Proceedings of IDETC/CIE 2005
ASME 2005 International Design Engineering Technical Conferences
& Computers and Information in Engineering Conference
September 24-28, 2005, Long Beach, California USA

DETC2005-84765

CAN A HOUSE WITHOUT A FOUNDATION SUPPORT DESIGN?

Andrew Olewnik
Graduate Research Assistant
Department of Mechanical and Aerospace Engineering
University at Buffalo-SUNY
Buffalo, NY USA 14260

Kemper Lewis[1]
Associate Professor
Department of Mechanical and Aerospace Engineering
University at Buffalo-SUNY
Buffalo, NY USA 14260

ABSTRACT

The House of Quality is a popular tool that supports information processing and decision making in the engineering design process. While its application is an aid in conceptual aspects of the design process, its use as a quantitative decision support tool in engineering design is potentially flawed. This flaw is a result of assumptions behind the methodology of the House of Quality and is viewed as an important deficiency that can lead to potentially invalid and poor decisions. In this paper this deficiency and its implications are explored both experimentally and empirically. The resulting conclusions are important to future use and improvement of the House of Quality as an engineering design tool.

KEYWORDS

Quality function deployment, house of quality, design-decision support, design models.

1. INTRODUCTION

By now, most people working in engineering design are aware of the management philosophy known as Quality Function Deployment (QFD) and the primary tool of the philosophy, the House of Quality (HoQ) [1]. At its root, the House of Quality is a conceptual tool for mapping attributes from one phase of the design process to the next. Referring to Fig. 1 as one representation of the design process, an example might be to utilize the HoQ in order to convert a set of "process design" specifications to "manufacturing" specifications in order to produce a particular product. The conceptual mapping provided by the HoQ within the design process is the transfer of information (arrows in Fig. 1) from one node of the design process to the next. This conceptual mapping allows a clear flow of information on a node by node basis in the design process from the identification of "perceived need" node all the way through the "manufacturing" node. This is a valuable tool in helping understand the role of different entities (management, engineering, marketing, etc.) and the general flow and type of information within the design process of Fig. 1. However, there is (in our view) a serious deficiency in the HoQ with potential to affect decisions so early in the design process, that later failures in the design or success of the product are unlikely to be traced to this issue. This deficiency results from the attempt to specify quantitative relationships in the mapping of customer attributes to technical attributes, i.e., mapping from the "perceived need" node to the "specification" node in Fig 1. The focus of this paper is to discuss this deficiency and to explore its effect empirically. Further, an experiment is performed on the HoQ in order to understand the significant factors that lead to final outcomes within the methodology. In the next section, the necessary background of the HoQ is presented to support the discussion and study.

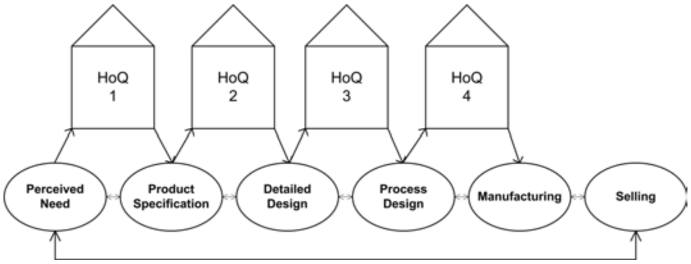


Figure 1 - The design process and the HoQ

---

[1] Corresponding author, kelewis@eng.buffalo.edu, (716) 645 – 2593 ext 2232

2. THE HOUSE OF QUALITY MODEL

QFD began as a management tool in Japan in the early 1970s [1] and in short time became popular within industry in North America at companies like General Motors [1], Ford [1], Xerox [2] and many smaller firms [3]. QFD's main component, the HoQ, is utilized as both a stand alone tool, as exemplified by Kaldate [4] and as a tool integrated in larger design processes, as at Praxair [5], to support product and process design. With such far reaching use and application, the HoQ might be assumed a fundamentally valid design tool.

However, in recent years, many methodologies and models utilized in a design-decision support role have come under scrutiny for flaws in their fundamental mechanics or assumptions [6-9]. Specifically, Barzilai [6] and Saari [7], showed the problems associated with pairwise comparisons and the conflicting decision results generated with methodologies that use such comparisons, like the Analytical Hierarchy Process (AHP). Hazelrigg [8] and Olewnik [9], review the validity of other popular design decision tools and discuss criteria for the validation of such tools in general. Understanding and classifying design models [10], specifically the topic of validation of those models with respect to engineering design is growing in importance both pragmatically [11] and philosophically [12]. The need for validation extends from the physical models utilized by designers [13], to validation of design practices like robust experimental design [14]. The discussion in this paper is related to the validation of design-decision support models [11].

While a valid decision process does not guarantee desirable outcomes, a flawed decision process confounds information used in the decision process and the process itself leaving no means of identifying what is at the root of the bad outcome, the information or the process. The validity of the HoQ and QFD in general has been challenged by Olewnik [9] and Hazelrigg [8], respectively. Though the goal of this paper is not to discuss the finer points of validity in terms of design-decision processes, it is helpful to frame the discussion of this paper in that context.

To support that discussion it is necessary to provide the background on the mechanics of the HoQ. Besides a conceptual mapping, the HoQ also functions as a model for understanding how attributes in one design node affect attributes in the subsequent design node. Consider, Fig. 2 which shows a standard HoQ as taken from [15] and provides obvious explanation for its reference as a "house". The Customer Attributes (CAs) represent what the customer wants in the product. CAs are posed in customer language. The Importance section represents the weight the customer assigns to each CA. The Customer Ratings section represents the customer perception of how well a current product performs on each CA. The ratings may also compare competitor products. Technical Attributes (TAs) represent the product characteristics necessary to meet the CAs. The TAs however, are in engineering design language. The Relationship Matrix is where relationships between CAs and TAs are identified and given a "weak", "medium" or "strong" relationship value. The Technical Test Measures and Technical Difficulty Ratings sections represent designer evaluations among the TAs. Target Value Specifications represent the target level the designers want each TA to reach. The Technical Importance section contains the calculated importance of each TA, which is a function of the Importance values and the values in the Relationship Matrix. Finally, the Correlation Matrix represents a matrix of the interrelationship among TAs.

