120th ASEE Annual Conference & Exposition
FRANKLY, WE DO GIVE A D*MN
June 23-26, 2013

Paper ID #7008

The Quality of Engineering Decision-Making in Student Design Teams

Mr. Nicholas D. Fila, Purdue University, West Lafayette

Nicholas is a Ph.D. student in engineering education at Purdue University. His research interests include engineering design, team learning, and instructional laboratories. He has conference publications on cooperative learning, engineering laboratories, innovation, and design.

Dr. Senay Purzer, Purdue University, West Lafayette

Senay Purzer is an Assistant Professor in the School of Engineering Education and is the Director of Assessment Research for the Institute for P-12 Engineering Research and Learning (INSPIRE) at Purdue University. Dr. Purzer is a NAE/CASEE New Faculty Fellow. She is also the recipient of a 2012 NSF CAREER award, which examines how engineering students approach innovation. Her expertise is on assessment and mixed-methods research.

The Quality of Engineering Decision-Making in Student Design Teams

Abstract

Engineers often must decide between multiple design options that present tradeoffs. Engineering students gain experience making and justifying such decisions during team design projects or capstone courses. While previous studies of engineering student design teams have focused on idea generation, conceptual design, team dynamics, and individual skills such as communication and leadership, few have studied decision-making within these teams. This paper presents a framework for evaluating justifications of engineering design decisions and investigates the quality of engineering decision-making during a team engineering design decision-making task in a first-year engineering course. Nineteen teams of three or four students were asked to investigate three design options and recommend one to a client. Decisions were evaluated based on the quality of indicated criteria, the quality of data used to support the decision, and the alignment between evidence gathered and criteria selected. Analysis of the team memos indicated that most teams supported their decisions with applicable data, but did not rely on accurate data or consider all appropriate data and criteria in their decisions. In addition to presenting a rubric that can be used to evaluate student engineering design decisions, this paper will contribute recommendations for instruction on engineering design decision-making.

Introduction

Decision-making is a critical aspect of the engineering design process. Throughout this process, engineers perform a variety of decision tasks such as selecting criteria and performance measures, targeting specific user needs, balancing benefits and tradeoffs of design options, identifying avenues of research and development, determining resource allocation, and opting for design modifications[1-4]. Decisions are so prevalent that some describe design as, essentially, a series of decisions[2].

Perhaps the most important decision task in engineering design is selecting one design alternative, among many, to further develop. Such a choice requires extensive commitment of time, money, and other resources[2], and thus, engineers must be able to employ strong decision-making practices. Further, engineers must be able to justify the resulting decisions to clients, users, and other stakeholders. The small number of studies on this type of decision-making suggest that engineers do not employ thorough decision-making processes[1,5] or provide sufficient documentation or justification of their decisions[1]. Further investigation is required in order to determine how engineering students justify their decisions and whether the resulting decisions and justifications reflect best practices in engineering design.

The purpose of this paper is to develop a framework to investigate and assess the quality of engineering students' formal justifications of their engineering decisions. Using this framework, we identify aspects of decision justification with which students struggle with an end goal of identifying need areas for instruction. Further, we present a rubric for evaluating engineering design decision justifications that can be used as a classroom assessment or a way to scaffold student decision-making during design projects.

Decision-Making in Engineering Design

Decisions require individuals to evaluate two or more optional outcomes and commit to a single option[6]. According to Jonassen[6], decisions can be either rational or naturalistic. Rational decision-making involves evaluating options based on criteria relevant to a given situation and selecting the option that provides maximum utility. An alternative to this view is naturalist decisions which are influenced by beliefs and prior experiences[6].

Engineering design has traditionally favored rational (or normative) approaches[5], such as multiple attribute utility theory, decision matrices, SWOT analysis, and fuzzy sets[1,6]. The benefit of such approaches is that alternatives are evaluated systematically and a single best option can be determined. Thus decisions can be considered comprehensive, objective, and rigorous[6]. Rational approaches, however, have been criticized because (1) not all design alternatives can be accurately evaluated on all criteria[1] and (2) they do not reflect the way people make decisions in authentic settings[6].

