Collegiate Aviation Review International

Volume 37 | Issue 1 Peer-Reviewed Article #4

1-1-2019

Choosing a Collegiate Training Aircraft with Confidence

Samuel M. Vance
Oklahoma State University

Qualitative preference is frequently used to make significant, programmatic choices between competing suppliers and products. Translating qualitative choice into defensible quantitative representation is possible with patience, method and, care. One example of this translation is the application of a popular Total Quality Management (TQM) tool, known as Quality Function Deployment (QFD), in the choice of a new/replacement, collegiate-training aircraft. Using QFD is especially important when a fleet replacement is being considered as the cost of a new, current aircraft can easily approach $500,000; thus a significant fiscal commitment is incurred in replacing a multi-aircraft fleet. A choice of this magnitude deserves multiple stakeholder inputs and requires respect from differing viewpoints. The successful outcome of any decision process ultimately hinges upon confidently exercising the best choice. The decision tool needs to be transparent, easy-to-understand and easy-to-apply. The corresponding choice of a preference scale can either mask or illuminate driving criteria in the decision process. This paper explores the application of QFD to the decision process across competing training aircraft choices and offers justification of the QFD non-linear "0, 1, 3, 9" preference scale. Application research into the mechanics of human preference showed that if 95% reliability in choice between alternatives is desired, then the perceived difference between the choices needs to be a factor of 3.0, as is the case in the employed QFD scale. Selection criteria used in the training aircraft decision, their dissimilar weighting, and the evaluation of competing aircraft in a recent collegiate-training aircraft selection are displayed as exemplars.

Recommended Citation:
Vance, S.M. (2019). Choosing a Collegiate Training Aircraft with Confidence. *Collegiate Aviation Review International,* 37(1), 59-72. Retrieved from http://ojs.library.okstate.edu/osu/index.php/CARI/article/view/7763/7191

Collegiate-training aircraft are consumable capital assets which require periodic replenishment. A complete fleet replacement is decision-intense and generates more attention and fiscal stress than a recurring investment which is less fiscally stressful but no less decision-intense.

The purpose of this paper is two-fold 1) to defend the illustrated scoring scale as necessary and proper for such an endeavor, and 2) to overview and recommend a repeatable methodology when choosing expensive, capital assets such as collegiate-training aircraft.

The two research questions probed in this paper both address this type of difficult, capital-intense decision:

- Research Question 1: How small a difference (in a criteria or attribute) can human preference reliably perceive?
- Research Question 2: Can a collegiate-training aircraft (specific make/model) be methodically chosen with confidence?

Decisions, even significant ones, are frequently are more of an art form than scientific. Often the decision logic is difficult to reconstruct, if it can even be reconstructed at all. It is a professionally respectful investment of the decision maker's time and finances if a repeatable process is used to organize the decision process. The U.S. aerospace industry Total Quality Management (TQM)/ISO 9000 revolution in in the late 1980s/early 1990s produced a bevy of tools. Stuart Pugh's Concept Selection process was a popular and simple tool to use to choose among competing alternatives (Burge, 2009). Numerous, Multi-Attribute Utility (MAU) tools also manifested themselves in this time frame, with two of the higher profile tools being The Analytic Hierarchy Process and Quality Function Deployment (QFD). All of these tools, which were once expensive software commercial purchases, are now freeware on web sites such as http://www.mindtools.com.

A typical flow in a decision process, as discussed in Hauser & Clausing (1988), Huber (1974), and Von Winterfeldt & Fischer (1975) follows these steps:

- A clear statement of the decision to be made
- Criteria with which to judge value – and the criteria's associated definitions
- User-defined preference weights on the criteria, especially if they are value judged as unequal
- Collecting the options from which to (ultimately) choose
- A method to evaluate the options v. criteria – and the associated scoring definitions
- Production of plausible results
- Ability to credibly explain the method used and the conclusion(s)/choice(s) drawn

QFD is a popular Japanese Total Quality Management (TQM) tool that migrated into the U.S. auto and aerospace industry in the late 1980s/early 1990s. Hauser & Clausing (1998) popularized the tool in a seminal Harvard Business Review article. To understand more about QFD's roots and the tool's mechanical, quantitative, decision-scale grounding, see Vance, Carstens, Gasper & Parker (2018).

