii), and the design of an electric toothbrush is used in (iii). The first example shows that people perceive a strong relationship between towel "quilting" and absorbency; a more general methodology to detect the "sentinel/crux" attribute relationship in products is given in Ref. [4]. The second example demonstrates that a large market potential for ecofriendly paper towels could be created through the activation of a particular preference construction for a group of users. The third example suggests that people can imagine delight more consistently than necessity. The inconsistent preference construction highlighted in the examples is not meant to discredit the method at hand but to add caution to its use and increased insight to the resulting design knowledge.

Contributed by the Design Automation Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received August 7, 2007; final manuscript received November 22, 2008; published online March 3, 2009. Review conducted by Timothy W. Simpson. Paper presented at the ASME 2007 Design Engineering Technical Conferences and Computers and Information in Engineering Conference (DETC2007), Las Vegas, NV, September 4–7, 2007.

2 Literature Review

Researchers have used context effects to test theories and assumptions of psychology and economic behavior. Slovic [1] and Kagel and Roth [5] offered excellent literature reviews. Preference reversal is well documented in the domain of context effects, observed when different phrasings of a choice question are shown to result in different experimental outcomes. Arguably, if a user appealed to an internal database through a query, the same preference should emerge independent from particular contextual variables. The finding that different but systematic choices result from manipulations of context is taken as evidence that users do not query a database but rather construct their preferences.

A classic example is that it is possible to construct pairs of lotteries with the property that many people, when asked at what price they would be willing to sell (or buy) the lotteries, put a higher price on one, but when asked to choose which they would prefer to participate in, choose the other [5]. Display effects, such as horizontal versus vertical positioning of choice sets, have also been demonstrated [5]. In multiple-choice surveys, such as discrete choice, it has been observed that how one feels about an attribute level (say, 35 mpg) depends critically on the competing levels of the other alternatives (say, 30 mpg or 40 mpg) and continuous attributes are biased upward compared with categorical attributes [6]. Examples of continuous attributes are vehicle fuel economy or laptop weight, while categorical attributes are discrete configurations such as number of passengers or USB ports. These findings suggest that when designers collect preference for continuous and categorical attributes together, people will place a greater emphasis on the continuous attributes in their choices.

The experimental economics community has demonstrated that willingness to pay and utility measures are contingent upon different decision contexts and preference constructions. Kagel and Roth [5] and Tversky and Kahneman [7] employed context effects to show that expected utility theory does not provide an adequate descriptive account of people's decisions. The psychology and experimental economics communities have documented violations of both "description invariance" and "procedure invariance," which claim that despite different representations and elicitation procedures, the same preferences should always result given the same choice problem. Camerer, an experimental economist, explained: "[i] nvariance violations are especially troublesome for utility theories…. The most famous violations of description invariance are 'framing effects.' Reversals of preference are induced by changes in the reference points…the most pressing question is whether framing effects are systematic and predictable. The evidence is mixed" [5].

3 Random and Nonrandom Preference Inconsistency

Studies show that judgments can change with mood, weather, and any number of random factors that a researcher cannot measure [8]. The stochastic nature of preference in response to such fluctuations is addressed in preference models by representing choice as a random variable. For example, random utility theory includes a stochastic term allowing for random changes in preferences over repeated decisions [9]. In the studies mentioned in Sec. 2, stochastic preference is represented in the statistical tests applied to the findings in order to determine their significance. A significant finding is one that is evident even under the assumption of stochasticity in people's choices.

We term preference inconsistency of a stochastic nature random preference inconsistency because it is frequently represented as a random variable in models. Furthermore, it is unobserved in the model; namely, researchers do not know the weather, level of hunger, or mood of each subject when they record their preferences, nor do they attempt to collect this information.

The preference inconsistencies documented in the studies mentioned in Sec. 2 are nonrandom in nature: The inconsistency is present at the group level, and the explanation for where inconsistency lies can be included in the model. A nonrandom preference inconsistency exists when a group of users is inconsistent in their preferences in a similar explainable manner; an example is when different contexts or question phrasings of identical choice situations lead to different choices and different model parameter estimates. This differs from preference heterogeneity in which model parameters account for differences in preferences between users or groups of users.

Table 1 Common approaches for identifying preference inconsistencies

Inconsistency identification approach	Preference Measurement I	Preference Measurement II	Subjects
Comparative	Experiment (Version A)	Experiment (Version B)	Different subject groups for the two preference measurements
External	Experiment	Market Data	Same or different subject groups for the two measurements
Internal	Experiment (Question A)	Experiment (Question B)	Same subject group for the two measurements

An analogy can be made here to design of experiments research, in which manufacturing parameters are varied in a methodical way to determine how they influence the final process output [10]. In experimental design and analysis, random fluctuation in the output is isolated from nonrandom changes in the output in order to identify the relationships between changes in parameters and changes in output. This is also the case in research on decisions: Using carefully designed experiments, factors influencing choice can be isolated from random fluctuations, and, in doing so, one can gain insight into the decision process. When user decisions about products are under investigation, gaining insight into the decision process leads to learning more about the relationship between the user and product, with the intent of improving the design of the product.

There are three main approaches to gain these insights through the identification of nonrandom preference inconsistencies in the decision theory, economics, and marketing literature. We call the approaches comparative, external, and internal. Table 1 summarizes these approaches, and more explanation is provided in the paragraphs that follow. Sections 6–8 provide an example of each in engineering design. A *comparative inconsistency* study compares preference constructions from different groups of users in response to very similar preference elicitation procedures; it is termed comparative because it requires comparisons of preferences between sets of users. The study includes a carefully designed decision context manipulation such that the versions of the manipulation represent seemingly equivalent decisions (i.e., they are equivalent mathematically). Either the decision or choice phrasing may change, or both. This is the most common approach in decision research, most likely due to the fact that it avoids exposing subjects to multiple versions of the same decision, which itself may influence their choices. Also, the tediousness of repeatedly answering very similar questions can be partially avoided if the experiment is separated across different groups of individuals.

An *external inconsistency* occurs when a group of users exhibits a systematic mismatch between stated preferences in a preference elicitation procedure and revealed preferences gathered from purchase history; it is termed external because it requires the examination of preferences determined outside the preference elicitation process used to identify it. The concept of context manipulation in comparing a real-world decision context to an experimental decision context is akin to the concept of ecological validity from behavioral psychology, which assesses the differ-

ences between conditions in a psychology experiment and real-world conditions and asks if these differences could influence the important findings of the experiment [11]. The weakness of the external consistency test is that there is typically not much knowledge about the specific differences between decision choices elicited during the experiment and revealed in real-world decisions. The researcher may never be able to specify why preference inconsistencies exist between the two situations or may find misleading explanations.

An *internal inconsistency* occurs when a preference structure determined for a group of users in one part of a preference elicitation procedure fails to explain or is contradictory to the preference behavior demonstrated separately by the same group of users in another part of the preference elicitation procedure; it is termed internal because it is identified by examining the results of one preference elicitation procedure, e.g., a survey, without using outside information, such as different versions of the same elicitation procedure or purchase information. The study manipulates decision context in two (or more) different fashions in decision instances within the same survey instrument. It requires careful planning to implement an internal inconsistency experiment successfully. Consider a simple internal test, such as asking the same subject a mathematically identical question in two different phrasings, one after the other. Subjects may infer the purpose and hypotheses of the experiment, which will affect the results. Subjects' responses to one decision may bias their choices in others. The experiment must be designed carefully (for example, it may include tasks that serve as a distraction) in order to minimize such unwanted effects.

A comparative preference inconsistency test faces the problem of mistaking preference heterogeneity for preference inconsistency because the inconsistency cannot be corroborated at the individual level. This inadequacy is extensively examined by Hutchinson et al. [12]. Without the ability to compare subjects' responses in the two conditions directly on a case-by-case basis, researchers can never know if the conclusions they draw hold at the individual level, but they may use demographic or other information to draw inferences about similar groups of subjects, informing them about the nature of the context effect. External and internal preference inconsistency tests can occur at the individual subject level because the subjects make decisions under the different decision context manipulations.