Professional and student designers employ some elements of rational decision-making, but in a largely haphazard manner. During both field observations and protocols, designers have been found to employ only a small number of criteria[1,5], thus neglecting many relevant criteria. Dwarakanath and Wallace[5] also observed that designers often revised criteria or repeated assessments on criteria throughout the design process, and further, that the purpose of some criteria could not be identified. Girod and colleagues[1] observed that both student and professional design teams spent little or no time discussing or justifying relevant importance of criteria (i.e. criteria weights). These results suggest that designers or design teams do not always strategically identify and select the limited number of criteria they do use.

In addition to weaknesses defining and weighing criteria, researchers have observed weaknesses in the way designers, particularly students, provide performance evidence for their criteria. Younker and McKenna[7] found that first-year student teams supported only 31% of their decisions with evidence. About half of those decisions were supported with self-generated evidence, which more frequently included pro/con assessments than performance results. The remaining decisions were supported by outside sources such as literature. While compiling relevant external information is essential to many design processes, Wertz and colleagues[8] suggest that engineering students often fail to assess the credibility of external sources. Thus, students infrequently support decisions with data and rely on possibly inaccurate or untrustworthy data.

Decision Justification Framework

Our framework for justifying engineering design decisions is based on the argumentation structure first suggested by Toulmin[9] and then elaborated by Driver, Newton, and Osborne[10].

According to this structure, arguments consist of four elements: *claims, evidence, warrants,* and *backing*. The claim is an assertion one is attempting to forward. Evidence consists of data or information used to justify that claim over other potential claims. Warrants describe how the evidence is relevant to the claim. Backing justifies the warrants. Backing and warrants may be implicit[10].

Argumentation can be further separated into *dialogical* and *rhetorical* argumentation[10]. The purpose of dialogical argumentation is to negotiate meaning, a decision, or a hypothesis. Engineers and engineering students participate in this type of argumentation when choosing one solution among many alternatives. Conversely, the purpose of rhetorical argumentation is to promote a conclusion an individual or group has already drawn. Engineers or engineering students perform rhetorical argumentation when justifying a design decision (e.g., selecting one design alternative over others) to some stakeholders (e.g., clients, bosses, instructors, etc.). In other words, dialogical argumentation describes the process of arriving at a decision, and rhetorical argumentation describes the content and appropriateness of that decision. Since this paper explores the quality engineering design decisions and justifications, we will focus on rhetorical argumentation.

Blair and Johnson[11] suggest a framework for evaluating rhetorical arguments based on three criteria: *relevance, sufficiency*, and *acceptability*. An argument is relevant if evidence used to support the claim is linked to the claim (i.e., the warrants and backing are strong/justified). An argument is sufficient if the evidence used is strong enough to support the decision (i.e., supporting evidence suggests the claim even when considering possible uncertainty or contradictory evidence). An argument is acceptable if the evidence is correct, reliable, and trustworthy. While Blair and Johnson do not present a rubric or metric for scoring arguments, their framework addresses key issues of arguments, and suggests a strong argument will be relevant, sufficient, and acceptable.

In many ways, justifications of engineering decisions are structured like rhetorical arguments. The decision, for example selection of a design alternative, is the claim one is attempting to forward. This decision is justified through the use of data (e.g., performance measures, estimates, qualitative evaluations). Criteria link data and the decision by justifying the data's applicability to the decision. Criteria, derived from stakeholder needs and constraints of the problem help explain what data will be relevant to the decision. For example, when selecting between power supplies for an electronic device, energy efficiency data and unit price on design alternatives would be relevant because energy efficiency and cost are likely to be important criteria. Material color is likely to be irrelevant unless the power supply is visible in the design and aesthetics are important to potential users or other stakeholders.

Because decision justifications can be thought of as rhetorical arguments, they might also be evaluated based on Blair and Johnson's framework (see Table 1). Relevance concerns the fit between the evidence and the claim (i.e. the quality of warrants and backing). In an engineering sense, this translates to the quality of the criteria, and the fit between the data and criteria. Acceptability concerns the correctness and trustworthiness of data. In an engineering sense, this means that data must provide reliable estimates of performance, and must be provided for all possible options. Sufficiency would indicate whether the decision is justified based on the data that are available. In an engineering sense, this means that the data used demonstrate the decision as the option for which benefits most outweigh the tradeoffs, and further, all available data factor accurately into that assessment. Table 1 demonstrates the indicators of strong engineering decisions.