Decisions should proceed from a requirements perspective. Asking "WHAT" is wanted/needed first is more important than examining "HOW" to satisfy the need; even though the latter is the more typical starting place. Focusing too early on solutions, especially in complex decisions, is almost always a false economy as the resulting choice(s) quickly become indefensible. "WHATs", can be categorized as the requirements or design criteria and must include their definitions and their priority. These properly precede "HOWs" which are represented by ideas, solutions, options, and concepts. While the "HOW" solution space also needs definitions, detailed solutions should not be assembled before an understanding of the decision requirements/criteria are documented.

"WHATs" are the harder of the two steps to take and, without process, will likely be avoided until they simply must be addressed for clarity. "HOWs" are the fun step and, without process, are all that likely may be attempted. Most decisions have multiple attributes for consideration, and most frequently these attributes are valued unequally by the evaluators chartered with the decision. In this type of decision, a MAU tool is needed and provides a proper balance between process and velocity. QFD employs a systematic identification of "WHATs", weights these requirements, and then mathematically evaluates each "HOW" against the "WHATs" to determine preference.

There are two critical, differentiating characteristics in the use of any decision tool 1) the ease of interpretation and credibility of the employed scoring scale, and 2) the level of documentation detail offered to support the decision, after the decision process is completed.

The first area of decision process differentiation is in the employed scoring scale. In the case of QFD, the 0-1-3-9 scale is purposefully non-linear and increments by a factor of three – why? This paper will overview the rationale for the scoring scale in an attempt to build confidence in its purposefulness, then apply the scoring scale to an exemplar application selecting a collegiate fleet training aircraft. Examples of QFD decision criteria (the "WHATs) in choosing a collegiate-training aircraft (the "HOWs") application include the macro balance between the aircraft's purchase price, characteristics and the aircraft's fit with the institution's current fleet, mission, vision and ethos.

The second area of decision process differentiation is in the care in which definitions are written. The value of definitions cannot be overstated. When employing QFD, a more complete application will offer definitions for the "WHATs", the "HOWs" and the scoring scale graduations applied to each "WHAT". While this may seem oppressive, the investment in time actually makes the scoring process easier and more efficient; and, more importantly, once the stakeholders have completed the decision process, the results are easier to reconstruct and defend. It is this ultimate defensibility that will distinguish a superior decision process application from a mediocre, or poor, application.

Literature Review

Multiple references concerning the history of QFD and using/optimizing QFD were easily located, starting with the QFD article by Hauser & Clausing (1988); however, no source was located that indicated any specific, QFD-scale basis or the rationale associated with the specific selection of QFD's nonlinear, 0-1-3-9 scale. An informative article tracing the history of QFD from 1978 to the present (Akao & Mazur, 2003) did not illuminate the scale basis. Burke, Kloeber, & Deckro (2002); Delano, Parnell, Smith, & Vance (2000); Fiorenzo & Alessandro, (1999); Franceschini & Rossetto (1995); and Franceschini & Rupil (1999) each produced excellent advice for employing and optimizing QFD, but none revealed nor discussed the basis for the preference scale.

When no basis was located, the literature review moved to the broader subject of human preference in which the concept of "Just Noticeable Difference" (JND) and Weber's Law was illuminated (Britt & Nelson, 1976). Weber's (and his mentor Fechner's) mathematical model, originally published by Weber in 1860, quantifies the perceived changes in the strength of a stimulus to any of the five senses (Britt & Nelson, 1976) as shown in equations (1) and (2).