4 Preference Inconsistency in Decision-Based Design

Random preference inconsistency has been addressed previously by Luo et al. [13] and Besharati et al. [14]. They use robust design to address the variance of consumer preference parameters in interdisciplinary marketing and engineering design optimization. As such frameworks allow variance in parameters, random inconsistencies can be dealt with in this approach but nonrandom inconsistencies cannot, as these inconsistencies manifest not necessarily in model variance, but in model parameters, with potential implications in variance and error terms as well. This manifestation will be demonstrated in the discrete choice analysis example in Sec. 6.

Pullman et al. [15] noticed the effects of nonrandom inconsistency, without noting it as such, in their comparison of preferences elicited from QFD and conjoint analysis, stating that optimal products designed using the two different preference elicitation processes varied on important features. They claim the difference stems from "what customers say they want and what managers think will best satisfy customer needs," which is another way of stating that managers are attempting to compensate for preference inconsistencies between point of evaluation, purchase, and use.

The potential effects of nonrandom preference inconsistencies are highlighted in the gray overshadow of two design methodology frameworks proposed by Michalek [16] and Wassenaar et al. [17], Figs. 1 and 2 respectively. The figures demonstrate how preference construction propagates in design processes. Michalek's framework takes only one measurement of customer preference, so designers cannot identify preference inconsistency unless the measurement procedure is conducted across two or more experimental manipulations. The framework of Wassenaar et al. includes three sources of preference construction and, with some modification, can allow designers to identify preference inconsistencies; for example, comparing collected customer preferences and existing market data can identify external inconsistencies.

Designers typically decide among their own conflicting preferences in the design process, and preference structure inconsistencies may lead to different products. Rewording questions can change indifference points and resulting preferences between two alternatives, impacting design decisions based on the hypothetical "equivalents-inequivalents" approach of See et al. [18], who stated that a designer's stated preferences may result in intransitive preference structures. This is a typical result of the context effect experiments mentioned previously. One may combat intransitivity or dampen preference inconsistency by asking designers to explain their preferences, but this leads to confirmation bias and bolstering, in which people reconstruct their preference structures after making a decision in order to rationalize their choice [19,20]. Kulok and Lewis [21] developed a method to correct for random preference inconsistency (referred to as mistakes) without, necessarily, input from the designer. A designer's inequality statements about competing product concepts are analyzed using calculated indicators that show whether values (or utility) of product attributes are consistent across the designer's statements. Inconsistencies are presented to the designer for approval of corrections as suggested by the analysis. Interestingly, Gurnani and Lewis [22] found that relaxing the assumption of rationality by introducing error into a model of designer decisions leads to convergence and optimality in decentralized design, which may parallel the fact that models of decision that include a random variable have been shown to be better predictors of choice.

5 Overview of Survey Used in Sections 6 and 7

Components of a 217-respondent online survey on paper towels are used to demonstrate comparative and external inconsistencies. The survey structure is summarized in Table 2. A sample question from each of the six parts is included in Appendix. The wording of the questions is identical to the web survey instrument, but the formatting is highly condensed. A screen shot of one question from Part I Version A is also included later in Fig. 3. The six-part survey was administered by Luth Research [23] via the Internet. Respondents received one dollar for participating. SAWTOOTH software [24] was used to design the survey and analyze the results. Part I results are summarized in Sec. 6 to demonstrate a comparative inconsistency; results from Parts II, IV, and V are analyzed in Sec. 7 and provide an example of external inconsistency. A companion paper [4] introduces a new methodology based on the identification of comparative preference inconsistency and uses Part I results extensively.

6 Comparative Inconsistency in Discrete Choice Analysis

In this comparative inconsistency study, we test whether subjects infer a relationship between certain product attributes. Complex attributes of high importance to the user we term *crux* attributes, while other attributes used to make inferences about the crux we term *sentinel* attributes. An example is the number of airbags, which is a sentinel attribute for the crux attribute of automobile safety; users perceive a relationship between these two attributes although they may or may not be related. Identifying this relationship helps designers to understand how users evaluate products and construct preferences. For example, if a new automobile design concept was *less safe* if it included more than one airbag, a unique education campaign would have to be launched to aid in the disassociation of number of airbags and increased

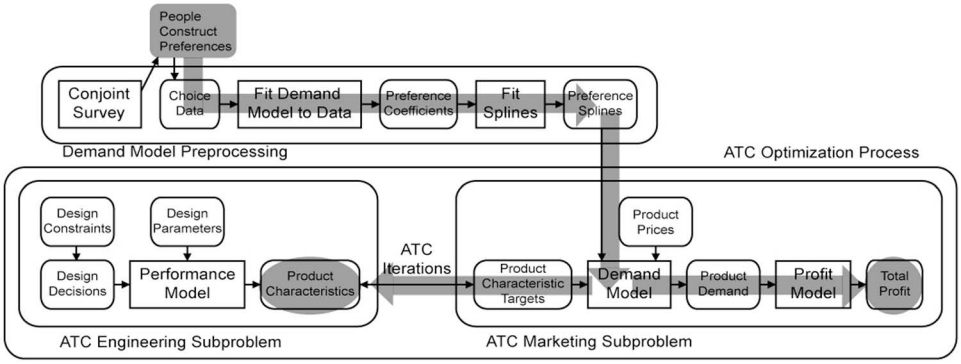

Fig. 1 Propagation of preference construction through Michalek's engineering/marketing ATC Formulation

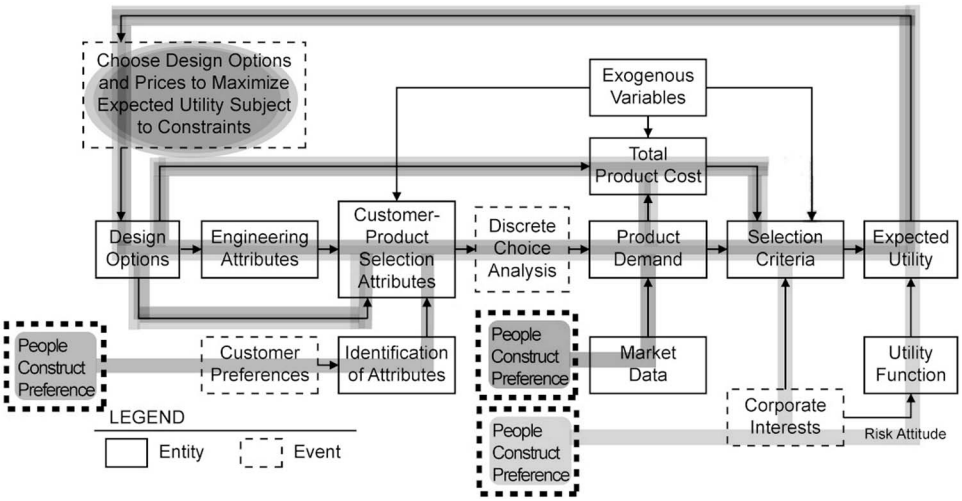

Fig. 2 Propagation of preference construction through decision-based-design flow chart of Wassenaar et al.

safety, and designers would want to quantify the strength of the crux/sentinel relationship before considering the implementation of the alternate design.

A method for rigorously identifying this crux/sentinel relationship is given in Ref. [4]. Here, we focus on a key hypothesis that defines the crux/sentinel relationship: The importance of the sentinel attribute in choice decreases as subjects gain more information about the associated crux attribute. In this case, we investigate if a towel's quilting (sentinel) is associated with its absorbency (crux) through a choice experiment that manipulates related decision parameters. Three groups consisting of 70, 73, and 74 different users were given a discrete choice survey (Part I of the survey described in Sec. 5), choosing between paper towels with different attributes and levels as described in Table 3. An example question from survey Part I, Version A, is shown in Fig. 3. Each version of the survey offers a different manipulation of the information given on towel strength, softness, and absorbency. We expect to find that the utility of quilting will be inconsistent when compared across groups. For survey Version A, where absorbency is not mentioned (yet is a very important product attribute), we expect quilting to have high importance in choice, as quilting will act as a proxy for absorbency.