Table 1. Elements of Arguments and Justifications

Element of Argument	Decision Focus	Indications of strength for design decision
Relevance	Criteria	Criteria reflect stakeholder needs and problem constraints; data match criteria
Acceptability	Data	Data are accurate; data are comprehensive
Sufficiency	Data usage in decision	Data support decision; all accessible data factor into decision

Methods

Setting and Participants

This study took place in the context of a first-year engineering design course. The course aimed to promote understanding of the engineering design process through activities and projects outlining the overall design process; highlighting key stages such as problem scoping, idea generation, concept reduction, concept selection, and solution detailing; and techniques and tools to aid the design process. Students were also introduced to computational tools such as spreadsheet programs and Matlab.

Participants were 73 students (8 female, 65 male) enrolled in a single section of the course. Early in the semester, students were grouped into 19 teams of three or four students using a computer-based team-development program[12]. Students completed in-class projects, an eight-week design project, and a practical exam in their teams. These teams were the unit of analysis for this study.

Data Collection

Data were collected during the course of a team practical exam. Teams were presented with a client, a public library, who prompted them to recommend one of three options to minimize energy cost: install solar panels, install a green roof, or make no changes. Students were given data sheets with information about the client, the design context, solar panels, and green roofs. Students did not have access to additional information but could ask an exam proctor (either teaching assistant, instructor, or researcher) to clarify information. Although the design situation was hypothetical, the data and information presented were based on a real-life library and its energy usage[13], and simplified cost and performance information for solar panels and green roofs. Students were required to perform calculations of energy usage and system cost for each of the three alternatives and document their process, data, decision, and justification in a short memo. We collected the final memo, spreadsheet calculation, and handwritten notes from each team. The teams were also video and audio recorded.

Rubric Development

We developed a rubric (Figure 1) to assess the quality of decision justifications based on the framework presented earlier. The rubric consists of six categories (see Table 2) representing aspects of well-justified decisions. Decisions are scored either 0 (not evident), 1 (needs improvement), 3 (acceptable), and 5 (excellent) in each category. Scoring is aided by descriptions of decision features that should result in each score for each category.

Category	Sub-category	0	1	3	5
Relevance (criteria link data to decision)	Appropriate criteria	Team did not use any of the most important stakeholder criteria	Team used most of the important stakeholder criteria AND used at least one irrelevant or redundant criterion	Team used the most important stakeholder criteria AND no irrelevant or redundant criteria were used	Team used all criteria important to stakeholders AND no irrelevant or redundant criteria
	Appropriate data	No data used in decision	Use at least one type of data that does not match a criterion	All data match at least one criterion BUT data do not cover all criteria OR all criteria covered but data do not exactly reflect criterion	All data match at least one criterion AND data cover all criteria
Acceptability (trustworthy data)	Accurate data	No data given	Data are inaccurate beyond reasonable error margin	Data are accurate within reasonable error margin (10%)	All data used in the decision are accurate
	Comprehensive data	No data given	Does not include data for all of the listed criteria	Includes data to measure each criterion listed, but does not include measurements for all design options	Includes data to measure each criterion listed for each option
Sufficiency (data support decision)	Comprehensive data usage	No explanation of data sufficiency given	Did not describe why selected system was better than all other options	Described why selected system was better than all others based on supporting evidence only (in presence of evidence that does not support decision)	Described why evidence supporting decision was sufficient, rebutted evidence that does not support decision (if applicable)
	Accurate data usage	No data used in decision	Data listed in support of decision do not support the decision OR positive evidence is clearly outweighed by negative evidence	All data listed in support of decision actually support the decision BUT positive evidence is weak or may be outweighed by negative evidence	All data listed in support of decision actually support the decision AND positive evidence outweighs negative evidence