(1) k = (perceived change [$p\Delta$]) = (change in stimulus [ΔS]) / (initial stimulus [S])

(2) $k = p\Delta = \Delta S/S = JND/S$

For example, if the weight of an article, which = 1, must be increased 33% before the difference is detectable, then the initial stimulus = 1 and the JND = 0.33. The k-value can be expressed as in equation (3):

(3) $k = .33/1 = .33$

Likewise, if the weight of the article must be increased 100% (doubled) before the difference is detectable, then the $k = 1$; correspondingly, if the weight must be increased 200% (tripled) to be detectable, then the $k = 2$.

Weber's Law is frequently quoted in consumer marketing with a reliability factor of 50% (Britt & Nelson, 1976), which seems inordinately low for consistent, dependable results in applications beyond marketing. Reliability of 50% would be unacceptable in aerospace engineering/program development applications where a 5% error (95% reliability) is typically considered the bare minimum of acceptance. How confident can a/should a decision maker be in their choice? To distinguish preference with confidence and a high reliability, does the QFD scoring scale need to be a factor of three (0-1-3-9)? Does the QFD scale of 0-1-3-9 achieve a

95% reliability in choice? Vance, Carstens, Gasper & Parker (2018) explored this exact question.

Preference Scale Investigation Methodology

Two, primary, data-collection methods were employed, each designed to contrast the other: an in-person, convenience sample taken on the Oklahoma State University (OSU) campus with physical weights taken in the spring of 2017, and an online survey of random respondents based on pair-wise, visual-size differences of circle areas taken in the fall of 2017. Both of these approaches were designed to elicit a preference response choice when presented with different stimuli. The on-line survey method drew from two sources: Amazon M-Turk and OSU students in the College of Education, Health & Aviation's online research solicitation system – called SONA (https://sona-systems.com).

The weight-based approach employed a structured, research script which contained 23, graduated-in-weight, pair-wise comparisons of ¼-20 plated nuts secured in small, uniform, cardboard boxes. Other ballast possibilities originally included the consideration of commercially identical shapes, such as marbles or batteries, and measured commodities such as clay, sand, or stones. Due primarily to cost considerations, ¼-20 plated nuts were chosen since over 650 were needed. Respondents were conveniently sampled on campus and randomly presented with five pairs of boxes and then queried, *"Can you detect a difference in weight? If so, please describe the difference? You can respond any way you wish such as 'Heavier', 'Lighter', or 'Equal'".* Figure 1 shows an exemplar box containing 10, ¼-20 plated nuts. Respondents were offered as much time as they desired to make the five comparisons; rarely did a single, pair assessment take more than three-five seconds.

Figure 1. One of the 23 boxes prepared for the research on top of a stock, new, unfolded box (note the small "A" inscription on the top edge).   This alphabetical code, which only faced the researchers, indicated the box contained 10, ¼-20 nuts.

The on-line survey was designed to translate the same 23 weight pairs into an electronic-friendly form. Consistent with Krider, Raghubir, & Krishna (2001), area was selected as the visual metric to be scaled as it proportionally represented total weight more accurately than either a proportional increase in diameter or circumference. Figure 2 shows an exemplar screen from the Qualtrics® elicitation software. Respondents were offered each pair-wise view for three seconds before the software automatically advanced to the collection screen for the respective sample. Total survey time to assess the 23 circle pairs was typically less than three minutes.

Figure 2. An online screen shot presentation of the 40-to-44, ¼-20 nut pair translated (normalized) to area. Note there is a counter displayed in the upper left corner which has incremented to 2 seconds and the orange slider bar at the top of the screen is showing survey completion progress.

Findings & Discussion

The employed-survey tools, number of respondents, and number of collected pair-wise comparisons from the respondents were tallied as follows:

- OSU in-person survey on campus (n = 127 respondents [not paid], 633 comparisons)
- Amazon's M-Turk (n = 524 respondents [who were paid $0.35 each], 12,048 comparisons)
- OSU SONA (n = 210 respondents [not paid], 4,841 comparisons*)*
- Totals: N = 861: with 17,522 comparisons

Preference Scale Investigation Results

Figure 3 plots the mean difference in either the weight or area pair for a k = 1, 2 or 3 versus the percent correct that the respondents were able to obtain in each pair-wise comparison. The percent correct is a direct indicator of reliability of choice.