Group A took survey Version A, which did not mention towel absorbency, softness, or strength. Group B took survey Version B, which was slightly "manipulated:" Absorbency, softness, and strength were mentioned as product attributes but with exactly the same "average" level of 2 out of 3 across all possible answers. This average level information was identical across all profiles, all

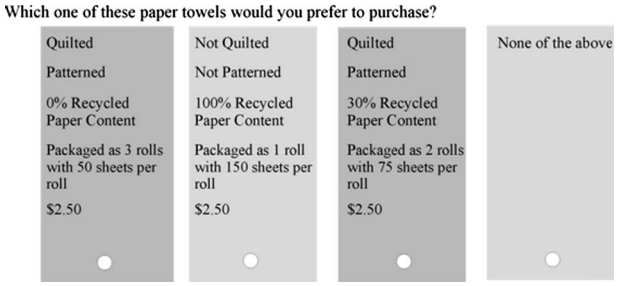

Fig. 3 Example questions from survey Part I, Version A

respondents, and all choice tasks. Group C took survey Version C, in which absorbency, softness, and strength were presented with varying levels: softness, absorbency, and strength, all with ratings of 1, 2, and 3 out of 3. These rating scales were ordinal; the scale points were defined objectively, as the survey began with rating descriptions, e.g., "absorbency: a rating of 1 out of 3 can absorb a 2.5 in. water spill (about the same size around as a tomato slice)." These rating descriptions were available as a pop-up window with every question.

A multinomial (McFadden's conditional) logit model was fitted to each group's answers using SAWTOOTH CBCHB [25]. The model is fitted to subjects' choices in the survey, as shown in Eqs. (1)–(4) below. The probability of choosing any response (product

Table 2 Overview of paper towel survey

	Version A "no quality"	Version B "fixed quality"	Version C "quality as attribute"
Total respondents	70	73	74
Part 1	Attributes:	Attributes:	Attributes:
Stated preference	Quilting	Quilting	Quilting
Conjoint	Pattern	Pattern	Pattern
10 choice tasks	Packaging	Packaging	Packaging
8 conjoint	Recycled paper	Recycled Paper Content	Recycled paper content
2 fixed	content	*All paper towels given	Softness
NOA (none of the above) option		equal softness, strength,	Strength
		and absorbency ratings	Absorbency
Part 2		Attributes:	
Stated preference Conjoint		Price ($1.29, $2.39, $3.49, $4.59)	
6 choice tasks		Strength (1/3, 2/3, 3/3)	
4 conjoint		Softness (1/3, 2/3, 3/3)	
2 fixed		Absorbency (1/3, 2/3, 3/3)	
No NOA option		Recycled paper content (0%, 30%, 60%, 100%)	
Part 3		Rate across brand: Price, strength, softness, absorbency,	
Rating of attributes		environmental responsibility, quilting, pattern, packaging	
		(above average, average, below average, do not know)	
		Rate across brand: Recycled paper content (0%, 30%, 60%, 100%)	
Part 4		Report paper towel last purchased: Price, brand, packaging, quilting,	
Past purchase Information		pattern	
Part 5		Report willingness to pay for three product scenarios:	
Buy/not buy scenarios		(if not willing to pay, explain why not)	
		(wtp1) Quilted, not patterned, 100% recycled paper content, 2 rolls	
		(wtp2) Quilted, patterned, 0% recycled paper content, 2 rolls	
		(wtp3) Quilted, not patterned, 60% recycled paper content, 2 rolls	
Part 6		Gender, age, income, where usually shop for paper towels, zip code,	
Demographic info		Ethnicity, education level	

configuration) j is equal to the probability that responses' utility U_j is greater than that of the other responses presented in the question. This utility has a measurable portion v_j and a random variable error term ε_j. The measurable utility is partitioned into part-worths β for attributes ζ, like pattern and quilting, and further into levels ω, like patterned/not patterned and quilted/not quilted. A dummy variable x indicates whether or not a certain attribute's level is included in a particular product configuration. Assuming normally distributed preference across individual subjects, and that the ε_j are independent and identically distributed (IID) with a double exponential distribution, the logit model is represented as in Eq. (4). For more information on this model, refer to Refs. [26,27].

$$P_j = \mathbf{P}[U_j > U_{j'} \quad \text{for all} \quad j' \neq j] \tag{1}$$

$$U_j = v_j + \varepsilon_j \tag{2}$$

$$v_j = \sum_{\zeta} \sum_{\omega} \beta_{\zeta\omega} x_{j\zeta\omega} \tag{3}$$

$$P_j = \frac{e^{v_j}}{\sum_{j'} e^{v_{j'}}} \tag{4}$$

The estimates of the part-worths β that provide the best fit between model and data were found using Bayesian estimation techniques [28]. The estimated parameters were normalized using the "full factorial marketplace normalization technique," an original method of the authors' explained in more detail in Ref. [4], in order to gain a sense of importance of the different attributes in choices across the three subject groups. Due to the nature of the multinomial logit model, utility values estimated in one model cannot be compared directly to those from another without some sort of normalization. A common approach is to link the models through a scaling parameter that adjusts the variance in the models [27]. While this is appropriate for some applications, when performing statistical tests between models we do not think that it is appropriate to adjust the variance on estimated terms in the models.

Table 3 Attributes and Levels present in Versions A, B, and C of survey Part I

	Strength	Softness	Absorbency	Quilting	Recycled paper content	Packaging	Pattern
Version A (N=70)	N/A	N/A	N/A	Quilted or not quilted	0%, 30%, 60%, or 100%	1, 2, or 3 rolls	Patterned or not patterned
Version B (N=73)	2 out of 3	2 out of 3	2 out of 3	Quilted or not quilted	0%, 30%, 60%, or 100%	1, 2, or 3 rolls	Patterned or not patterned
Version C (N=74)	1, 2, or 3 out of 3	1,2, or 3 out of 3	1,2, or 3 out of 3	Quilted or not quilted	0%, 30%, 60%, or 100%	1, 2, or 3 rolls	Patterned or not patterned

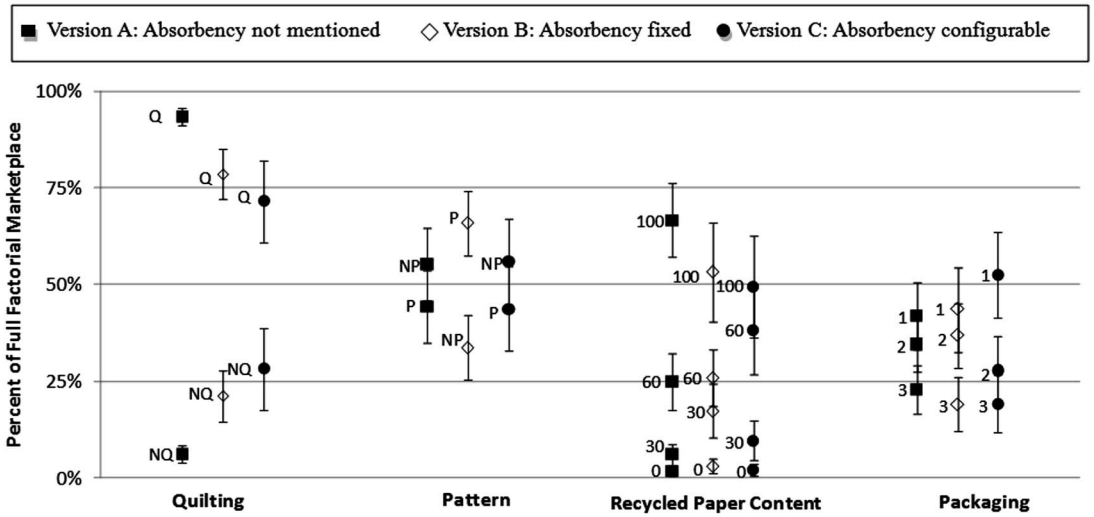

Fig. 4 Aggregated full factorial market shares across survey versions

In the full factorial market, part-worth utilities are used to predict the choice share of every possible product (every combination of attribute/levels), assuming that every possible product is available for selection. Full factorial choice shares are calculations of the utility of each product in the marketplace divided by the sum of the utility of all possible products, thus the full factorial idea provides a normalization of the utility scales across models and groups of decision makers. The full factorial marketplace includes all products j with all possible levels of attributes in all possible combinations, thus avoiding related model artifact problems. Full factorial choice shares are aggregated across all products that contain a particular attribute/level; to create a measurement we term the aggregated full factorial market share, $\hat{P}_{\zeta\omega}$.

