Contrasting cases: What we can learn from students' perceptions of "good" design

Joan M. T. Walker, Long Island University, School of Education, 720 Northern Boulevard, Brookville, NY 11458, joan.walker@liu.edu
Paul H. King, Department of Biomedical Engineering, Vanderbilt University, Nashville, TN 37325, paul.h.king@vanderbilt.edu

Abstract: Set in the context of an engineering design course, this study used contrasting cases to examine undergraduates' (1) ability to discriminate poor and excellent examples of student design projects, and (2) students' justifications for their ratings. At the beginning of the yearlong course, only 76% of students correctly identified the better project. Toward the end of the course, the majority of students correctly identified the superior example; however, many failed to note serious flaws in the poor example. Analyses of student justifications showed that most students used superficial reasons (e.g., readability) initially and deeper structural reasons (e.g., process) on the post-test. Results are discussed in terms of expert-novice differences, and how educators can support students' ability to notice salient problem features.

Introduction

Expertise is the goal of education in any domain. However, educators struggle with how best to teach for expertise (e.g., Bransford, Brown & Cocking, 1999; Hatano & Oura, 2003; Lajoie, 2003). Research in recent decades has advanced our understanding of expertise and points to interesting avenues for designing instruction that supports its development. Among the most striking findings to emerge is evidence that novices and experts view the world very differently. For instance, when asked to organize several physics problems, experts grouped them according to their deeper principles (e.g., conservation of energy) whereas novices arrayed them according to their surface features (e.g., involved inclined planes; Chi, Glaser & Farr, 1988). In other work expert teachers, advanced beginners and novice teachers were asked to state what they noticed when observing a teacher engaged in classroom instruction (Sabers, Berliner & Cushing, 1991). Experts were far more likely to offer evaluative comments and interpretations of what they observed and to make suggestions for improvements. By contrast, novices and advanced beginners were only able to offer descriptions of what they saw, much like a sports announcer offering play-by-play during a basketball game.

Differences between experts and beginners are consistent with theoretical and philosophical arguments that the world is invariant; it is only our perceptions that change (Gibson & Gibson, 1955). Further, this phenomenon supports the idea that learning can be measured as changes in perception (Lave & Wenger, 1998). However, assessments that tap into this phenomenon are rarely included in classroom instruction. One recently developed technique that may help students develop the critical thinking skills that are a hallmark of expertise is contrasting cases, which involves using carefully arranged contrasts to underscore conceptual differences (Barron, Schwartz, Vye, Moore, Petrosino, Zech & Bransford, 1998; Bransford & Schwartz, 1999).

From the perspective of cognitive psychology, the ability to discriminate features of a problem is developed through experiences that provide opportunities to notice important contrasts and similarities among problems (Bransford, Franks, Vye & Sherwood, 1989). For instance, Broudy (1977) has argued that in addition to "knowing that" and "knowing how" (i.e., replicative and applicative knowledge, respectively) people "know with." Knowing with means that an individual "thinks, perceives and judges with everything he has studied…, even though he cannot recall these things on demand" (p. 12). This suggests that novices fail to differentiate problem features because, relative to experts, they have experienced fewer opportunities to notice what is similar and different or relevant and irrelevant across problems they encounter. Thus, it is logical to assume that educators can support the development of critical thinking skills by creating opportunities for students to reflect on the underlying structural features of problems and the quality of solutions that address them.

This study used contrasting cases to reveal differences in how experts and novices view the practice of engineering design. This work is important because a common task for many engineers is the analysis of an existing design with an eye toward its improvement. Given its importance in the profession of engineering, design is a required course for seniors in the majority of engineering programs in the US. However, design instruction and research often focuses on students' ability to generate a design solution rather than critique the merits of existing ones. For instance, a major requirement in most design courses is the completion of a senior design project and research in design education often focuses on students' responses to problem-solving scenarios and their conceptions of the design process (Atman & Turns, 2001; Cross, 2001; Newstetter & McCracken, 2001; Walker, Cordray, Brophy & King, in press; Walker, Cordray, King & Fries, 2005). Promoting analytical skills in the field of design is important because if engineers are responsible for critiquing and improving existing designs then they must be trained to notice important design features. You cannot fix what you do not see.

Specifically, excellent and poor examples of design were drawn from the archives of former student design projects collected at a major university's engineering program. Students' perceptions of the projects' quality were then compared to the evaluations of experts. The study had several goals. The first goal was to understand what students valued as attributes of good design and their ability to differentiate between two designs of very different quality. It was also intended to help students focus attention on concepts of design quality and to understand the requirements by which their own design projects would be judged. Finally, the study was designed to offer engineering educators an alternative and 'user-friendly' form of student assessment.