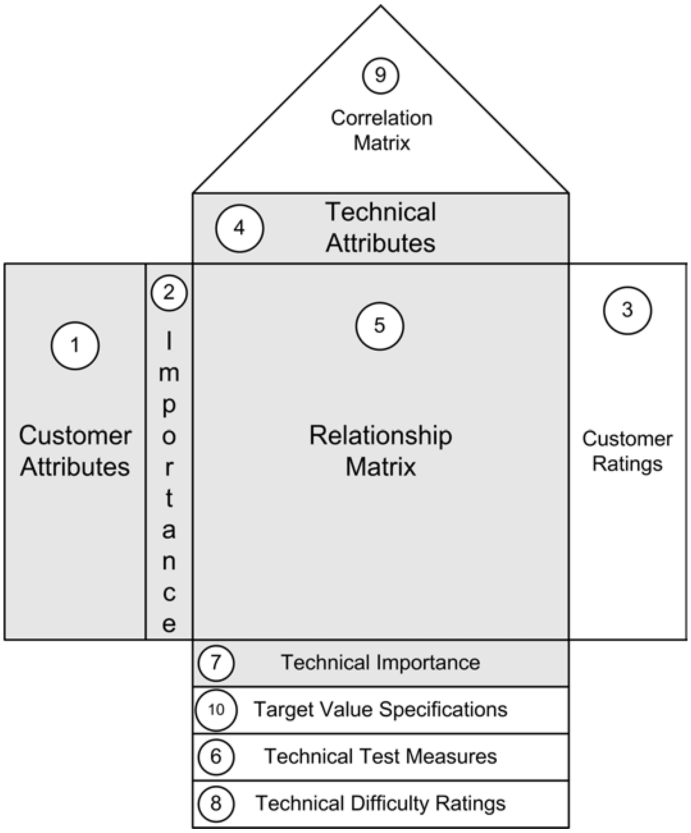


Figure 2 - Standard HoQ

Taking our starting point as the beginning of the design process of Fig. 1, the goal is to translate the "fuzzy voice of the customer" into measurements in the company language [16]. The steps to follow to complete this "translation" are provided by Breyfogle [15]. These steps are labeled in the HoQ of Fig. 2 and are as follows:

1. *Make a list of customer attributes*. This list is usually identified through customer interviews and/or surveys.
2. *Identify the importance of each customer attribute*. This information is also determined from customer surveys.
3. *Obtain customer ratings on existing design and competitor design*.

4. *Designers compile a list of technical attributes to meet the customer attributes*. These attributes should be scientifically measurable terms that can be assigned target values [15] and designers should avoid concept specific terms [16].
5. *Relationships should be identified in the relationship matrix and assigned qualitative value (weak, medium, strong)*. These qualitative relationships are later replaced by a quantitative three number scale.
6. *Technical tests should be performed on existing design and competitor designs to gauge objective measures of difference*.
7. *Importance of each technical attribute should be calculated in either absolute values or relative weights*. This is done using Eq. (1) or Eq. (2) respectively, where there are *m* CAs and *n* TAs and *w* represents the customer importance for the i^{th} CA.

$$\left. raw\,score \right|_{j=1}^{n} = \sum_{i=1}^{m} score_{i,j} \times w_i \qquad \textbf{(1)}$$

$$\left. \begin{matrix} relative \\ weight \end{matrix} \right|_{j=1}^{n} \quad \frac{rawscore_j}{\sum_{j=1}^{n} rawscore_j} \qquad \textbf{(2)}$$

8. *Difficulty of engineering each technical attribute should be assessed*.
9. *The correlation matrix should be filled out*.
10. *Target values for each technical attribute should be set*. This may be based on customer ratings from step 3.
11. *Select technical attributes to focus on based upon technical importance calculations of step 7 and technical difficulty assessment of step 8*.

These primary steps can be carried out in subsequent HoQs used between other stages of the design process. These are the steps for a standard HoQ. Of course simplified and more complex HoQs can be constructed depending on the designers and company utilizing the tool. With this essential background in mind, discussion can turn to the deficiency previously described.

3. THE HOUSE OF QUALITY AND THE DESIGNER

The HoQ is most commonly applied between the "perceived need" and "product specification" nodes of the design process, i.e. the phase described specifically in the steps from Section 2. The role of the HoQ here is critical, as it is meant to *model* the relationship between the customer attributes of a product and the technical attributes of the product. This "language translation" and subsequent characterizations made about the importance of technical attributes based upon that translation is vital to the *potential* success of the product. That is, the HoQ model is meant to identify the most important technical attributes. As long as those technical attributes are the center of the product design, the customer attributes will be satisfied to a level that makes the product desirable and ultimately successful. On a conceptual level, the fundamental mechanics behind the HoQ are well suited to this goal, however, there are two complexities that arise in the implementation of the methodology that raise suspicion about the ultimate value of the results. To investigate these difficulties, it is beneficial to first discuss the implicit assumptions behind the HoQ model as it is implemented between the first two nodes of the design process in Fig. 1.

To aid this discussion of model assumptions, a reduced HoQ is shown in Fig. 3 (including only the gray components of Fig. 2). This section of the house represents the components necessary to support discussion and empirical study in this paper. Note the representation shown in Fig. 3 is representative of the form of all examples used in the paper. In order to fill out this HoQ only a subset of the eleven steps from Section 2 are necessary. Those steps are {1, 2, 4, 5, 7, 11}. The assumptions behind these steps are critical to the results of Technical Importance.

- The first assumption is that the CAs and their individual importance (steps 1 and 2) are truly representative of the potential customer base. The validity of this assumption is a matter of marketing study and not contended herein. However, confidence in these two components is paramount.

- The second assumption is that the TAs (step 4) are the appropriate, measurable product characteristics to meet the CAs. This assumption might be considered the very crux of design. That is, the ability to understand perceived need and convert that understanding into a product or system seems to be the most basic function of a designer. Consider that one definition of design is, *to create or contrive for a particular purpose or effect* [17]. Thus, to take issue with this assumption would be to take issue with the fundamental notion of design. That brings us to the third and fourth assumptions (step 5).

- The third assumption is that designers can indeed identify when a particular TA relates to a particular CA and the qualitative strength of that relationship, i.e. "weak", "medium" or "strong". It is likely that for the most part designers will be able to identify the existence of relationships, especially since they generated the TAs. However, it is possible that some "weak" relationships could be missed due to their subtle nature. The importance of this assumption will become evident in the experimental study. It is also reasonable to believe that designers can distinguish the *qualitative* level of the relationships. However, the fourth assumption deals with a designer’s ability to distinguish the quantitative level of the relationships.

- The fourth assumption is that the designer can later assign *quantitative* values to represent the qualitative relations and further, the quantitative values are always the same. As suggested by Breyfogle [15], an example quantitative scale to utilize might be 1 for weak, 3 for medium and 9 for strong. However, Breyfogle indicates that this is an example possibility, not necessarily the scale to use. It is also not dictated that one quantitative three number scale be used. Instead the choice of quantitative scale(s) is left to the designers to decide. It is worth noting however, that throughout the literature on QFD and the HoQ the most common scale seen is (1-3-9) and that only one three number scale is typically used in any given example.