$$\hat{P}_{\zeta\omega} = \sum_{j} x_{j\zeta\omega} P_j \tag{5}$$

As with utility part-worths, a higher aggregated full factorial market share means a higher preference, and a larger spread between aggregated market shares for the levels of one attribute means that that attribute is more important in the choice decision than others. The concept of importance is calculated formally as

$$I_{\zeta} = \sum_{\omega} (\hat{P}_{\zeta\omega} - n^{-1})^2 \tag{6}$$

where n equals the number of levels ω for attribute ζ. As mentioned previously, here we will test the hypothesis that the importance of quilting is highest when it stands as a sentinel for absorbency in Version A of the survey:

$$I_{\text{quilting,A}} > \max\{I_{\text{quilting,B}}, I_{\text{quilting,C}}\} \tag{7}$$

Other hypotheses associated with the determination of this relationship are presented in Ref. [4].

Figure 4 shows that the aggregated full factorial market shares for the attributes and levels of quilting, pattern, packaging, and recycled paper content across the three groups. Abbreviated indications of levels are noted on the graph; for example, the levels of packaging are abbreviated as 1 (1 roll), 2 (2 rolls), and 3 (3 rolls). There are four sets of points, one for each attribute. The first column (of squares) in each set displays the aggregated full factorial market shares for Version A of the survey, the second for B, and the third for C. The larger the share of the full factorial market any level has (represented by the points on the graph), the more that level was preferred in the subject's choices. The more dispersed the points in any given column, the larger the importance calculated per Eq. (6) and the more important that attribute was in the choices subjects made in the corresponding survey version. The error bars included are standard errors associated with the aggregated full factorial market shares, calculated using the covariance matrix on the estimated part-worths. The overlapping error bars for the attributes of pattern and packaging indicate that these attributes do not show a significant pattern across different versions of the survey. Recycled paper content's importance in choice decreases slightly as more information on strength, softness, and absorbency is presented. However, note that 100% recycled paper content is always preferred, which would likely not be the case if subjects were inferring that higher recycled paper content meant lower strength, softness, and/or absorbency. This counterintuitive and questionable preference manifests in Part II survey results, which are discussed in Sec. 7.

For quilting, we see that subtle changes in the question wording lead to significant differences in utilities as indexed by the percentage share of the full factorial marketplace. Group A, which knew nothing about the towels' absorbency, had a relatively high preference for quilted paper towels. Absorbency was included as average across all choices for Group B, and their preference for quilting was significantly less. For Group C, which saw varying levels of absorbency in their available towel choices, quilting approaches the importance of the more trivial attributes of pattern and packaging. This difference in preference structure between Versions A and B is interesting because these versions shared exactly the same choice options. This comparative preference inconsistency is summarized in Table 4.

The calculated importance of the attribute quilting is 0.38, 0.16, and 0.09 for Versions A, B, and C, respectively. The hypothesis in Eq. (7) is found to be statistically significant ($p < 0.05$). Testing this as a set of five hypotheses discussed in Ref. [4], quilting is found to be a sentinel attribute for absorbency.

A simple main effect analysis was also performed using SAWTOOTH's SMRT "Counts" interface [29]. The independent variables were the attributes, and the dependent variable was choice.

Table 4 Comparative preference inconsistency

Inconsistency identification approach	Preference Measurement I (Version A)	Preference Measurement II (Version B)
Comparative	Quilted full factorial market share: 94%	Quilted full factorial market share: 74%

In Version A, recycled paper content and quilting were significant predictors of choice ($p<0.01$), and pattern and packaging were not significant predictors of choice. In Version B, recycled paper content and quilting were significant predictors of choice ($p<0.01$), and pattern and packaging were not significant predictors of choice. In Version C, strength, softness, absorbency, and recycled paper content were significant predictors of choice ($p<0.01$). Quilting was not a significant predictor of choice, nor was packaging or pattern. As an indication that respondents evaluated the attributes as independent from each other, no significant two-way interaction effects were found between attributes in any survey version.

If the integrated marketing and design robust optimization approach of Besharati et al. [14] had been used with Version A of this survey, it would have found that quilting was an important attribute in the design of the paper towel but not so for Versions B and C. Because the nonrandom preference inconsistency affects the parameter and potentially also variance and error, their approach would lead to different robust optima dependent on the construction of preferences. Thus, such utility elicitations do not provide windows into consumers' or designers' true preferences, but this example suggests that people construct their preferences case-by-case and subtle wording changes or different contexts can lead to different utility elicitations.

Surveys that miss crux attributes will give inconsistent preferences for sentinel attributes when compared with surveys that include both crux and sentinel attributes. In the above example, quilting serves as a sentinel for the crux attribute of absorbency. Imagine that designers create a new towel that absorbs twice as much liquid as current towels, but the towel cannot have quilting lines due to structural limitations of the new design. The designers would have to weigh the advantages and disadvantages of introducing the new product: It may be the best product available, but without an expensive advertising campaign to delink quilting and absorbency in the minds of users, the design advantages may unintentionally lead to decreased utility and popularity. Also, consider a budgetary decision where one nonessential attribute, quilting or pattern, must be cut from a towel's design. Marketers may argue that, while quilt lines serve no functional purpose, patterns allow for easy product differentiation, a marginally attractive feature in this price- and brand-driven market. Designers who have a working knowledge of customer perceptions could counter that eliminating quilting may decrease the absorbency of the towel in the users' minds and thus negatively impact associations with the brand's quality. A preference inconsistency experiment enables such design discussion to take place on a scientific and quantitative level using testable hypotheses.

7 External Inconsistency in Ecodesign

In this section, we undertake an external preference inconsistency test as a reality-check on the preferences for recycled paper content reported in Part II of the paper towel survey. We learn that the preferences reported are out of step with actual market behavior. We show how to use other information in the survey to understand why this may be.

An ecofriendly product purchase requires users to make a complex attribute trade-off between public good (preserving the environment) and private good (sacrifice in price or performance, or both). Preference construction for ecofriendly products is prone to external inconsistency between preferences gathered in the design process and preferences exhibited in the marketplace. Three preference construction phenomena can account for these construction inconsistencies. Social desirability bias (SDB) refers to the propensity for people to answer a survey in accordance with an accepted social norm, in the case of ecodesign, in accordance with advancing the public good rather than the private one [30]. *Embedding* occurs when, for example, users state that they are willing to pay a 50 cent premium for a recyclable yogurt container in a survey on yogurt, when, in fact, the user is only willing to add 50 cents to their entire weekly shopping bill to purchase ecofriendly goods [31]. In other words, such premiums do not necessarily scale for an individual's multiple purchases. *Pseudosacredness* occurs when the individual's values for the environment are manifested as "sacred" in preference elicitation, i.e., they will not trade them for other desirable qualities; but in another context, their ecovalues are traded off [32].