Figure 1. Decision Assessment Rubric

Table 2. Rubric Category Descriptions

Quality of Argument	Rubric Category	Explanation
Relevance	Appropriate criteria	Criteria accurately reflect stakeholder needs. Important criteria are weighted more heavily and irrelevant or redundant criteria are excluded.
	Appropriate data	Data presented accurately reflect each criterion, thus sufficiently linking data used to decision.
Acceptability	Accurate data	Data provide accurate or reasonable estimates of performance measures.
	Comprehensive data	Data provided for each solution based on each criterion.
Sufficiency	Comprehensive data usage	All data factor into final decision, indicated decision is informed by all available information.
	Accurate data usage	Decision is supported by the given data, i.e. data do not suggest another decision. Teams may have to justify balance of benefits and tradeoffs in light of conflicting data.

The development of this rubric was an iterative process resulting in five drafts. During each iteration, we applied the rubric to a sample of students decisions. We made changes to the descriptions to achieve greater clarity, parsimony, and reliability. We also identified changes to the category structure that (1) resulted in independent and non-overlapping categories and (2) covered all features of well-justified decisions based on the framework.

Drafts 1 and 2 contained five categories: *comprehensive criteria* (criteria cover the range of stakeholder needs), *measurable criteria* (criteria could be measured based on available data), *accurate data* (data arrived from correct interpretation of information sheets and error free calculations), *comprehensive data* (data were included for each criterion), and *sufficient data* (data were used to compare alternatives and supported the team's recommendation). In draft 3, we split *sufficient data* because it represented two separate elements: *comprehensive data usage* and *accurate data usage*. *Comprehensive data usage* indicated the amount of data collected that teams factored into their final decision and *accurate data usage* indicated whether the data that were used actually supported the decision. We added another category, *appropriate data usage*, to draft 4 in order to indicate whether data matched the intended criterion, thus securing the link between data and the decision.

After draft 4 was completed, we prepared another rubric with wording specific to the design scenario. For example, we included the specific criteria that should be included in the *comprehensive criteria* category (system cost and CO_2 emissions). We also prepared examples of decision elements that would result in scores of 0, 1, 3, and 5 for each category. One graduate research assistant and one undergraduate research assistant scored each team's decision and justification based on the rubric and its supplementary materials. The raters were given access to the each team's memo and calculations. The raters did not have access to the video during the rating process to ensure that assessments were made based only on the information each team intended to present to the client, thus reflecting the decision justification collectively decided upon by each. This process ensured that rating was based on rhetoric argumentation of the team and not any dialogic argumentation performed during the actual task.

For each category but *comprehensive data usage*, raters' scores indicated sufficient degrees of inter-rater reliability. The raters discussed any differences in their assessments and found areas where they differently interpreted the rubric. Most of the differences occurred in the *comprehensive data usage* category, as was reflected by the reliability statistics. These differences were clarified and resolved and the wording in the rubric was modified based on these discussions. We also noted that the raters' scores were nearly identical for the categories *appropriate data usage* and *measurable criteria*. Inspection of their wording revealed similar constructs, so we merged the categories into a new *appropriate data* category. We created a final draft of the rubric based on these changes.

Two raters (one original rater and an undergraduate research assistant) again scored all student decisions based on the updated rubric. We switched one of the raters to ensure at least one rater was not familiar with the development of the rubric and thus was rating based on the wording of the rubric and not internal discussions. In order to determine reliability of each finalized scale, we compared each rater's scores using Cronbach's alpha and Pearson's r. After this iteration, Pearson's r values were all statistically significant and Cronbach's alpha values were near or above .80.