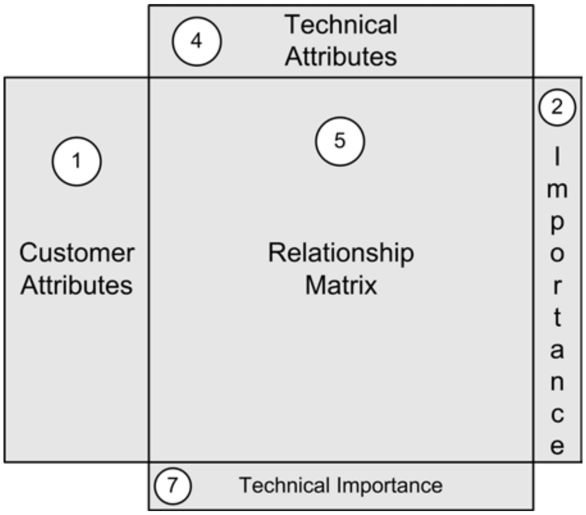


Figure 3 - Reduced HoQ

The use of one three number scale is an understandable simplification in the HoQ model. That is, it would be confusing and difficult for designers to try to apply multiple three number scales throughout. For example, using (1-3-9) across one row of CAs and (2-5-8) across another. However, this simplification points to the larger issue, i.e. that designers have no reason to choose a particular quantitative relationship represented by a three number scale. The fact that the scale consistently appears as (1-3-9) in the literature is further suggestive of this. The assumption that designers can choose an appropriate scale means that they know ahead of time both the range on which the relationship scale lies and the relative difference between weak, medium and strong. Put another way, this assumes the designers can put a quantitative value to reflect how a given TA will affect the perception of customers. It is difficult to accept that designers could indeed make this kind of assessment, yet this is exactly what they must do to generate the final "Technical Importance" as per step 7 from Breyfogle [15]. This assumption is the primary deficiency explored in this paper through experimentation and an empirical study of a HoQ example beginning in the next section.

4. EXPERIMENTING WITH THE HOQ

In order to understand the influence of assumption three and four in filling in the HoQ an experiment was established. Recall, assumption three is that the designers can identify where relationships exist in the "Relationship Matrix" and the qualitative nature of that relationship and assumption four is that the designers can appropriately identify a three number scale that captures the relationships quantitatively. Thus, an experiment would need to make factors that represented what the designers control from these two assumptions. The factors are identified as *column density*, *qualitative tendency* and *quantitative scale*.

To aid in describing the experiment set up, consider the HoQ of Fig. 4. The *column density* represents the number of CAs that a given TA effects and is calculated using Eq. (3). For example, the second TA in the HoQ of Fig. 4 has a column density of one-fifth.

The *qualitative tendency* represents the most common qualitative relationship for a given TA. Thus, if a TA has a "weak" tendency it will have one more than half of the active cells in the column designated as "weak". For the example HoQ in Fig. 4, a TA with a column density of four-fifths and a "weak" tendency will have at least two cells with a weak score inserted for calculation of technical importance. In the case of a tie in qualitative tendency (e.g., a TA column with one "weak", "medium" and "strong" relationship), there is some initial evidence that suggests using the *lowest* qualitative score to dictate the qualitative tendency for the TA column in question. However, this issue is still under study and is not a focus in the experiment discussed herein as ties are ensured to not occur.

Finally, the *quantitative scale* is the three number scale that is utilized to replace the qualitative scale for calculation of the technical importance. The experiment was performed on a five by five HoQ similar to Fig. 4.

Customer Attributes	**Technical Attributes**					**Customer Importance**
	TA1	TA2	TA3	TA4	TA5	
CA1			*		*	*
CA2			*	*	*	*
CA3		*		*		*
CA4					*	*
CA5				*	*	*
						raw score
						relative weight
						rank

Figure 4 - Experimental HoQ

$$\left.\begin{matrix}column\\ density\end{matrix}\right|_{j=1}^{n} = \frac{no.of\ CAs\ affected}{m} \qquad \textbf{(3)}$$

To perform the experiment and study the effect of each factor, only one TA was varied on all three factors. In this case, TA1 of Fig. 4 was varied on all three factors. The levels and their corresponding values for each factor are shown in Table 1. A full factorial experiment was performed (using Matlab) yielding 48 experiments (4 levels × 3 levels × 4 levels) each representing a different house configuration. The remaining TA columns were held constant in column density and are denoted by asterisks in Fig. 4. For example, TA3 has a constant column density of 2/5 in each design, thus two asterisks in the column. At each experiment, 500 simulations were performed allowing the relationship locations, score from current quantitative scale and customer weights to be randomly selected (except for TA1 where the *qualitative tendency* controlled some of the score selections). Essentially, this treated these other components as noise in the experiment.

Factors										
Column density				Qualitative tendency			Quantitative scale			
Levels										
1	2	3	4	1	2	3	1	2	3	4
Settings										
$1/5$	$2/5$	$3/5$	$4/5$	weak	med	strong	1 3 9	2 5 8	1 2 3	1 50 100

Table 1 - Factors and level settings for experiment

There is an obvious expectation for the effects of *column density* and *qualitative tendency*. Namely, any TA that affects multiple CAs, i.e. has a high *column density*, will naturally have a high relative weight and favorable rank position, since it will have more relationships with CAs than other TAs. Similarly, the more often a TA has a high *qualitative tendency*, the more likely it is to have high relative weight and improved rank position, since its quantitative scores will be higher than average. In the results of this experiment, as each of these factors increases for TA1, the relative weight should increase and rank should improve (first place is best). The primary goal then in this experiment is to study the effect of scale choice on these two importance metrics. Resulting main effects plots with mean and ninety-five percent confidence intervals for the rank and relative weight of TA1 are shown in Figs. 5-7.

Note, Figs. 5 and 6 show the results expected for *column density* and *qualitative tendency*. As the *column density* increases in Fig. 5 the mean relative weight for TA1 increases from 9% to 27% and the mean rank decreases from 4.3 to 2.2. Similarly, as the *qualitative tendency* increases in Fig. 6 the mean relative weight increases from 10% to 28% and the mean rank decreases from 4.2 to 2.2. However, based on the results in Fig. 7 there is evidence that the choice of a three number quantitative scale has no effect on the final relative weight and rank of a given TA. Effectively, the use of a three number scale pushes the importance calculations to the expected average, in this case a relative weight of twenty percent and rank of three, for a given five by five HoQ.