The following analysis is based on the assumption that people perceive a link between recycled paper content and ecofriendly paper towels. The survey avoided directly evaluating the strength of this link in order to minimize respondent "priming" for thinking about the environment during the survey. We purposefully never mentioned the word "environment," as this could strengthen social desirability bias. We also could not ask respondents for the relative importance of recycled paper content as compared with other attributes in ecofriendly paper towels at the end of the survey, as the large amount of exposure to this attribute during the survey would bias their answers. Furthermore, direct assessment of importance weights has not led to reliable estimates [33]. Instead, we relied on a statistical model of stated decisions to infer importance weights. With respect to the perceived importance of recycled paper content versus other potential ecofriendly attributes, respondents did not care as much about packaging in their product choices as they did about recycled paper content in Sec. 6. Previous choice analysis research on toilet paper found that the average respondent was not willing to pay more for unbleached paper alone but was willing to pay more for recycled paper alone [30]. With respect to the perceived importance of buying products made from recycled materials versus other ecofriendly actions, Guber, in her excellent compilation of polls on environmental issues, stated that environmentalists and nonenvironmentalists alike are more likely to buy products made from recycled materials whenever possible than buy a product because the label or advertising said it was environmentally safe or biodegradable or avoid purchasing products made by a company that pollutes the environment [34].

A multinomial logit model, discussed in Sec. 6, was fitted to Part II of the survey, which was similar to Part I. Refer to Table 2 and the Appendix for a detailed explanation of the discrete choice design. Product choice options were described in terms of softness, strength, absorbency, recycled paper content, and price ($1.29, $2.39, $3.49, and $4.59). Recall that price was previously held constant at $2.50 for Part I of the survey. The part-worth utilities for the attributes and levels from survey Part II are presented in Table 5.

The estimated part-worths for price, recycled paper content, and softness are shown in Fig. 5. Price has the largest range of utility

Table 5 Part-worth utilities estimated from responses to survey Part II

Attribute: Level	Part-worth	Standard error
Softness: 1 out of 3	−1.42	0.29
Softness: 2 out of 3	0.70	0.25
Softness: 3 out of 3	0.72	0.24
Strength: 1 out of 3	−2.50	0.35
Strength: 2 out of 3	0.74	0.27
Strength: 3 out of 3	1.76	0.28
Absorbency: 1 out of 3	−3.37	0.44
Absorbency: 2 out of 3	1.11	0.29
Absorbency: 3 out of 3	2.26	0.31
0% recycled paper content	−2.18	0.42
30% recycled paper content	0.07	0.29
60% recycled paper content	0.84	0.33
100% recycled paper content	1.27	0.35
$1.29	3.34	0.44
$2.39	2.30	0.38
$3.49	−0.14	0.36
$4.59	−5.49	0.70

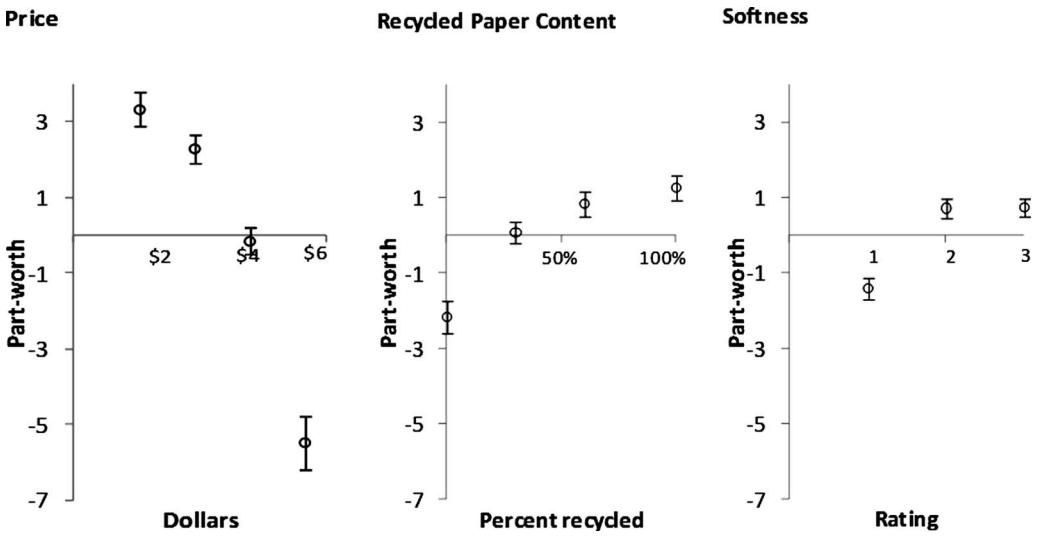

Fig. 5 Estimated part-worths for price, recycled paper content, and softness

values, indicating that it is the largest predictor of choice, with subjects receiving the least utility from $4.59 out of all available attributes' levels. Softness has the smallest range of utility values, with preference estimated as indifferent between a rating of 2 out of 3 and 3 out of 3. Recycled paper content's range lies somewhere in between these two attributes, showing an increase in utility as recycled paper content increases.

Table 6 presents market research on the percentage of people who have used various brands of towels [35]. Green Forest, which 2% of people have used, is the only towel listed that includes 100% recycled paper content (although in some markets, store brand towels may be 100% recycled paper). The leading brands—Bounty, Brawny, Scott, Sparkly, and Mardi Gras—include none. Therefore, the preferences exhibited in Part II of the survey do not mesh with market data, providing an example an external preference inconsistency, described qualitatively in Table 7.

A related external inconsistency can be identified using the results from Part V of the survey, where users are asked what price they would pay for a paper towel with average strength, softness, and absorbency, with 0% recycled paper content, as shown in Table 2 and the sample question from survey Part V in the Appendix. 60 out of 217 users stated that they would not buy the towel for any price due to lack of recycled paper content and/or concerns for the environment. Noting that there were three different versions of Part I of the survey that could have influenced later results, we checked for residual effects of this condition. Table 8 indicates that the 60 respondents with "environmentally worded no buy explanations" were not influenced by the version of Part I of the survey that they saw.

Because they refuse to pay any price, even a heavily discounted price, for the 0% recycled paper towel, these respondents appear to be staunch environmentalists that have associated sacred preferences expressed in their purchases. However, in Part IV of the survey, 52 out of 60 of these respondents reported buying a towel brand that has 0% recycled paper content the last time they went shopping. This conflict in preference is an external preference inconsistency, summarized in Table 9.

Table 6 Brands of paper towels used, January–September 2005. Base: 23,750 adults aged 18+ whose household uses paper towels [35]

Brand	Percentage who have used	Brand	Percentage who have used
Bounty	65%	Hi-Dri	11%
Brawny	36%	Kleenex Viva	9%
Scott Towels	29%	Coronet	6%
Sparkle	23%	Marcal	5%
Mardi Gras	17%	Green Forest	2%
Other brands	15%	Store brand	35%

Table 7 External preference inconsistency, qualitative

Inconsistency identification approach	Preference Measurement I (survey)	Preference Measurement II (market data)
External	100% is the most preferred level of recycled paper content.	Most customers use towels that do not include any recycled paper content.

Table 8 Distribution of "environmentally worded no buy explanations" in survey Part V indicates that responses were not influenced by versions of survey Part I

Part I version	Environmentally worded no buy explanations in Part V	Total respondents
A	19	70
B	20	74
C	21	73

Table 9 External preference inconsistency, quantitative

Inconsistency identification approach	Preference Measurement I (scenario question)	Preference Measurement II (Past purchase)
External	60 respondents opted not to consider purchasing a towel with 0% recycled paper content.	52 out of the 60 respondents purchased a towel brand with 0% recycled paper content the last time they bought paper towels.

It can be partially explained by the fact that only 3 of the 52 users gave a correct rating of 0% recycled paper content for their brand of towel in Part III. We term these 52 users idealists, as they are both idealistic in their own preferences and the configurations of the products they purchase. In identifying the inconsistent preferences of the idealists, we may now theorize that it is not simply social desirability bias that is influencing the part-worth utilities for recycled paper content shown in Fig. 5, but also a lack of education about the products currently available, and perhaps the presence of pseudosacred preferences for some respondents.