Each dot in Figure 3 represents one of the pair-wise comparisons in either the in-person, weight-based survey [blue], area in the M-Turk survey [orange], or area in the SONA survey [grey]. Recall the circle areas were normalized to follow the exact mathematical relationships scheduled for the ¼-20 nuts. The horizontal axis was plotted to show the results by weight (or area). For example, in the case of the 27-81 pair (k = 2.0), the mean of this pair is computed as (27 + 81)/2 = 108/2 = 54. Data for the 27-81 pair was plotted on the horizontal axis at 54, not 27 or 81.

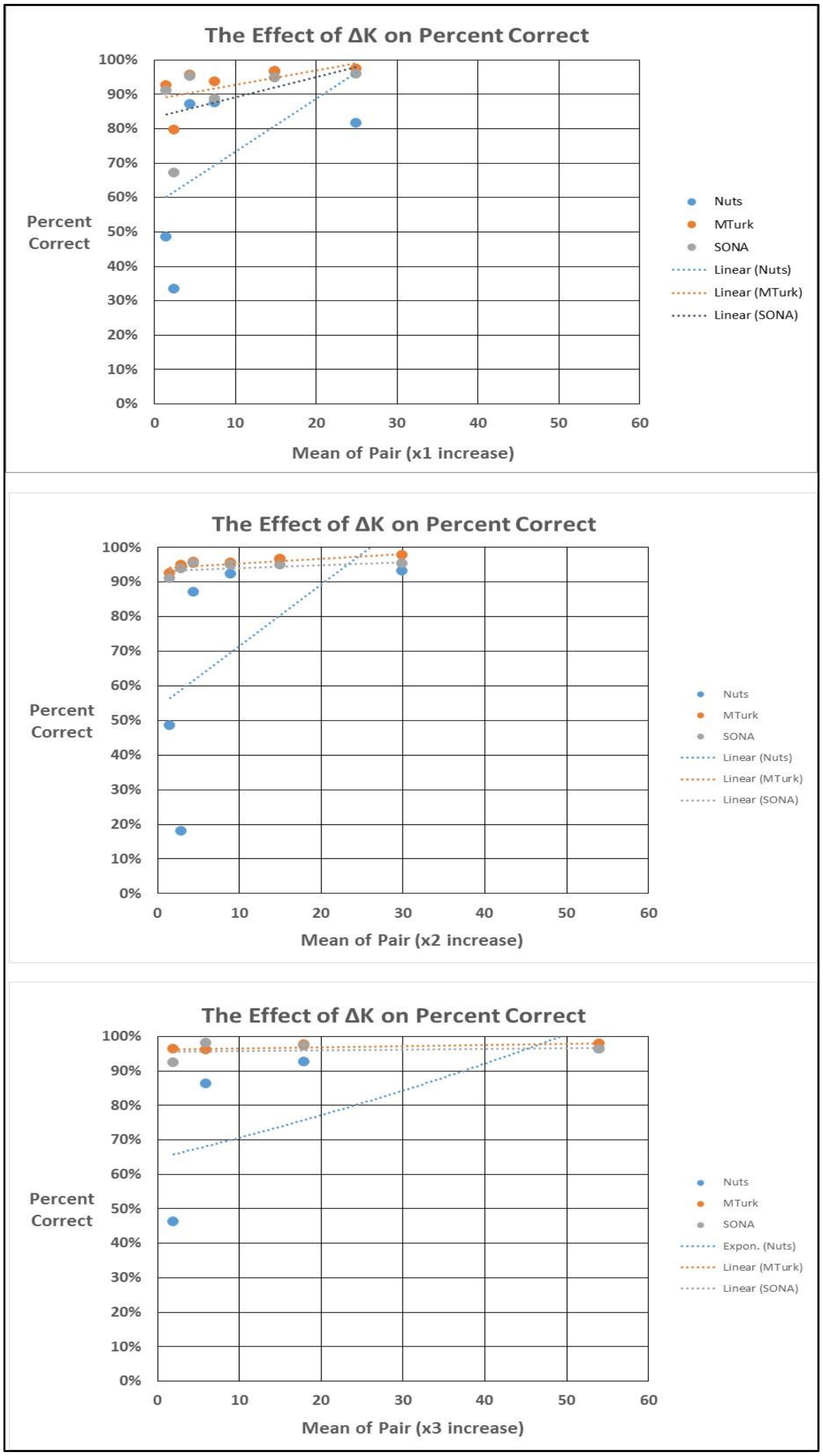

Figure 3. These three graphs show mean weight/area on the x-axis versus percent correct on the y-axis. From top-to-bottom, the graphs show a 1x, 2x, and 3x factor in the mean weight/area increases. All three data streams are shown in respective color codes.

Predictably, when the weight/area was increased at a larger rate (3x versus 2x, or 1x), the reliability (summarized by the least squares/best fit lines) increased quickly to more than 90% and, in many cases for the online survey data, more than 95% even at lower weight differences. These graphs document the area-based results produced superior reliability over the weight-based results. All of the OSU in-person survey results [blue] scoring less than 50% reliability were for small weight differences (1 or 2, ¼-20 Nuts).