Table 3. Reliability Statistics

Category	Cronbach's alpha	Pearson's r	P-value
Appropriate Criteria	.86	.75	<.001
Appropriate Data Usage	.99	.97	<.001
Accurate Data	.79	.72	<.001
Comprehensive Data	.92	.86	<.001
Comprehensive Data Usage	.82	.72	<.001
Accurate Data Usage	.89	.81	<.001

Evaluation of the Quality of Student Decision Justifications

The final decision justification scores represent the average score on each category awarded by each of two raters. Table 4 presents the mean and standard deviation of the scores for the sample of 19 student teams. We used the value three (out of five), representing "acceptable" scores based on the rubric's wording, as a threshold for indicating potential instructional need areas. Students scored above three on the categories: *appropriate data*, *comprehensive data*, and *accurate data usage*. Thus, for the most part, the teams selected data that were well matched to their criteria, collected data relevant to their decision, and reported those data and their implications accurately. Students scored below three on the categories: *appropriate criteria*, *accurate data*, and *comprehensive data usage*. The mean score for *accurate data* was particularly low. These scores indicate that teams indicated irrelevant or redundant criteria or missed important criteria, relied on inaccurate data, and only used a portion of the data they collected or intended to collect to justify their decisions.

Table 4. Team Decision Justification Scores

Category	Mean	SD
Appropriate data	4.81	0.54
Accurate data usage	4.25	0.86
Comprehensive data	3.88	1.54
Comprehensive data usage	2.88	1.50
Appropriate criteria	2.47	1.36
Accurate data	1.00	0.37

Discussion

Findings on Student Decision Justification

Rubric results indicated that students performed well on a number of aspects the decision-making task completed in this study. The teams employed appropriate data that corresponded to the criteria they selected. They also, for the most part, compiled most the data relevant to the decision task and made decisions consistent with these data. These results suggest that, at least for the task in this study, students will (1) make strong decisions given relevant data and (2) students can identify data that are relevant to the decision task.

The areas in which students had difficulty justifying their decisions correspond to areas highlighted by previous research studies. Wertz and colleagues[8] note that students often fail to discriminate between credible and unreliable data sources. While Wertz and colleagues referred to credibility of external sources, in this study, students themselves produced the inaccurate data. Informally, observations of the videos showed that students often failed to check their own calculations or the calculations of their teammates. The failure to check work was particularly egregious when considering the most common data error was calculating the cost of the full coverage green roof option to be $26M, or roughly 40 times the correct cost (which was comparable to the cost of the other alternatives). While our results did not reflect failure to support decisions with evidence, as found by Younker and McKenna[7]—all teams supported their decisions with some evidence—it is concerning when paired with the finding that half of first-year evidence-based decisions were supported with self-generated evidence, the type of evidence students used in this study. It is important that students learn to support decisions with data, but they also must learn to ensure that those data are accurate or trustworthy.

Despite relying on inaccurate data, students did link those data to appropriate criteria. Those criteria, however, did not always reflect stakeholder needs. Rubric scoring results showed that students failed to select all appropriate criteria. Most often, students did not account for the CO_2 emissions criterion, which the clients listed as important but secondary, and for which had no obvious numerical data was made available. They also suggested criteria, such as overall energy savings or maintenance costs, which were already accounted for by overall system cost. Research on engineering decision-making, and human decision-making in general suggests that individuals and teams often miss key criteria or suggest criteria that are minimally relevant to the situation[14]. Further, students did not compare all alternatives on all criteria they selected, a finding corresponding to those by Girod and colleagues[1].

Findings on Decision Justification Rubric

Two independent raters (one graduate research assistant and one undergraduate research assistant) each scored 19 student decisions based on the rubric to high degrees of consistency. Since one rater was unfamiliar with the rubric prior to rating, this indicates that the rubric is easy to understand and use to assess decisions reliably. An additional strength of the rubric is that key decision justification areas received orthogonal rating. In essence, penalization on one area of justification does not indicate penalization on another area. Because of this, the rubric can identify specific areas students may be struggling in their decision justifications.

It is appropriate here to discuss one practical aspect of rubric implementation. Student decisions in this study were often poorly detailed due to time constraints of the task. When individual raters discussed use of the rubric, and based on notes taken during assessment, they indicated some difficulty interpreting exact criteria and data use structures. For example, students used inconsistent wording or only vaguely alluded to data contained in different sections of the report.

The raters scored reports consistently, so this issue did not seem to affect the results of the study, but the raters' lack of certainty with their assessments could portend issues on future implementations of the rubric. We suggest employing negotiated coding[15] on an initial sample of decisions so that raters can identify potential issues and reach agreement on difficult cases.