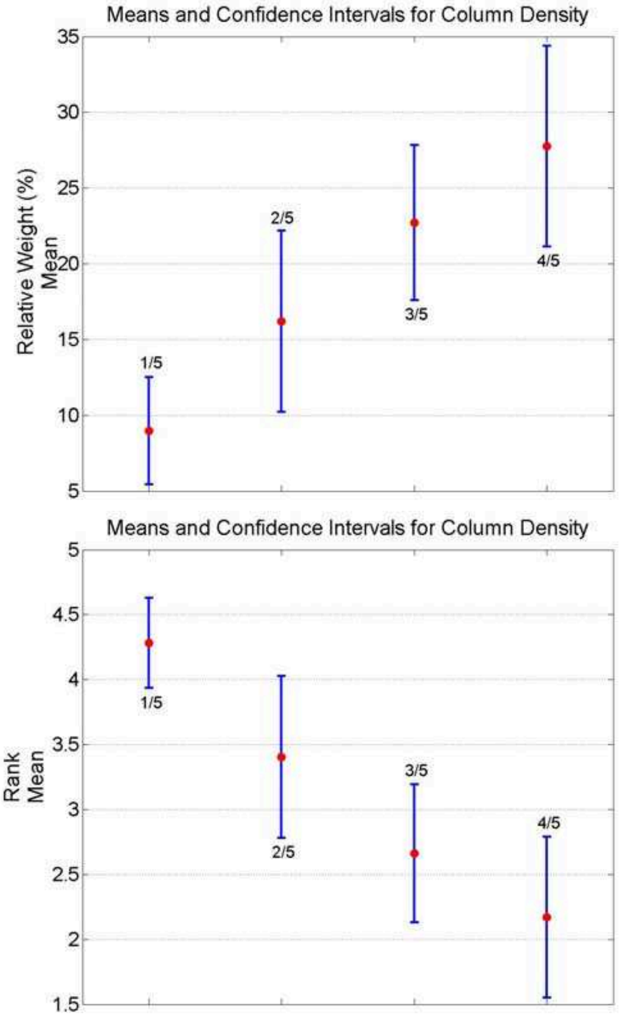


Figure 5 - Effect of column density

The purpose of using unreasonable quantitative scales such as (1-2-3) and (1-50-100) was to show that even if the range is changed there is no effect on the mean and little effect on the confidence interval. These scales are thought of as "unreasonable" because they do not represent perceptual distinctions that a choice like (1-3-9) is intended to represent. For example, the relative difference between (1, 2 and 3) is so slight, it does little to differentiate weak, medium and strong qualitative relationships that designers perceive. Similarly, (1-50-100) seems too extreme in its relative difference. While the scales typically utilized are meant to reflect expert knowledge, they are nothing more than the designers' best guess to the quantitative level of the relationship. Further, the limitations applied through simplification of the process, i.e. use of *one* three number scale that is assumed to exist on a range from one to nine and is typically dictated by practice, severely limits the extent to which conclusions may be drawn from the process.

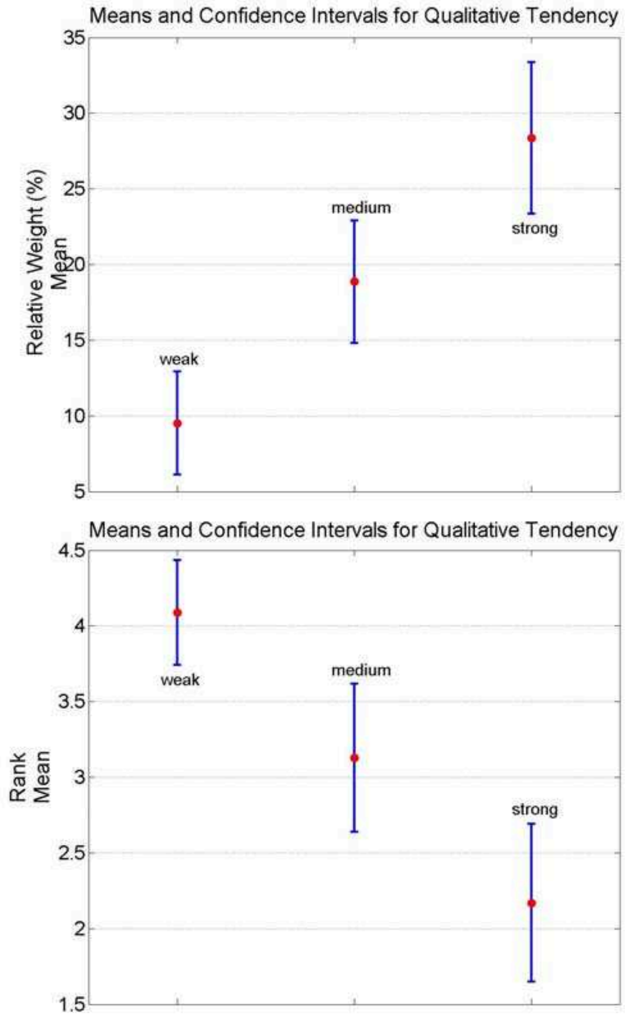


Figure 6 - Effect of qualitative tendency

In order to add statistical significance to the qualitative evidence presented thus far, comparison of the resulting relative weight and rank distributions is performed. The comparison is facilitated through a t-test performed on the resulting distributions of relative weight and rank order for each factor at each level. For example, looking at the *column density* factor there are four distributions of twelve data points (48 total experiments, 12 for each of the 4 levels). Similarly, for *qualitative tendency* there are 3 distributions with 16 data points and for *quantitative scale* there are 4 distributions with 12 data points. The t-test allows a comparison of distributions per each factor to assess differences in response (relative weight or rank order) due to the factor level. In using the t-test, it is assumed that the distributions are normally distributed and have equal variance [18]. The null hypothesis for the t-tests performed is that the distributions have equal means, or H_o:$\mu_1=\mu_2$, where μ_1 and μ_2 represent the means of the two distributions in question. In other words, the null hypothesis is that there is no effect due to changing the factor levels. The t-tests were performed at a significance level of α=0.05. From the t-test, a P statistic is calculated. If the value of P is greater than α, the null hypothesis is rejected (i.e., there is a difference in the distributions due to the changing levels). If the value of P is less than α we fail to reject H_o (there is no difference in the distributions due to the changing levels).