Significantly, only when $k = 2$ (a tripling of the stimulus/factor of 3), and only for the M-Turk online, area-based elicitation method, does reliability consistently equal or exceed 95%. Note in the lower plot, the two on-line “Least Square/Best Fit” lines remain above the desired 95% threshold, but with the in-person, weight-based elicitation, three data points fall below this level, thus the associated least square/best fit lines does not consistently exceed 95%. Therefore, if the required level of reliability is 95%, then a 200% increase (a tripling) of the original stimulus is necessary; this result is more dependable if the physical weight or area difference is larger. These results validated the QFD factor of three (0-1-3-9) scoring scale is, in fact, necessary to achieve a 95% reliable choice.

Exemplar Application of Quality Function Deployment (QFD)

Hauser & Clausing (1988) describe the foundational elements of the QFD process as five sequential steps shown in Figure 4.

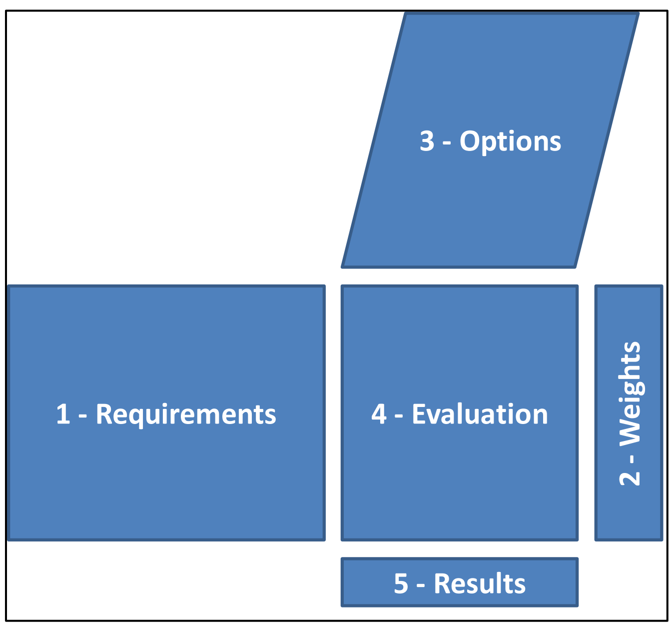

Figure 4. The five primary QFD steps, as described in Hauser & Clausing (1988), are designed to be competed in sequential order. The proportions of the boxes’ areas are representative of level-of-effort and coincidentally connote a roughly equal number of Step-1 requirements with Step-3 solution options.