It is possible to create a model of preference from the results of survey Part II that separates the part-worth utilities of the idealists from the other respondents. A heterogeneous model can isolate preference inconsistency to one portion of the model, as will be discussed in a future publication. Preference inconsistencies for ecofriendly product attributes can be addressed proactively in design, for example, by including design details to "activate" a set of customer preferences that lead to a large market share shift. Preference inconsistency can also be addressed reactively, by designing an ecofriendly paper towel that is preferred in the market over a variety of preference constructions. A new methodology for proactive and reactive incorporation of an inconsistent preference model in a design optimization framework will also be discussed in the subsequent paper.

The study here highlights the intricacies of using revealed (market) preference data exclusively to assess preference for products that do not yet exist or are unknown to the survey respondents. While such data can be helpful for identifying preference inconsistencies, they can create unnecessary limitations on design possibilities. For example, in this study, assessing only revealed preferences would have neither identified the idealists nor indicated the potential market share change with user education. Without the perspective that preference inconsistencies provide, the towel market appears price- and brand-driven with little room for new competition. The one brand listed in market research with 100% recycled paper content has only been used by 2% of people. The outlook on a green product's chances of success changes when we discover that people may have pseudosacred preference structures paired with a lack of education about recycled paper content in towels. It is this type of investigation that can uncover new design challenges for stagnant markets. In this example, designers are challenged to create a new product attribute that educates about recycled paper content, perhaps designing a visual indication into the product.

8 Internal Inconsistency in the Kano Method

This section investigates an internal inconsistency in an approach similar to the classical decision theory studies mentioned in Sec. 2 in that no further analysis, such as the estimation of part-worth utilities, is required to identify the inconsistency. The experiment manipulates the decision question that is asked of research participants; there is a classification question that asks participants to separate essential product attributes from simply delighting product attributes, and a purchase question that should provide the same separation.

A survey conducted to investigate the effectiveness of the Kano method [36] revealed internal preference inconsistencies in which related questions in the survey were answered in an inconsistent manner for 24% of respondents. Other conclusions from the survey have been presented elsewhere [37]. This web-based survey investigated preferences for electric toothbrushes, studying the product attributes: variety of vibration speeds, brush head replacement indicator, recyclable brush heads, low battery indicator, automatic overcharge protection, and drip catch. The 80 respondents, all students at the University of Michigan, were educated regarding the meaning of the Kano categories—must-be, one-dimensional, delighting, and indifferent—and then were asked to classify the product attributes into these categories and provide written rationale for their classifications. Attributes classifications were explained to respondents exactly as follows:

> Now we would like you to classify product features by category. The different categories are:
>
> Delighting: If a product feature is "Delighting," it means that the feature provides extra product satisfaction for you when it is present, but the product still does its job perfectly well when the feature is absent. An example of a potentially delighting feature in a laptop computer would be a security system that included a fingerprint scanner.
>
> One-dimensional: If a product feature is "One-dimensional," it means that the more attention we, the company, pay to this feature in the design of the product, the more satisfied you, the customer, are with the product and the better the product performs for you. An example of a One-dimensional feature for a laptop computer would be "Weight" (the lighter the laptop, the better).
>
> Must-be: If a product feature is "Must-be," it means the lack of this feature would definitely cause you dissatisfaction, and probably make the product not as useful to you. However, unlike a One-dimensional feature, extra design effort spent on improving a Must-Be product feature would not make much difference to you—it just needs to be included and functioning normally. An example of a Must-be feature of a laptop computer would be a "Spacebar" on the keyboard.
>
> Indifferent: If a product feature is "Indifferent," it means that the feature does not provide either satisfaction or dissatisfaction to you. An example of a potentially indifferent feature of a laptop computer might be infrared communications ability.
>
> If your feelings for the specific product feature do not match any of the above described categories, pick None of the Above.

Respondents' classifications and written explanations for these classifications were examined by independent judges who identified 39 incorrect classifications. Two independent judges were used to identify incorrect classifications, which were removed from further analysis. The judges were given the same attribute descriptions as the respondents and were asked to use these descriptions to identify respondent classification and written explanation pairs that *clearly* indicated that the respondent had misclassified the attribute. The judges had a proportion of overall agreement of 0.90 when independently identifying misclassifications, and discussed discrepancies to 100% agreement. Proportion of overall agreement is a common technique used to judge interrater agreement in qualitative analysis [38]. Only 2 must-be and 5 delighting classifications were removed. Examples are shown in Table 10 below, along with the reclassification suggested by the judges.

In the remaining data, the respondents provided 90 must-be classifications and 181 delighting classifications. A later part of the survey revisited attributes previously classified as must-be and delighting. First, the survey presented a description of a basic toothbrush that included none of the attributes previously mentioned, and told respondents it cost $40.00.

> Basic electric toothbrush:
> - Price: $40.00
> - One bristle vibration speed
> - No indication when brush head needs to be replaced
> - Disposable (throw-away) brush head, non-recyclable
> - Toothbrush simply stops vibrating when it needs to be recharged
> - Must manually stop charging toothbrush to prolong battery life
> - No drip catch that prevents drips from the head of the brush from reaching your hand

Table 10 Examples of misclassifications identified by independent judges

Attribute	Misclassification	Explanation	Reclassification suggested by judges
Brush head replacement indicator	Delighting	Superfluous details. I will know when the brush needs to be replaced by using it. Only I can know this, not the brush.	Indifferent
Variety of vibration speeds	Delighting	Yeah, different speed is necessary. But too much varieties may cause confusion.	Must-be
Low-battery indicator	Must-be	Would be very helpful. Although when the bristles stop, that is as clear an indicator as I need.	Delighting

To review this basic toothbrush description at any point during the rest of the survey, just click the pop-up window on the following pages:

"Click here for the price and description of a basic electric toothbrush"

Then, the survey presented a toothbrush description that included all attributes *except* those they previously classified as must-be (the "no-must-be toothbrush"), and asked new questions: whether or not they would consider buying the no-must-be toothbrush, and if they would, to specify the price they would pay. If they would not consider buying the toothbrush, the survey asked them to explain why, as shown in the sample question below where the must-be attributes are automatic overcharge protection and drip catch:

You are shopping for a new electric toothbrush. The following toothbrushes are available for purchase, amongst other alternatives. For each toothbrush shown, please tell us if you would consider purchasing the toothbrush or not. If you wouldn't consider purchasing it, please tell us why.

Consider the following toothbrush:

- Variety of bristle vibration speeds
- Indicates when brush head needs to be replaced
- Recyclable brush head
- Low-battery indicator
- Must manually stop charging toothbrush to prolong battery life
- No drip catch that prevents drips from the head of the brush from reaching your hand

(Will select one of the following)

○ I would purchase this toothbrush for: $ (Numeric response)

○ I would not consider purchasing this toothbrush because: (Write-in response)

On a separate webpage, the survey gave another description that included all attributes except those they classified as delighting (the "no-delighting toothbrush") and repeated the questions above.

If a user would not consider buying a toothbrush and specifically mentioned lack of an attribute in explaining unwillingness to purchase, this attribute was labeled as a "dealbreaker." If a user would consider purchasing the toothbrush, all attributes *not* included in the scenario were labeled as "negotiable." By examining responses and explanations from users willing to purchase the no-must-be toothbrush, 19 respondents were identified with at least one negotiable must-be attribute, providing a total of 29 must-be negotiable classifications, or 32% of must-be classifications (N=90). One might expect that the 19 users' displeasure could alternatively manifest as a large discount in the price they would quote for the toothbrush lacking their must-be attribute(s). This was not a clear response, as the average willingness-to-pay for the no-must-be toothbrush was $39.37, almost exactly the same price as for the basic toothbrush, with a standard deviation of $13.46.