The study posed four research questions: (1) can students correctly identify the better of two design projects, (2) are students' quantitative ratings similar to the ratings of experts, (3) how do students justify their ratings, and (4) do students' perceptions of the projects' quality become more expert-like over time? It was expected that at the beginning of a yearlong course, many students would fail to correctly identify the better project, and that this would stem, in part, from basing their judgments on surface features (e.g., meeting content requirements and presentation style) rather than deeper structural features such as the underlying engineering process and emergent product. In the second semester, after students had experienced the challenge of developing their own senior design project, the number of students who correctly identified the better project was expected to increase. Students' perceptions were expected to become more expert-like with increased references to engineering content as a basis for judging the quality of a design.

Methods

Participants included three experts (senior biomedical engineering faculty) and 42 seniors enrolled in an undergraduate engineering design course at a private university in the mid-South of the United States. Experts evaluated the projects in the spring of 2003 and 2004 at the conclusion of the design course. Projects could earn a maximum of 100 points. Criteria for judging the projects included four categories: *Efficiency*, which includes the project's engineering goals (20 points), and demonstrated ability to design system, component or process (25 points); *Innovation,* which includes the project's creativity (20 points) and consideration of alternatives (20 points); *Clarity* of communication and persuasiveness (10 points), and *Ethics* (e.g., issues of exclusion, 5 points).

The contrasting cases were selected on the basis of the experts' ratings. One project explored a technique for imaging the heart and had been rated as "excellent" ($M = 92.00$, $SD = 6.55$; range = 86-99). The other project sought to improve devices designed to assist newborns' breathing and was rated "below average" ($M = 65.66$, $SD = 16.19$; range = 47-76). Descriptive statistics for each project by category are presented in Table 1. Strengths of the excellent example included the level of challenge, creativity when setbacks were encountered, and the development of a working prototype. Weakness of this project included the fact that the paper describing the project's development was "annoyingly redundant at times" and contained technical jargon without explanation of those terms. Strengths of the below average paper included a thorough literature review, and clear writing and formatting. Weaknesses of this project were considerable and included the fact that no quantitative analyses were performed, and that the authors "drew intuitive and appealing conclusions, which did not require any engineering." In short, the excellent project demonstrated engineering and research skills, whereas the below average project demonstrated excellent research but no follow-through in engineering. Although the number of expert ratings is small, there was clear consensus among this group's evaluations. Specifically, they were expected to provide sufficient contrasts in quality and a basis for evaluating whether students valued 'function over form.'

Students rated the projects at two time points during a yearlong course analogous to the experience of student teaching for teacher education majors. The first time point was in the fall semester when students attended a traditional lecture-driven course focused on design principles and knowledge. The second time point was in the spring when students spent a full semester in direct field experience, working as a team under the guidance of an advisor, to develop a design project. Students were e-mailed instructions to download papers describing the two former student projects from the course web page. After reading each paper students evaluated it with the same scoring rubric used by faculty; however, students were not informed of the faculty evaluations prior to completing the task. In addition to assigning numeric ratings, students were asked to explain their evaluations. When both papers had been graded, students were asked to explicitly compare the two papers, decide which one was better and explain why.

Analyses and Results

Seventy-six percent of students correctly identified the best project. Variability in student ratings was considerable. Descriptive statistics for student ratings for each project by category are summarized in Table 1.

Table 1. Descriptive statistics for student and expert ratings by project and by category at pre-test

		Below average project			Excellent project			
Category		Mean	SD	*Range*	Mean	SD	*Range*	*t(43)*
Efficiency	Student	34.47	5.52	*22-44*	40.40	3.95	*32-45*	
	Expert	28.00	4.35	*23-31*	41.33	3.51	*38-45*	*1.90+*
Innovation	Student	31.88	5.09	*18-40*	34.05	4.04	*21-40*	
	Expert	25.67	11.01	*13-33*	36.67	2.88	*35-40*	*2.12**
Clarity	Student	7.53	1.53	*3-10*	8.27	1.51	*3-10*	
	Expert	7.00	1.00	*6-8*	9.00	1.00	*8-10*	*1.08, ns*
Ethics	Student	4.68	0.67	*2-5*	4.36	.83	*2-5*	
	Expert	5.00	0.00	*5-5*	5.00	0.00	*5-5*	*.54, ns*
Total	Student	78.57	10.51	*49-97*	86.62	9.02	*63-100*	
	Expert	65.67	16.20	*47-76*	92.00	6.56	*86-99*	*2.23**