We also expect that, to some degree, decision scores are related to the specific task and may have been different had the task been altered. For example, students were given a limited time to complete the task. They may have spent more time ensuring the accuracy of their data if given more time. Further, they were limited to information on data sheets that they had to transform into data. One on hand this ensured that they perform somewhat complicated calculations, on the other, it ensured that they had at least some data on which to base their decision. Given the freedom to pursue their own data, students may have supported their decision with less data, untrustworthy data, or may have found ready-made data that were more accurate but less related to the criteria. Those who employ the rubric should consider the nature of the specific decision task and its potential implications on evaluation.

Conclusions

In this paper, we have presented a rubric that can be used to assess the quality of student decision justifications. The rubric is easy to use and has strong inter-rater reliability. It is useful not only for assessing the quality of decision justifications, but diagnosing areas in which the decisions are strong or weak.

Future work should focus on three key areas. First, the rubric should be tested on a variety of addition decision-making tasks such as decisions made during authentic design projects. Such work could identify context-specific differences in the quality of student decision-making and justification. Second, the rubric can be use in instructional settings. It can be employed as an assessment tool for student design projects or activities. Further, this type of structured decision justification template can be used as a way of scaffolding student learning until they are able to support their decisions with accurate, appropriate, and sufficient evidence. Finally, this study addressed the quality of decision outcomes, but not the process used to arrive at those decisions. In future studies, we plan to investigate elements of team decision-making process that contributed to strong decisions and justifications.

Acknowledgments

We would like to the acknowledge the contributions of Kavin Nataraja and Julian Wang, who acted rubric raters for this study. This material is supported by funding from the Purdue Research Foundation.

References

1. Girod, M., Elliott, A.C., Burns, N.D. & Wright, I.C. Decision making in conceptual engineering design: an empirical investigation. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* **217**, 1215-1228 (2003).
2. Ullman, D.G. Robust decision-making for engineering design. *Journal of Engineering Design* **12**, 3-13 (2001).
3. Bucciarelli, L.L. *Designing engineers*, (MIT Press, Cambridge, Mass., 1994).
4. Crismond, D.P. & Adams, R.S. The Informed Design Teaching & Learning Matrix. *Journal of Engineering Education-Washington* **101**, 738 (2012).
5. Dwarakanath, S. & Wallace, K.M. Decision-making in engineering design: observations from design experiments. *Journal of Engeering Design* **6**, 191-206 (1995).
6. Jonassen, D.H. Designing for decision making. *Educational Technology Research and Development*, 1-19 (2012).
7. Younker, J. & McKenna, A. Examining student use of evidence to support design decisions. in *2009 ASEE Annual Conference and Exposition, June 14, 2009 - June 17, 2009* BOEING (American Society for Engineering Education, Austin, TX, United states, 2009).
8. Wertz, R.E.H., Fosmire, M. & Purzer, S. Assessing Information Literacy Skills Demonstrated in an Engineering Design Task. *Journal of Engineering Education* (In press).
9. Toulmin, S.E. *The uses of argument*, (Cambridge University Press, 2003).
10. Driver, R., Newton, P. & Osborne, J. Establishing the norms of scientific argumentation in classrooms. *Science education* **84**, 287-312 (2000).
11. Blair, J.A. & Johnson, R.H. Argumentation as dialectical. *Argumentation* **1**, 41-56 (1987).
12. Ohland, M.W., *et al.* Smarter teamwork: System for Management, Assessment, Research, Training, Education, and Remediation for Teamwork. in *ASEE Annual Conference and Exposition* (American Society for Engineering Education, Vancouver, BC, Canada, 2011).
13. Rezaie, B., Esmailzadeh, E. & Dincer, I. Renewable energy options for buildings: Case studies. *Energy and Buildings* **43**, 56-65 (2011).
14. Proctor, R.W. & Van Zandt, T. *Human factors in simple and complex systems*, (CRC PressI Llc, 2008).
15. Garrison, D.R., Cleveland-Innes, M., Koole, M. & Kappelman, J. Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education* **9**, 1-8 (2006).