Step-1 is the collection of requirements against which the solution options will be judged. Not explicitly shown in the figure are the accompanying definitions of each requirement which are critical for both the evaluators and anyone interested in a deeper understanding of the resulting decision. The evaluators need definitions so the solution options can be judged from the same point of reference. After the decision is completed, the definitions help explain and document the decision for anyone wishing to reconstruct how the decision was made. A typical requirements list would include no fewer than four and no greater than nine requirements. When there are three or less requirements, there is a risk of them becoming too broad to accurately define; correspondingly, if there are more than nine the requirements can become too granular and/or over defined.

Step-2 is a weighting of the requirements. Typically, two perspectives are possible 1) all equal, or 2) unequal. In either case, using a 100-point pool is a common reference to most monetary systems and is easily understood. The question at this point is how would the 100 points be divided in whole point increments across the requirements to acknowledge which are more important than others? If all of the requirements are indistinguishable in value, then simply divide the 100 point pool by the number of requirements and set all to the same value. This approach is also an easy sensitivity analysis once the unequal weights are set. If equality is inappropriate, the evaluators must embark on a mutually-agreed upon distribution of the 100-point pool across the requirements to arrive at unequal weights.

Step-3 is the identification and definition of the solution options. Each should be defined to a similar level of detail so they can be fairly judged against the requirements.

Step-4 is the pair-wise comparison of each solution option to each requirement. Ideally, this comparison is done one requirement at a time, not one solution option at a time. Completing this matrix horizontally helps ensure each requirement is being evaluated consistently against each solution option. Completing the matrix vertically (evaluating one solution option against all the requirements then moving to the next solution option) risks a biased view of each requirement. QFD relies on a non-linear, decision scale to complete the translation of subjective/qualitative judgement into quantifiable/quantitative scores. The relationship scale is 0-None, 1-Weak, 3-Moderate, and 9-Strong. Ideally each of the 1-3-9 levels also has an associated definition for each requirement to ensure consistent scoring by the evaluators. A non-linear, ratio scale such as this is designed to purposefully over weight and accentuate solution options that strongly interact with, or strongly satisfy, a requirement. This helps focus attention of drivers and de-focus attention on the many potential details.

Step-5 is the sum-product-matrix math of the Step-2 requirement's weights with Step-4's 0-1-3-9 evaluations; therefore, when a 100-point pool is used to weigh the requirements, the maximum possible score obtainable by any solution option is 900. To more easily interpret output, the scores can be and typically are normalized back to a 0-100 scale simply be dividing each raw score by 9.

Figure 5 shows an exemplar taken from an actual application of QFD in a recent (2014-2015) collegiate-training aircraft purchase.

		A	B	C	D		
Non-Differentiating Criteria	HP	175	200	180	180		
	Avioncis package	GNS-1000	GNS-1000	GNS-1000	GNS-1000		
	Purchase Price	Nominal	Higher	Nominal	Nominal		
Differentiating Criteria						Equal	Unequal
Fleet Commonality	Make a logical transition aircraft moving from DA-20 to PA-28-201R	9	3	3	1	9	5
Business Basis	Investment value coupled with associated Level 5 simulator purchase	1	3	3	9	9	15
Ownership / Operating Costs	Direct and indirect operating costs	3	1	3	3	9	5
Enrollment Appeal	Ramp/curb/bling appeal - diversifies fleet, adding additional training basis	3	9	1	3	9	30
Pre-PVT Curriculum	Backwards compatible with Pre-PVT student abilities, supplements DA-20 fleet	9	1	3	3	9	5
Post PVT Curriculum - IMC Capable	Forward compatible with IMC need and adds value to overall professional training experience	9	9	9	9	9	10
Strategic Seed	Opens relationship with manufacturer with whom long-term fleet replacement would be considered	1	9	3	3	9	10
Ergonomics	Interior phyiscal layout, panel, controls, front / rear seats accommodate diverse student / insturctor profiles	3	9	1	1	9	5
Training Safety	Demonstrated compatibility with the demands of a training environment	3	3	3	9	9	5
Parental Safety Impression	A non-aviator's view of safety features and/or record	3	9	1	1	9	5
Flight Qualities	Overall quality of the flying experience; solidity, crispness of control, stability, airframe integration......	1	9	3	1	9	5
	Equal	405	585	297	387	99	100
	Unequal	360	670	280	440		