+ = p < .06, * = p < .05

Second, we compared students' quantitative ratings to those of experts. Comparison of the groups' average ratings showed that students and experts had similar perceptions of the excellent project but that students rated the poor project much higher than experts. Experts' total ratings of the two projects differed by an average of 25 points, whereas students' ratings differed by less than 10 points. Ratings differed most for the categories of Efficiency and Innovation. The two groups had similar perceptions of the projects' Clarity and Ethics

Preliminary analyses of students' justifications for their ratings suggested that those who correctly identified the excellent project found the task relatively straightforward (e.g., *"Comparing these two papers is like night and day;" "This exercise points out that it is imperative to use math and engineering principles to help prove your design;"* and *"The second paper simply tests a design that has already been made"*). By contrast, students who chose the below average project appeared unaware of its flaws or judged the project on the basis of form rather than content (e.g., *"It was … easier to read;" "good overall readability; the [other] paper was boring and seemed to keep going and going"*). This led to the development of a coding scheme with four categories. The first two categories, meets content requirements (e.g., "contained everything it was supposed to") and readability (e.g., "easy to read, understand"), represent a *Surface* approach to evaluation. The second two categories, product (e.g., "generated a working prototype") and process (e.g., "tested and revised their idea") indicate attention to engineering content or a *Deep* approach to evaluation. Almost all of the students (92%) used surface reasons to justify their ratings, whereas 71% used Deep reasons. Wilcoxon signed rank tests showed the distribution of these related variables to be significantly different ($z = 3.00$, $p < .01$).

Following this assignment, the course instructor revealed the faculty ratings to students and held an in-class discussion of the observed differences between student and faculty evaluations. Students also received extensive feedback from teaching assistants about their performance. This feedback focused on the fact that while many students could identify the strong points of the excellent project, they were less successful in identifying the poor project's weaknesses. The instructor emphasized how this would be problematic when students faced the task of troubleshooting an existing design. Teaching assistants provided students with a list of this project's problems and urged students to "consider this list seriously, because your understanding of this assignment can make or break your final project." Students were told that they would complete a similar assignment again in the spring semester.

Post-test: Procedures, Analyses and Results

Participants and procedures were identical to the pre-test; however, the contrasting cases changed. Experts had rated one project, a new approach to a dosage inhaler device, as excellent (M = 92.67, SD = 4.93; range = 87-96); the other project explored alternative approaches to the design of a cardiac catheter and was rated as "average" (M = 72.33, SD = 5.85; range = 68-79); both projects were chosen because they shared many of the strengths and weaknesses identified in the pre-test exemplars. For instance, the average project had a strong literature review but failed to generate a prototype, drawings or specifications. Descriptive statistics for student and expert ratings for each project by category are summarized in Table 2.

Table 2. Descriptive statistics for student and expert ratings by project and by category at post-test

		Average project			Excellent project			
Category		Mean	SD	*Range*	Mean	SD	*Range*	*t (43)*
Efficiency	Student	32.17	6.88	*9-20*	40.97	3.07	*12-41*	
	Expert	32.67	.57	*32-33*	41.33	2.88	*38-43*	*.03, ns*
Innovation	Student	30.05	5.45	*13-40*	35.17	2.85	*26-40*	
	Expert	27.00	6.24	*22-34*	37.66	1.52	*36-39*	*1.49, ns*
Clarity	Student	6.43	1.71	*3-10*	8.31	1.54	*3-10*	
	Expert	8.00	1.00	*7-9*	8.67	1.15	*8-10*	*1.01, ns*
Ethics	Student	3.83	1.43	*0-5*	5.00	.00	*0*	
	Expert	4.66	.58	*4-5*	5.00	0.00	*5-5*	*.17, ns*
Total	Student	72.10	12.43	*40-90*	86.62	9.02	*63-100*	
	Expert	72.33	5.85	*68-79*	92.67	4.93	*87-96*	*.42, ns*

Analyses were identical to the pre-test. On the post-test 94% of students correctly identified the better project; the variability of student ratings remained considerable. Similar to the pre-test, students rated the excellent example higher on all dimensions with the exception of ethics, which was at ceiling. Unlike the pre-test, students and experts saw the same degree of difference in the two projects' quality (see t-test results in Table 3). Wilcoxon signed rank tests for students' reasons used on pre- and post-tests showed no change in use of surface reasons ($z = 1.00$, $p < .31$; pre-test N= 39; post-test N = 37) and increased use of deep reasons ($z = 1.73$, $p < 05$; pre-test N = 30, post-test N = 36).