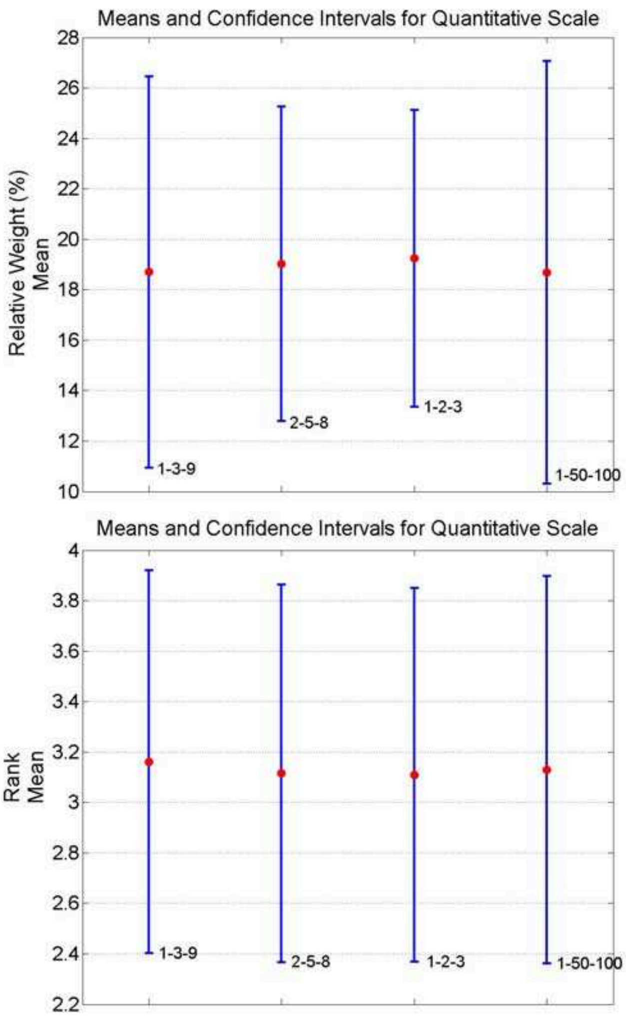


Figure 7 - Effect of quantitative scale

The results of performing the t-tests are shown in Table 2. Note that a test of equality for the variances of the distributions for each factor level was performed and it was found that the assumption of equal variance is valid. For both *column density* and *qualitative tendency* the value of P is almost always less than α, indicating a rejection of the null hypothesis, H_o. This provides evidence that changing the levels for these two factors indeed affects the final relative weight and rank order of TAs in the HoQ. There are only two cases in which the value of P is larger than α for *column density* level comparisons. The first case occurs when comparing 2/5 and 3/5. However, since the P value is only slightly greater than α, it suggests that there is some difference in the distributions, i.e. there is an effect due to changing from 2/5 to 3/5. Only in the case of changing *column density* from 3/5 to 4/5 is there clear statistical proof that there is little or no effect on the final relative weight and rank of TA1. This is understandable since a change from 3/5 to 4/5 only represents a 33% increase in *column density*, the number of CAs affected for the TA already exceeds half, thus reducing its relative effect.

For the *quantitative scale* factor however, the value of P is much greater than α in every case, giving statistical creditability for accepting H_o. Namely, it can be concluded that there is no affect on the final quantitative results in the HoQ due to quantitative scale choice.

Results of t-test for 5x5 HoQ experiment

Factor	Levels compared	Relative weight P value	Rank P value
Column density	1/5 vs. 2/5 1/5 vs. 3/5 1/5 vs. 4/5 2/5 vs. 3/5 2/5 vs. 4/5 3/5 vs. 4/5	0.038 0.000 0.000 0.078 0.010 0.220	0.014 0.000 0.000 0.061 0.006 0.200
Qualitative tendency	weak vs. medium weak vs. strong medium vs. strong	0.000 0.000 0.003	0.002 0.000 0.008
Quantitative scale	1-3-9 vs. 2-5-8 1-3-9 vs. 1-2-3 1-3-9 vs. 1-50-100 2-5-8 vs. 1-2-3 2-5-8 vs. 1-50-100 1-2-3 vs. 1-50-100	0.930 0.909 0.984 0.979 0.915 0.895	0.927 0.917 0.950 0.991 0.977 0.968

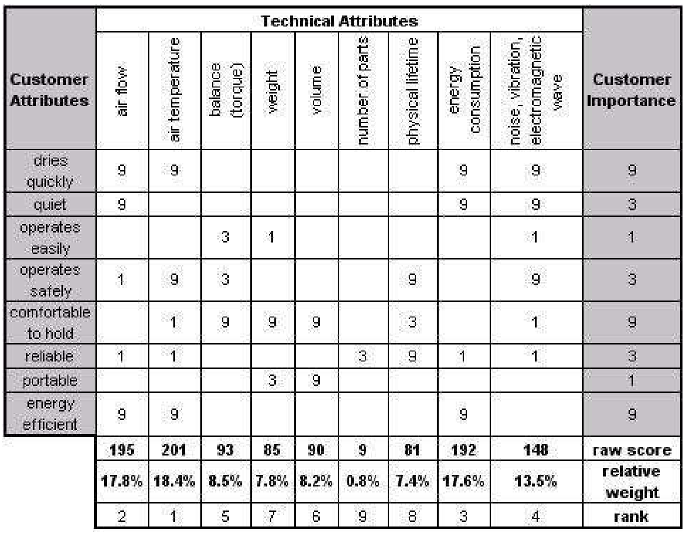
Table 2 - Statistical significance of factors

Given these conceptual limitations and the evidence provided by the experiment, both qualitative and quantitative, it is likely that resulting relative weight calculations should not be utilized for any type of quantitative comparison. For example, designers should not utilize the relative weights as a reflection of relative importance of one TA over another. At best, the results from the experiment suggest that it may be possible to get a sense of the rank importance of one TA over another, since it is evident that the *column density* and *qualitative tendency* of a TA seem to have dominating effect. However, there is still a danger in this as is discussed and shown in the empirical studies of the next section.

5. EMPIRICAL INVESTIGATION OF THE HOQ

Given the doubts raised in the assumptions of the HoQ method and the results of the experiment of Section 4, exploring an example empirically would be beneficial. To conduct this empirical investigation an example HoQ from the literature is utilized. The example is shown in Fig. 8. The example in Fig. 8 is a HoQ for the design of a hair dryer adapted from an example in [19]. The example represents an instance of potential design decision making using the HoQ. The goal is to show how erroneous conclusions and decisions could be made regarding this product example due to the assumption of quantitative relationship scale.

Consider how conclusions might be drawn from a given HoQ. Step 11 of the procedure for utilizing the HoQ suggests utilizing the results from step 7, i.e. to look at the raw score (rank) or the relative weight calculated for each TA. The raw score, rank and relative weight are given for each TA in the example of Fig. 8. Designers must now draw conclusions based either on the raw score (ranked priority) or the relative weight, as per step 11. The choice between using rank to prioritize and using relative weights to make decisions provides several possible courses of action for designers. It is likely that every company that utilizes the HoQ has different approaches for handling this information, which may even change for each new design. To support the empirical investigation here, two possible approaches are used.



Customer Attributes	air flow	air temperature	balance (torque)	weight	volume	number of parts	physical lifetime	energy consumption	noise, vibration, electromagnetic wave	Customer Importance
	Technical Attributes									
dries quickly	9	9						9	9	9
quiet	9							9	9	3
operates easily			3	1					1	1
operates safely	1	9	3				9		9	3
comfortable to hold		1	9	9	9		3		1	9
reliable	1	1				3	9	1	1	3
portable				3	9					1
energy efficient	9	9						9		9
	195	**201**	**93**	**85**	**90**	**9**	**81**	**192**	**148**	**raw score**
	17.8%	**18.4%**	**8.5%**	**7.8%**	**8.2%**	**0.8%**	**7.4%**	**17.6%**	**13.5%**	**relative weight**
	2	1	5	7	6	9	8	3	4	**rank**

Figure 8 - Hair dryer HoQ example

The first option is that the designers could utilize the relative weights to determine how resources should be allocated in the course of the design. Specifically, the designers could allow the relative weights to dictate the percentage of resources to spend in designing around each TA. The difficulty here however is that the designers do not truly know if the range and relative difference in relationship scale is representative of the actual relationship between CAs and TAs, as evidenced by Fig. 7. Thus, the relative weights could be potentially no better than those generated by some random process[2]. To investigate this idea, random processes that work within the framework of the HoQ were designed and used to

---

[2] This notion was put forth in open discussion at the Decision-Based Design Open Workshop held at the 2002 DETC.