Value Scale	
blank	None
1	Marginal
3	Measurable
9	Cleary Superior

Figure 5. QFD fully applied to a collegiate-training aircraft selection. All five steps are shown completed. Note the output scores for options "A", "B", "C" and "D" in this example were not normalized to a 0-100 scale, rather they were left in their raw, 0-900 format.

In this application, the requirements space was first differentiated between three, non-differentiating and 11, differentiating criteria. *Non-differentiating requirements* were adjudged to be incapable of distinguishing one solution option over another and included engine horsepower, the avionics package (which was identical across the four, solution options), and purchase price (which was simply documented as either 'nominal' or 'higher'). The evaluator's recognized a difference in price, but did not consider it to be a driving consideration.

The 11 *differentiating criteria* reflect that the collegiate program was considering adding a third aircraft type to an already mixed fleet. Tactically, this new, third aircraft type was desired to be primarily an instrument flight trainer. Strategically, the addition was envisioned to be a seed-purchase which would open a relationship with the manufacturer to ultimately facilitate a complete fleet replacement with the same aircraft make/model type. Due to this decision complexity, there were more than the typical maximum nine requirements in this application – which potentially risked dilution and duplication of the requirements. Note the accompanying definitions which helped to defend the evaluator's desire for 11, distinct, requirements. Both an equal and unequal set of requirements weights were include on the right side of the evaluation matrix. In the equal weighting case, 100/11 = 9 whole points per requirement. The unequal array of weights was agreed to by the evaluators in a facilitated discussion and essentially contained three tiers of requirements by value. One requirement, 'Enrollment Appeal', at 30 of the 100-available points, dominated over the remaining 10 requirements. There were three requirements in the next tier, weighted at 10 or 15 points each, with the remaining seven requirements each valued at 5 points.

The evaluators also chose to modify the QFD "None/Weak/Moderate/Strong" scoring scale verbiage to "None/Marginal/Measurable/Clearly Superior" prior to completing the scoring of each solution option. No math is required to determine that one solution, option "B," will almost surely lead the others in point value, simply because it has the largest number of "Clearly Superior" satisfied requirements.

The matrix, sum-product math was completed between the evaluations of each option's ability to satisfy each requirement and each array of requirement weights and documented at the bottom of the matrix. For ease of discussion, the raw scores from Figure 2 were each divided by 9 and are shown below in Table 1.

Table 1.

QFD-output scores (Figure 2) re-normalized from the previous 0-900 scale to a 0-100 scale.

	A	B	C	D
Equal	45	65	33	43
Unequal	40	74	31	49

*Note: The resulting 0-900 point scale proportions were preserved with the ability to more efficiently interpret the data on a 0-100 point scale.

In this application, solution option "B", scored better than the other three options and should have been considered the preferred option. Because all weights and the scoring scale are ratio scales, the outputs can be directly compared to each other in magnitude. Observe the equal Step-2 requirement's weight scores show "B" as the preferred choice with twice the value of the "C" and half again as much as either "A" or "D". However, when the unequal Step-2 requirement's weight scores are used, option "B" retains the highest preference while option's "A" and "C" preference erode and option "D" rises, but not enough to be competitive with option "B". The option "D" score rose because it satisfied a valued requirement ('Business Basis') at the maximum level (9).