Discussion

Grounded in evidence that experts and beginners view the world differently, this study used contrasting cases to learn about the development of students' abilities to notice important conceptual features of engineering design. Results suggest that advanced beginners can often recognize the better of two design solutions; however, they do not always recognize the serious flaws and limitations of poor design work. Specifically, these students appeared to mistake good research for good design. They did not recognize that merely exploring and describing alternatives was essentially a failure to design. Moreover, students' reasons for evaluating the quality of a design is not necessarily related to its deeper structural features such as engineering content or the quality of the underlying design process. These findings are consistent with arguments that novices and beginners sometimes engage in critical thinking without explicit attention to the bases for their judgments (Broudy, 1977).

Students demonstrated enhanced critical thinking skills on a similar task at the end of their design experience. Improvements on the post-test included an increased ability to recognize the better project, greater alignment with experts' quantitative evaluations, and increased use of deep reasons for choosing the better project. This improvement occurred despite the fact that the contrast between the post-test materials was less stark than the contrast between the pre-test materials. We believe that students' improved performance on the post-test is due to their immersion in the field and the opportunity to take on the duties and identity of a designer. Through engaging in the practice of design students gain insight into the myriad of conditions and details associated with the design process. This experience helps them to notice the design decisions of others. In the context of these students' educational experiences, it appear that moving from an abstract understanding of design developed during a lecture-based course to a more intimate knowledge of the practice of design helped students 'see with new eyes' and supported their ability to discriminate the quality of two design projects.

An outstanding question is how students' design knowledge might become further differentiated with instructional support. Addressing this question, we are currently conducting several studies using contrasting cases as an intervention. First, we are currently replicating this study. Second, we have developed and used an additional contrasting cases task, which asks students to choose the better of two proposals addressing the same problem (rather than choosing the better of two solutions addressing different design issues, which is the task described here). Preliminary analyses of these data show that students perform much better on this task at pre-test, suggesting that an optimal scaffolding of students' critical thinking skills would initially demand near comparisons followed by comparisons across an increasingly wider range of problems. This task is also important because it offers a window into how students rate the work of designers who are not their peers. For instance, students might have hesitated to be critical of the projects used here because they knew they were the work of former design students. We are also increasing the size of our expert sample, and assessing the generalizability of this work by exporting the measures described here to other engineering programs.

In sum, by drawing from the products created in their own classrooms engineering educators can develop authentic, meaningful assessments for students that acknowledge the legacy of former design students while promoting the development of their current students' critical thinking skills and differentiated knowledge of design.

References

Atman, C. J. & Turns, J. (2001). Studying engineering design learning: Four verbal protocol studies. In C. Eastman, M. McCracken & W. Newstetter (Eds.) *Design knowing and learning: Cognition in design education,* pp. 37-62. New York: Elsevier.

Barron, B., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998).Doing with understanding: lessons from research on problem- and project-based learning. *Journal of the Learning Sciences, 7,* 271-311.

Bransford, J.D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad and P. D. Pearson (Eds.) *Review of Research in Education, 24*, 61-100. Washington, DC: American Educational Research Association.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds). *Similarity and analogical reasoning* (pp. 470-497). New York, NY: Cambridge University Press.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999) *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Broudy, H. S. (1977). Types of knowledge and purposes of education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.). *Schooling and the acquisition of knowledge* (pp. 1-17). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Cross, N. (2001). Design cognition: Results from protocol and other empirical studies of design activity. *Design knowing and learning: Cognition in design education,* pp. 79-103. New York: Elsevier.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review, 62,* 32-41.

Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational Researcher, 32*(8), 26-29.

Kuhn, D., & Udell, W. (2001). The path to wisdom. *Educational Psychologist, 36(4),* 262-264.

Lajoie, S. P. (2003). Transitions and Trajectories for Studies of Expertise. *Educational Researcher, 32*(8), 21-25.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral practice*. New York: Cambridge University Press.

Newstetter, W. C. & McCracken, W. M. (2001). Novice conceptions of design: Implications for the design of learning environments. In C. Eastman, M. McCracken & W. Newstetter (Eds.) *Design knowing and learning: Cognition in design education,* pp. 63-78. New York: Elsevier.

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal, 28*(1), 63-88.

Sternberg, R. J. (2001). Why Schools Should Teach for Wisdom: The Balance Theory of Wisdom in Educational Settings, *Educational Psychologist, 36*, 227-245.

Walker, J. M. T., Cordray, D. S., Brophy, S. P., & King, P. H. (in press). Design scenarios as a measure of adaptive expertise in design. *International Journal of Engineering Education*.

Walker, J. M. T., Cordray, D. S., King, P. H., & Fries, R. C. (2005). Novice and expert definitions of biodesign: Developmental differences with implications for educators. *International Journal of Engineering Education*, 21, 467-479.

Acknowledgements

This work was supported by the Engineering Research Centers Program of the National Science Foundation (Award #EEC9876363).