"simulate" results. Three different random processes were compared to the results in the example of Fig. 8. The empirical results were generated as follows:

1. ***Insertion of discrete uniform random number***: In this recreation method, random numbers from a discrete uniform distribution (range 1 to 9) are inserted wherever a relationship exists in the original HoQ relationship matrix. The relative weight of each TA is calculated for each of the thousand recreated HoQs and the average relative weight for each TA over all recreations calculated. The goal of this simulation is to observe if using random numbers with known relationship locations yields results similar to the original HoQ.

2. ***Arbitrary insertion of a three number scoring scale***: In this recreation method, a three number scale consistent with the example is used (1-3-9). However, a score (zero, 1, 3 or 9) is arbitrarily inserted in the relationship matrix locations. The controlling factor is the "column density" metric, which is calculated from the original HoQ for each TA, using Eq. (3). Using a discrete uniform random number generator and moving down each column of the relationship matrix for each TA, a random number [0,1] is generated and if it is greater than the column density for that TA a zero is inserted. Otherwise, a relationship is assumed to exist and another random number [0,1] is generated. If the number is less than one-third a low score is inserted, if the number is greater than or equal to one-third but less than two-thirds a medium score is inserted, and if the number is greater than two-thirds a high score is inserted. Once this procedure is completed for every position in the relationship matrix, the relative weight for each TA is calculated for each HoQ recreation and average relative weight for each TA over the total number of recreations is calculated. The goal of this simulation is to observe whether reducing the certainty of where relationships exist and the quantitative level of that relationship yields similar results to the original HoQ.

3. ***Arbitrary insertion of a discrete uniform random number***: Similar, to the previous recreation method this recreation procedure also uses the column density to control the number of relationship scores inserted for each TA. However, when a relationship is assumed to exist in this case (i.e., if a uniform random number [0,1] is less than or equal to the column density), a discrete random number from a uniform distribution is inserted (range 1 to 9). Again, the relative weight of each TA for each HoQ recreation is calculated and the average relative weight for each TA over all recreations calculated. The goal of this simulation is similar to the previous approach but the certainty of the quantitative level of the relationship has been reduced further.

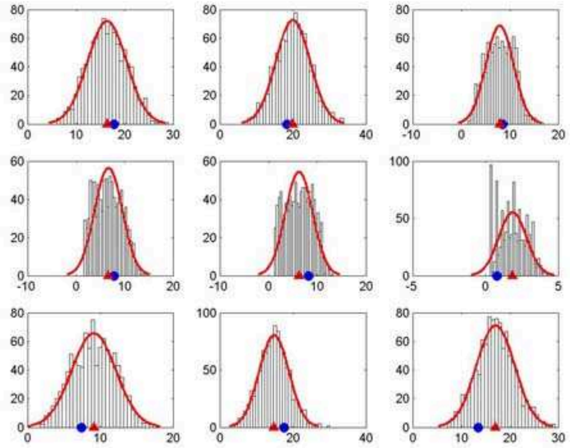


Figure 9 - Random HoQ 1 results (hair dryer)

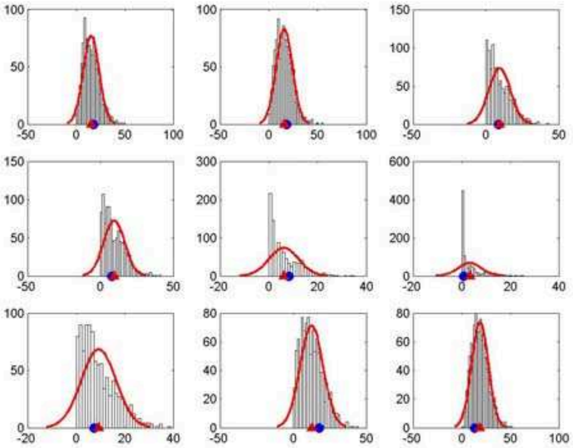


Figure 10 - Random HoQ 2 results (hair dryer)

The distribution of relative weights that result from each of these three random procedures for the hair dryer example of Fig. 8 are shown in Figs. 9, 10 and 11, where the circle represents the actual relative weight from the original HoQ and the triangle represents the average of the distribution. Note that the TAs listed left to right in the hair dryer HoQ appear left to right and row by row in the figures. The averages of each procedure are shown in Table 3 for the hair dryer example.

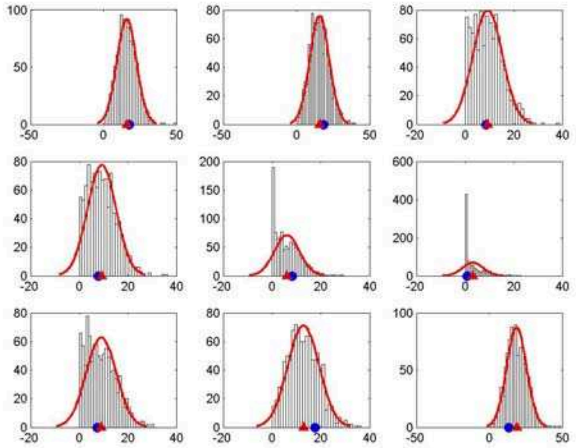


Figure 11 - Random HoQ 3 results (hair dryer)

Average relative weight	air flow	air temperature	balance (torque)	weight	volume	number of parts	physical lifetime	energy consumption	noise, vibration, electromagnetic wave
original HoQ	17.8	18.4	8.5	7.8	8.2	0.8	7.4	17.6	13.5
Approach 1	16.4	20.0	7.9	6.7	6.2	1.9	9.1	14.8	17.0
Approach 2	15.5	15.8	9.3	9.4	6.1	3.3	9.2	12.6	18.8
Approach 3	15.9	15.7	9.3	9.3	6.1	3.2	9.1	12.9	18.6
column density	0.63	0.63	0.375	0.38	0.3	0.13	0.38	0.38	0.75

Table 3- Random HoQ results (hair dryer)

Numerically, the average relative weights generated using the random procedures appear similar to the relative weights from the original hair dryer HoQ. The Wilcoxon signed rank test, which can be used to test whether the median of a distribution is equal to a scalar value [20], gave no verification that the distribution means were the same as the actual relative weights (scalar values) in the original HoQs. So, while a random number generator does not behave exactly as the HoQ method, the numerical proximity of results is hard to ignore.

The authors realize the simulations are not purely random. That is, maintaining relationship locations and using the "column density" metric in the simulations provides a form of bias. However, this bias is representative of assumption three, i.e., designers' can indeed recognize where relationships exist. The fact that the resulting average relative weights via random score insertion never vary widely from the original HoQ suggests that the quantitative information provided by the scales lacks meaning. Thus, the true relationship, i.e., relative difference, among TAs is not well defined. The fact that random number generation has produced results numerically similar to the scales typically used in HoQ implies that these scales are not necessarily meaningful representations of the relationship between customer and technical attributes.