A significant, positive benefit of using QFD is the ability to successfully translate subjective and qualitative data into a quantitative and defensible choice and ultimately a defensible procurement position for presentation to executive leadership decision makers to purchase aircraft "B". In this case, the weights on the requirements did not matter, either an equal or unequal view produced the same domination of option "B". It is thus possible to methodologically choose a collegiate-training aircraft.

Conclusions & Recommendations

In response to Research Question 1) - using QFD with a 0-1-3-9 scoring scale offers a 95% confidence that if care has been taken with definitions and evaluation, the resulting preference choice is defensible. The trade for the evaluator's investment in time is reaped in dependable, after-the-fact, decision-logic reconstruction. Being able to defend with confidence capital-intense decisions, such as choosing a collegiate-training aircraft, has great value for stakeholders, students, executive leadership, and the institution.

In response to Research Question 2) – decisions without a method may seem easy but the employed decision logic is nearly impossible to reconstruct after the fact. Using a method (such as QFD), which is based on definitions and a ratio scoring scale, requires significant evaluator input and process buy-in but in exchange, offers, easy-to-follow, defensible preference of the resulting choice.

References

Akao, Y. & Mazur, G.H. (2003). The leading edge in QFD past, present and future, *The International Journal of Quality & Reliability Management*; 2003: 20, 1; ABI/INFORM Collection, p. 20

Britt, S.H. & Nelson, V. (1976). The marketing importance of the "just noticeable difference", *Business Horizons*, 19: 38-40. doi:10.1016/0007-6813(76)90063-X

Burke, E., Kloeber, J.M., & Deckro, R.F. (2002). Using and abusing QFD scales, doi:10.1081QEN-120006707

Burge, S. (2009). The systems engineering tool box, Retrieved from https://www.burgehugheswalsh.co.uk/uploaded/1/documents/pugh-matrix-v1.1.pdf

Delano, G., Parnell, G.S., Smith, C., & Vance, S.M. (2000). Quality function deployment and decision analysis-A R&D case study, *International Journal of Operations & Production Management*, Vol. 20 Issue: 5, p.591-609, https://doi.org/10.1108/01443570010318959

Fiorenzo, F. & Alessandro, P. (1999). Rating scales and prioritization in QFD, *International Journal of Quality & Reliability Management*, Vol. 16 Issue: 1, pp.85-97, https://doi.org/10.1108/02656719910250881

Franceschini, F. & Rossetto. S. (1995). The problem of comparing technical & engineering design requirements, *Research in Engineering Design*, Vol 7:270-278, https://doi.org/10.1007/BF01580463

Franceschini, F. & Rupil, A. (1999). Rating scales and prioritization in QFD, *The International Journal of Quality & Reliability Management*, Bradford Vol. 16, Iss. 1, p.85-96, https://doi.org/10.1108/02656719910250881

Hauser, J. & Clausing, D. (1988). The house of quality, *Harvard Business Review*, vol. 66, no. 3, pp. 63–73. Retrieved from https://hbr.org/1988/05/the-house-of-quality

Huber, G.P. (1974). Multi-Attribute Utility Models: A Review of Field and Field-Like Studies, M*anagement Science*, Vol 20, Issue 10, https://doi.org/10.1287/mnsc.20.10.1393

Krider, R.E., Raghubir, P., & Krishna, A. (2001). Pizzas: π or square? Psychophysical biases in area comparisons. *Marketing Science*, 20(4), 405-425. https://doi.org/10.1287/mksc.20.4.405.9756

Vance M., Carstens J., Gasper M. & Parker, P. (2018). Quantitatively Exploring Human Preference, Journal of Management Science and Business Intelligence, 2018, 3–1, July 2018, pages 15-30, doi: 10.5281/zenodo.1246313

Von Winterfeldt, D. & Fischer, G.W. (1975). Multi-Attribute Utility Theory: Models and Assessment Procedures, *Utility, Probability, and Human Decision Making. Theory and Decision Library*, Vol 11, https://doi.org/10.1007/978-94-010-1834-0_3