The same question and evaluation procedure was repeated for the no-delighting-toothbrush, with 161 out of 181 delighting attributes receiving a negotiable status. Only three delighting attribute classifications were identified as dealbreakers. The average purchase price for the no-delighting-toothbrush was $41.30 with a standard deviation of $10.99.

Table 11 summarizes these findings. Delighting attributes should not be dealbreakers when it comes to a product purchase consideration because "the feature provides extra product satisfaction for you when it is present, but the product still does its job perfectly well when the feature is absent." Few (3 out of 181) inconsistencies with this rationale were found during this study. Many (29 out of 90) inconsistencies were found for must-be attributes when purportedly "the lack of this feature would definitely cause you dissatisfaction, and probably make the product not as useful to you." Table 12 presents examples of the statements provided when subjects classified attributes as must-be attributes alongside the prices they later stated they were willing to pay for a toothbrush that did not include this attribute; in the first three examples, strongly worded explanations of why the attribute is must-be are paired next to price premiums over the basic toothbrush.

Note that a classification of "unknown" is given to any attribute (i) not included in a no-must-be/no-delighting scenario that a respondent would not purchase and (ii) respondents did not specifically mention in their written explanation of why they did not purchase under that scenario. Some unknown classifications, therefore, may be silent dealbreakers.

Recall that independent judges checked respondent explanations for must-be and delighting classifications for consistency with the descriptions mentioned previously and that instances of

Table 11 Negotiable and dealbreaking attributes

Attribute	Must-be	Delighting
Cleaned total	90	181
Dealbreakers	29	3
Negotiable	29	161
Unknown	32	17

Table 12 Example must-be explanations that are actually "negotiable"

Attribute	Classification	Explanation	Written-in price for brush that does not include attribute
Automatic overcharge protection	Must-be	Very necessary for me.	$52
Variety of vibration speeds	Must-be	Quite critical when using the product for different area or when some medical conditions come up.	$41
Automatic overcharge protection	Must-be	It is expected for the charger to take care of the battery.	$43
Low-battery indicator	Must-be	If I am brushing my teeth and the thing turns off halfway, I will throw the tooth brush away and never buy it again.	$15

deviation from these descriptions were identified and removed. A negotiable must-be attribute represents a preference inconsistency: in one preference construction, people claim the attribute *must be* or *needs* to be included and functioning normally, while in another, they would buy the product without it. 19 respondents (24%) exhibited this inconsistency. Congruously, a dealbreaker delighting attribute also represents a preference inconsistency, but this inconsistency was found only in one respondent's answers. The preference inconsistency is summarized in Table 13.

The difficulty of consistently imagining necessity versus ease of consistently imagining delight must be kept in mind when eliciting design needs and preferences from users. In an optimization framework, this shift in preference from dealbreaker to negotiable (e.g., from noncompensatory to compensatory) would indicate a fundamental change in the problem formulation, whereby a parameter (possibly a bound in an active constraint) could change to a variable depending on how preferences were elicited.

In the example at hand, it can be seen that designers may exert great effort toward a difficult-to-design yet seemingly essential function identified in the design process, such as overcharge protection, only to discover in the market that users do not find that they need that function. This situation leads to statements such as "the customer does not know what they want," but in light of these findings it is important to distinguish between necessities and other desires, as it would appear the user does know some of what they want, but not always what they need. When working with people during the design process, designers must be cognizant of the limitations of the imagination, and ask questions about product desires in different manners. If inconsistency in answers to these questions is found, it is perhaps best to ask new questions that require increasingly less imagination until consistency in answers is found. For example, performing this study with real prototypes that do and do not include must-be negotiable features would require less imagination and perhaps more consistency in reported needs.

Table 13 Internal preference inconsistency

Inconsistency identification approach	Preference Measurement I (Question Framework A)	Preference Measurement II (Question Framework B)
Internal	Total must-be classifications: 90	Number of these classifications found not to be a "must" for purchase consideration: 29

9 Conclusion

In this paper, we demonstrated insights that can be gained when designers expand their concerns to include not only user preferences but also how these preferences were formed. We focused on three instances of preference construction in product design methods. A crux/sentinel attribute relationship suggests to the designer that keeping certain sentinel attributes in a product may be preferable to users, even if they become superfluous through design advancements, until users adapt to the new design. The designer must keep in mind that if users think a relationship exists between product attributes, this thought will most likely influence their preferences until proven otherwise through education or experience. The highly inconsistent construction of preferences with respect to ecofriendly paper towels highlights the potential market gains that could appear if careful design can trigger the construction of a set of favorable preferences over another less favorable set. Identification of negotiable must-be attributes suggests that stated customer "needs" should be taken as a starting point for building a set of required attributes in a product, rather than as defining criteria. The finding that people seem to consistently imagine delightful attributes but inconsistently imagine necessary ones is intriguing and should be studied further.

Merging research in the behavioral sciences with that in engineering design is a rewarding but delicate task. The preference inconsistency examples provided above do not render the design methods invalid in which they appear. Rather, they provide additional insight for product designers: understanding of the user's ability to assess their own needs and preferences, learning about a customer's perceived associations between product attributes and the manner in which they evaluate product functionality, and increasing the sophistication of a linked engineering and marketing design process to access new market share.

Acknowledgment

This work was partially supported by a National Science Foundation Graduate Research Fellowship, the Donald C. Graham Chair Endowment, and the Antilium Project at the University of Michigan. Survey respondents in Secs. 6 and 7 were recruited by Luth Research (www.luthresearch.com). Software was provided by SAWTOOTH software (www.sawtoothsoftware.com). Martin Backsell contributed to the building and administration of the survey in Sec. 8. Colleagues Elea Feit, Fred Feinberg, and W. Ross Morrow provided helpful suggestions. This generous support and help is gratefully acknowledged.

Appendix: Sample Questions From Survey

Part I

Which one of these paper towels would you prefer to purchase?

Version A choices:

(A) quilted, patterned, 0% recycled paper content, packaged as 3 rolls with 50 sheets per roll, $2.50
(B) not quilted, not patterned, 100% recycled paper content, packaged as 1 roll with 150 sheets, $2.50
(C) quilted, patterned, 30% recycled paper content, packaged as 2 rolls with 75 sheets per roll, $2.50
(D) none of the above

Version B choices:

(A) quilted, patterned, 0% recycled paper content, packaged as 3 rolls with 50 sheets per roll, strength 2 out of 3, softness 2 out of 3, absorbency 2 out of 3, $2.50
(B) not quilted, not patterned, 100% recycled paper content, packaged as 1 roll with 150 sheets, strength 2 out of 3, softness 2 out of 3, absorbency 2 out of 3, $2.50
(C) quilted, patterned, 30% recycled paper content, packaged as 2 rolls with 75 sheets per roll, strength 2 out of 3, softness 2 out of 3, absorbency 2 out of 3, $2.50
(D) none of the above

Version C choices:

(A) quilted, softness 2 out of 3, patterned, absorbency 3 out of 3, 0% recycled paper content, strength 1 out of 3, packaged as 3 rolls with 50 sheets per roll, $2.50
(B) not quilted, softness 3 out of 3, not patterned, absorbency 1 out of 3, 100% recycled paper content, strength 2 out of 3, packaged as 1 roll with 150 sheets, $2.50
(C) quilted, softness 1 out of 3, patterned, absorbency 1 out of 3, 30% recycled paper content, strength 2 out of 3, packaged as 2 rolls with 75 sheets per roll, $2.50
(D) none of the above

Part II

Which one of these paper towel products would you prefer to purchase?

(A) softness 2 out of 3, strength 1 out of 3, absorbency 3 out of 3, 0% recycled paper content, $1.29
(B) softness 3 out of 3, strength 2 out of 3, absorbency 1 out of 3, 100% recycled paper content, $2.39
(C) softness 1 out of 3, strength 2 out of 3, absorbency 1 out of 3, 30% recycled paper content, $4.59

Part III

For each brand of paper towels, please give us your best estimate of the postconsumer recycled content:
Bounty: 0%, 30%, 60%, 100%, or do not know (they select one, repeat for each brand)

Part IV

Thinking back to the last time you purchased paper towels, what brand did you purchase?