A second course of action the designers could follow is to use the rank order of the result from the HoQ to prioritize some or all of the TAs. Thus, allocation of resources would be left to designer discretion based upon the rank order. However, another difficulty with the lack of certainty in the relationship scale is the uncertainty that can lie about a particular relationship score. For example, the designers likely use "3" to represent a "medium" qualitative relationship; however, there is no reason to believe that the actual value is represented. Thus, the question becomes what affect does uncertainty in the qualitative relationships have on the final rank order of TAs. That is, maybe "3" is utilized to represent the value but in reality the exact value could be slightly less than or greater than "3". Add to this the complication that designers should not assume (1-3-9) is necessarily better than another scale, like (2-5-8), and the result is uncertainty in the true final rank. To investigate the effect of quantitative relational uncertainty, another simple simulation was performed using triangular distributions to represent uncertainty in a particular three number scale. The triangular distributions used and the resulting rank shifts for the hair dryer TAs are shown in Figs. 12 and 13.

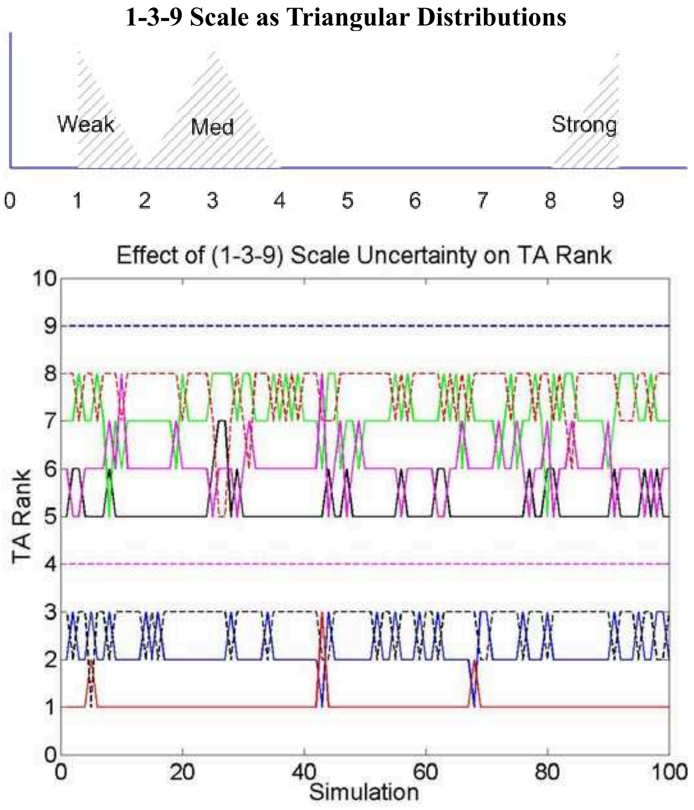
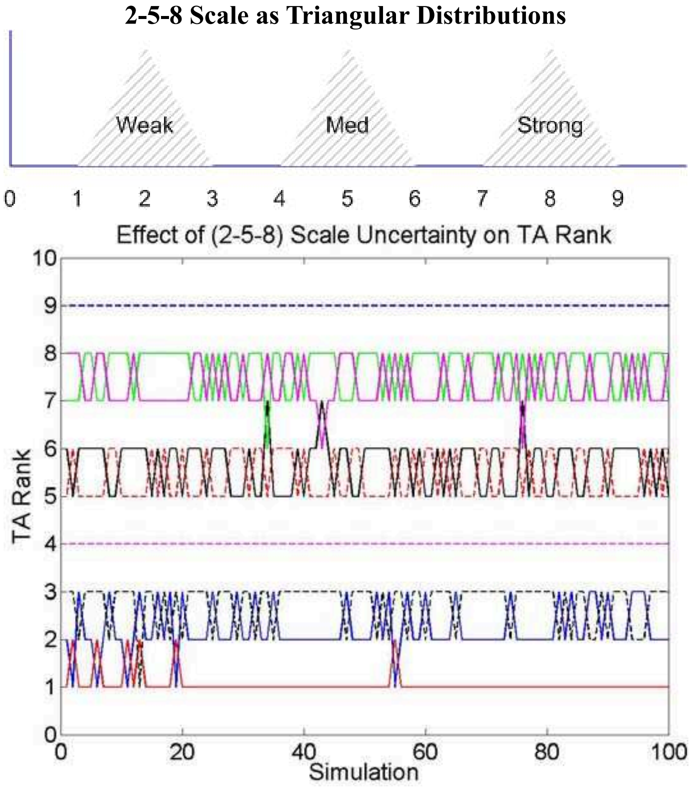


Figure 12 - Rank shifts for hair dryer (1-3-9)

The idea behind using the triangular distributions shown in Figs. 12 and 13 is to represent the uncertainty that exists in the actual quantitative scale choice and the fuzzy nature in a given qualitative relationship. The charts below the distributions show how the nine TAs of the hair dryer example in Fig. 8 can change rank position over one hundred recreations. The recreations are intended to represent one hundred individual design situations where the true quantitative value of "weak", "medium" and "strong" may not be (1-3-9) or (2-5-8) exactly. The result is that the **true** rank order is different than what is achieved in assuming the whole number scales. The results are hypothetical but serve to show the sensitivity of rank order to various forms of uncertainty in the actual quantitative relationships. That uncertainty might be the choice in scale (1-3-9 vs. 2-5-8) or the actual quantitative strength of a particular qualitative relation (distribution) or both.


2-5-8 Scale as Triangular Distributions

Effect of (2-5-8) Scale Uncertainty on TA Rank

Figure 13 - Rank shifts for hair dryer (2-5-8)

Looking at various scales and uncertainty quantifications provides a glimpse into the sensitivity of TA rank order with respect to the scale choice and the amount and form of uncertainty between qualitative relations. The TAs that shift only a few times and/or only one place in the rank order are not necessarily of concern and seem relatively robust to the uncertainty depicted in the scales here. However, note that at times some TAs can have large rank order changes. In Fig. 12 one TA shifts rank between fifth and eighth place and sixth and eight place in Fig. 13 as a result of a shift in quantitative scale and the amount and shape of the relationship uncertainty. These drastic changes that can result from the slightest uncertainty in the quantitative scales and the lack of knowing the true quantitative scale should worry designers as the robustness of the final rank order comes into question. This is especially true when designers are looking for a subset of TAs to focus on. Fig. 12 shows that it may be difficult to know which TAs are the top five in importance since several TAs shift among the fifth through eighth rank order positions.

From these investigations, it is clear that however the designers choose to influence design decisions, i.e., from relative weights or from rank order, confidence in those results should not be high. More commentary on these results and their impact is discussed in the Conclusion section.