(A) Bounty
(B) Brawny
(C) Scott
(D) Viva
(E) Sparkle
(F) Store brand
(G) Seventh Generation
(H) Green Forest
(I) Other
(J) Do not know
(K) Do not regularly purchase paper towels (they select one)

Part V

Consider the following paper towel:
Quilted, patterned, 0% recycled paper, 2 rolls with 75 sheets per roll, softness: 2 out of 3, absorbency: 2 out of 3, strength: 2 out of 3 (they select one)

(A) I would purchase this paper towel for $ (numeric free response)
(B) I would not purchase this paper towel because: (free response)

Part VI

Age (Write-in)

References

[1] Slovic, P., 1995, "The Construction of Preference," Am. Psychol., **50**(5), pp. 364–371.
[2] McGettigan, P., Sly, K., and O'Connell, D., 1999, "The Effects of Information Framing on the Practices of Physicians," J. Gen. Intern. Med., **14**(10), pp. 633–642.
[3] Kerstholt, J., and Jackson, J., 1998, "Judicial Decision Making: Order of Evidence Presentation and Availability of Background Information," Appl. Cognit. Psychol., **12**(5), pp. 445–454.
[4] MacDonald, E., Gonzalez, R., and Papalambros, P., 2007, "The Construction of Preferences for Crux and Sentinel Product Attributes," Proceedings of the 2007 International Conference on Engineering Design, Paris.
[5] Kagel, J., and Roth, A., 1995, *Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ.
[6] Fitzsimons, G., Hutchinson, J., and Williams, P., 2002, "Non-Conscious Influences on Consumer Choice," Mark. Lett., **13**(3), pp. 269–279.
[7] Kahneman, D., and Tversky, A., 1979, "Prospect Theory—Analysis of Decision Under Risk," Econometrica, **47**(2), pp. 263–291.
[8] Sudman, S., Bradburn, N., and Schwarz, N., 1996, *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*, Jossey-Bass, San Francisco.
[9] Luce, R. D., and Suppes, P., 1965, Preference, Utility, and Subjective Probability, *Handbook of Mathematical Psychology*, R. D. Luce, R. R. Bush, and E. Galanter, eds., Wiley, New York, Vol. III, pp. 249–410.
[10] Wu, C. F. J., and Hamada, M., 2000, *Experiments: Planning, Analysis, and Parameter Design Optimization*, Wiley, New York.
[11] Bem, D. J., 1979, "Template Matching: A Proposal for Probing the Ecological Validity of Experimental Settings in Social Psychology," J. Pers. Soc. Psychol., **37**(6), pp. 833–846.
[12] Hutchinson, J. W., Wagner, A. K., and Lynch, J. G., 2000, "Unobserved Heterogeneity as an Alternative Explanation for "Reversal" Effects in Behavioral Research," J. Consum. Res., **27**(3), pp. 324–344.
[13] Luo, L., Kannan, P. K., and Besharati, B., 2005, "Design of Robust New Products Under Variability: Marketing Meets Design," J. Prod. Innov. Manage., **22**(2), pp. 177–192.
[14] Besharati, B., Luo, L., and Azarm, S., 2006, "Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach," ASME J. Mech. Des., **128**(4), pp. 884–892.
[15] Pullman, M., Moore, W., and Wardell, D., 2002, "A Comparison of Quality Function Deployment and Conjoint Analysis in New Product Design," J. Prod. Innov. Manage., **19**(5), pp. 354–364.
[16] Michalek, J. J., 2005, "Preference Coordination in Engineering Design Decision-Making," Ph.D. thesis, University of Michigan, Ann Arbor.
[17] Wassenaar, H. J., Chen, W., and Cheng, J., 2005, "Enhancing Discrete Choice Demand Modeling for Decision-Based Design," ASME J. Mech. Des., **127**(4), pp. 514–523.
[18] See, T., Gurnani, A., and Lewis, K., 2004, "Multi-Attribute Decision Making Using Hypothetical Equivalents and Inequivalents," ASME J. Mech. Des., **126**(6), pp. 950–958.
[19] Tetlock, P., Skitka, L., and Boettger, R., 1989, "Social and Cognitive Strategies for Coping With Accountability: Conformity, Complexity, and Bolstering," J. Pers. Soc. Psychol., **57**(4), pp. 632–640.
[20] Nickerson, R., 1998, "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," Rev. Gen. Psychol., **2**(2), pp. 175–220.
[21] Kulok, M., and Lewis, K., 2005, "Preference Consistency in Multiattribute Decision Making," *2005 ASME Design Technical Conferences*, American Society of Mechanical Engineers, Los Angeles, Paper No. DETC2005-84764.
[22] Gurnani, A., and Lewis, K., 2006, "Decentralized Design at the Edge of Rationality," *2006 ASME International Design Engineering Technical Confer-*

ences, American Society of Mechanical Engineers, Philadelphia, Paper No. DETC2006–99520.
[23] Luth Research website, 2007, http://www.luthresearch.com.
[24] SAWTOOTH software website, 2007, http://www.sawtoothsoftware.com.
[25] 2005, The CBC/HB System for Hierarchical Bayes Estimation Version 4.0 Technical Paper, Sawtooth Software Technical Paper Series, SAWTOOTH Software, Sequim, WA.
[26] Train, K., 2003, *Discrete Choice Methods With Simulation*, Cambridge University Press, Cambridge, Chap. 6.
[27] Louviere, J., Hensher, D., and Swait, J., 2000, *Stated Choice Methods*, Cambridge University Press, Cambridge.
[28] Rossi, P., Allenby, G., and McCulloch, R., 2005, *Bayesian Statistics and Marketing*, Wiley, West Sussex, UK.
[29] 2008, http://www.sawtoothsoftware.com/support/issues/ssiweb/online_help/index.html.
[30] Bennett, J., and Blamey, R., 2001, "Yea-Saying and Validation of a Choice Model of Green Product Choice," *The Choice Modelling Approach to Environmental Valuation*, J. Bennett and R. Blamey, eds., Edward Elgar, Cheltenham, UK, pp. 178–201.
[31] Kahneman, D., and Knetsch, J., 1992, "Valuing Public Goods: The Purchase of Moral Satisfaction," J. Envir. Econom. Manage., **22**(1), pp. 57–70.
[32] Thompson, L., and Gonzalez, R., 1997, "Environmental Disputes Competition for Scarce Resources and Clashing Values," *Environment, Ethics, and Behavior: The Psychology of Environmental Valuation and Degradation*, M. Bazerman, D. Messick, and K. Wade-Benzoni, eds., New Lexington, San Francisco, pp. 75–104.
[33] Goldstein, W., and Howard, C., 1992, "The Relative Importance of Relative Importance: Inferring Other People's Preferences From Relative Importance Ratings and Previous Decisions," Org. Behav. Hum. Decis. Process, **51**(3), pp. 382–415.
[34] Guber, D., 2003, *The Grassroots of Green Revolution: Polling America on the Environment*, MIT, Cambridge, MA, Chap. 8.
[35] Mintel Oxygen, 2006, Mintel International Group Limited, Household Paper Products Report.
[36] Kano, N., Seraku, N., Takahashi, F., and Tsuji, S., 1984, ""Attractive Quality and Must Be Quality," Quality The Journal of Japanese Society for Quality Control, **14**(2), pp. 39–48.
[37] MacDonald, E., Backsell, M., and Gonzalez, R., 2006, "The Kano Method's Imperfections, and Implications in Product Decision Theory," *Proceedings of the 2006 International Design Research Symposium*, Korean Society of Mechanical Engineers, Seoul, Korea.
[38] Spitzer, R., and Fleiss, J., 1974, "A Re-Analysis of the Reliability of Psychiatric Diagnosis," Br. J. Psychiatry, **125**(4), pp. 341–347.