6. CONCLUSION

"The principle benefit of the house of quality is quality-in-house. It gets people thinking in the right directions and thinking together. For most U.S. companies, this alone amounts to a quiet revolution" [1]. This conclusion from Hauser speaks to the benefit of conceptual mapping alluded to in the Introduction. For companies just implementing QFD and the HoQ, there is undoubtedly an improvement in information structure, flow and direction. However, "thinking in the right directions" is only a qualitative notion. As design processes and methodologies improve and companies become more efficient in their use of information and knowledge, qualitative direction alone will yield fewer and fewer gains. While the HoQ is viewed as a tool that brings the proper entities into communication within a company, the work in this paper highlights the fact that the current methodology is limited to qualitative assessment at best.

Any quantitative conclusions drawn from the method are potentially flawed for reasons discussed in this paper. Specifically, the experiment shows that one choice of quantitative scale over another has no effect on the final outcome in terms of rank and relative weight. This implies that quantitative conclusions are likely flawed since quantitative importance calculations rely on a scale choice. The empirical investigation using random processes adds credence to this conclusion and gives the impression that the scales lack any meaningful information in representing the relationships between customer and technical attributes.

The empirical results generated from the representation of quantitative scales as uncertain triangular distributions shows the sensitivity of rank order to scale choice and uncertainty in the fuzzy qualitative value they represent. While a rank order may be seen as a hybrid compromise between quantitative and qualitative conclusions about the TA importances, there is still obviously room for error that could lead to disastrous decisions. Further adding to this difficulty is the fact that designers may miss the existence of weak relationships and this affects the *column density* factor, which was shown to be extremely important to TA importance in the experiment.

In all, the results of this paper show that designers should limit the importance placed on results from the HoQ method, especially those regarding quantitative value. Of course, the authors realize that the limitations laid out in this paper are likely known to varying degree by designers who utilize the HoQ regularly. However, it is important that these limitations are studied and reported in a rigorous fashion to ensure they represent global rather than local knowledge.

The results of this paper provide motivation for improving upon the conceptual soundness of the HoQ method to make it a more rigorous tool for supporting design. Improvement would support the level of confidence designers have in the conclusions drawn from the HoQ methodology and help avoid conflict during the design process. That is, if multiple designers are in disagreement over the conclusions (e.g. most important TAs) because they are aware of the limitations described in this paper, resolving the conflict becomes difficult. In such cases, while the HoQ provides some support to resolving the conflict, it falls short of ending the conflict definitively. In the authors' view, the HoQ method has the potential to overcome these limitations and go beyond conceptual mapping. In order to continue achievement through this methodology a more thorough quantitative mechanism must be identified. Concurrent work in this vein is ongoing in terms of finding a design of experiment based approach to identifying the relationship values between CAs and TAs through application of conjoint analysis [21, 22]. Further work is also planned in expanding the HoQ experiment to include factors such as number of TAs.

ACKNOWLEDGMENTS

We gratefully acknowledge the National Science Foundation, grants DMI-9875706 and DMI-0140411, in support of this research.

REFERENCES

[1] Hauser, J. and Clausing, D., 1988, "The House of Quality," Harvard Business Review, **66**, (3), pp. 63-74.

[2] Hauser, J., 1993, "How Puritan-Bennet Used the House of Quality," Sloan Management Review, Spring 1993.

[3] *Abstracts from Symposia on QFD*. 2005, accessed: February 2005, http://www.qfdi.org/books/symposium_%20proceedings.htm.

[4] Kaldate, A., Thurston, D., Emamipour, H., and Rood, M., 2003, "Decision Matrix Reduction in Preliminary Design," *ASME Design Engineering Technical Conferences*, Chicago, IL, DETC2003/DTM-48665.

[5] Olewnik, A., Hammill, M., and Lewis, K., 2004, "Education and Implementation of an Approach for New Product Design: An Industry-University Collaboration," *ASME Design Engineering Technical Conferences*, Salt Lake City, UT,

[6] Barzilai, J., 1997, "A New Methodology for Dealing with Conflicting Engineering Design Criteria," *18th Annual Meeting of the American Society for Engineering Management*.

[7] Saari, D., 2000, "Mathematical Structure of Voting Paradoxes. I; Pair-wise Vote. II; Positional Voting," Economic Theory, **15**, pp. 1-103.

[8] Hazelrigg, G., 2003, "Validation of Engineering Design Alternative Selection Methods," Journal of Engineering Optimization, **35**, (2), pp. 103-120.

[9] Olewnik, A. and Lewis, K., 2005, "On Validating Engineering Design Decision Support Tools," Journal of Concurrent Engineering Design Research and Applications, to appear.

[10] Dym, C. and McAdams, D., 2004, "Modeling and Information in the Design Process," *ASME Design Engineering Technical Conferences*, Salt Lake City, UT, DETC2004-57101.

[11] Hazelrigg, G., 2003, "Thoughts on Model Validation for Engineering Design," *ASME Design Engineering Technical Conferences*, Chicago, IL, DETC2003/DTM-48632.

[12] Pedersen, K., Emblemsvag, J., Bailey, R., Allen, J.K., and Mistree, F., 2000, "Validating Design Methods and Research: the Validation Square," *ASME Design Engineering Technical Conferences*, Baltimore, Maryland, DETC2000/DTM-14579.

[13] Malak, R.J. and Paredis, C.J.J., 2004, "On Characterizing and Assessing the Validity of Behavioral Models and Their Predictions," *ASME Design Engineering Technical Conferences*, Salt Lake City, UT, DETC2004-57452.

[14] Frey, D. and Li, X., 2004, "Validating Robust Parameter Design Methods," *ASME Design Engineering Technical Conferences*, Salt Lake City, UT, DETC2004-57518.

[15] Breyfogle, F.W., 1999, *Implementing Six Sigma: Smarter Solutions Using Statistical Methods,* John Wiley & Sons, Inc. New York.

[16] Hofmeister, K., 1995, "QFD in the Service Environment," in *Quality Up, Costs Down: A Manager's Guide to Taguchi Methods and QFD*, W. Eureka and N. Ryan, Editors, pp. 57-78.

[17] *The American Heritage Dictionary of the English Language,* 2000, 4th ed., Houghton Mifflin Company.

[18] Montgomery, D., 2001, *Design and Analysis of Experiments,* 5th ed., John Wiley & Sons, Inc. New York.

[19] Masui, K., Sakao, T., Aizawa, S., and Inaba, A., 2002, "Quality Function Deployment for Environment (QFDE) to Support Design for Environment (DFE)," *ASME Design Engineering Technical Conferences*, Montreal, Canada, DETC2002/DFM-34199.

[20] Gibbons, J.D., 1985, *Nonparametric Methods for Quantitative Analysis,* 2nd ed., American Sciences Press, Inc. Columbus, OH.

[21] Dolan, R., 1990, "Conjoint Analysis: A Manager's Guide," Harvard Business School Publishing, Boston, MA.

[22] Green, P. and Srinivasan, V., 1978, "Conjoint Analysis in Consumer Research: Issue and Outlook," Journal of Consumer Research, **5**, pp. 103-123.