Assessing Students’ Knowledge of Design Process in a Design Task

Judith E. Sims-Knight¹, Richard L. Upchurch², and Paul Fortier³

Abstract – Today’s engineering educators widely agree that engineering curricula need to incorporate a larger emphasis on design. ABET has instituted the requirement that programs build continuous improvement loops around student learning outcomes, including the ability to design a system, component, or process to meet desired needs. Success at this enterprise depends on having appropriate assessment tasks, ones that are manageable in the context of coursework and provide pointers toward improvement, yet are authentic. The most widely used way of assessing design skill is to give a design task, but the requirements to do so usually are too cumbersome to use in the context of a class. The paper will describe our struggles to create a procedure within a design task whereby we might assess students’ design process. This paper will describe the tasks, the findings obtained with them, and how they could be used within a continuous improvement loop.

Index Terms - Design Process, Assessment, Continuous Improvement, Design Task

INTRODUCTION

Research on the nature of expertise has revealed that experts’ decision making is not an explicit, logical process but rather an implicit and intuitive process based on domain-specific knowledge acquired over years of practice [1]. Even within a domain such as software design, different kinds of experience (university training vs. practical experience) result in differences in practice [2]. Because expertise is the result of such extensive practice and because much of expertise is implicit, it is very difficult for experts to describe their design process, let alone to teach it by principle.

An alternative approach is to provide a learning environment in which individuals can develop their own expertise through the quality model [3]. Applied to student learning of design process, this would have students identify their own process, assess their process, evaluate the results of their assessment, formulate improvement plans, and repeat the process. To do this requires an appropriate assessment task, one that is manageable in the context of coursework and that helps elucidate the design process to students and provides pointers toward improvement, yet is authentic. Under the auspices of a National Science Foundation grant to develop such assessments, we explored alternative ways of assessing students’ design process within a design task.

CHOICE OF DESIGN TASKS AND PROCEDURE

We had several goals in developing the design task. First, we wished to use a design task of depth sufficient to warrant the use of design process (e.g., design alternatives, feasibility, trade-offs) and that could be assessed by professional standards (e.g., the use of abstractions, models, or diagrams, cohesion, coupling), but that could be done in three hours. We also wanted to find tasks that could be used across disciplines and across the undergraduate curriculum. Finally, we wanted students to concentrate on the design phases after the requirements are completed through the final design, but before implementation.

For mechanical and electrical engineering students, we first tried a task in which students had to develop a windshield wiper that automatically turned on when rain started. This task worked well for students who had no prior knowledge, but, once the nature of the task leaked out, some students looked up the solution, which was readily available on the Internet, and thus invalidated the assessment. We then turned to a task in which students had to design a solar power panel that would turn to catch the sun. This proved to be too difficult for students. We have developed a third task, which is an embedded control system to improve LCD advertising billboards by changing the backlighting depending on conditions.

For computer science and computer engineering courses we developed three design tasks. First was a library bibliography task, second was a date e-minder system, and third was a polynomial grapher. In each task students were given the requirements and asked to create a design. The first two were not only too complex, but they could have been solved more easily by database methodology. Because we wished to use the tasks across the curriculum, the students were told not to use a database for the solution. This made the task more difficult for students who knew database design. They had to inhibit their desire to make a database design and often got their teams off track. The third one worked well in that it could be completed in two hours.

We count four major lessons learned in our experimentation with design tasks. First, two hours is as good as three hours. In this context—a design task that is required but not graded—students do not use additional time effectively. Neither their performance nor their process was substantially better when given three hours to do the same task. This is not to say that we encountered a ceiling effect.
Students exhibited very little process and tended to leave out chunks of product. We believe, however, that to improve students’ behaviors requires a more scaffolded approach. Smith, Eppinger, and Gopal [4] found that providing additional information about the design environment improved designers’ speed.

Our second lesson was that it is important to develop tasks the performance of which is not confounded by specific knowledge. One must find tasks that one can expect all students to know the domain, such as how to graph polynomials.

The third lesson is that it is easy to make the task too complex for the context. One must walk a fine line between a task that is so simple that design processes such as feasibility are irrelevant and one that is so complex students can not master it.

The fourth lesson is that it is important to instruct the students to gather their notes and create a final design document. Fifteen minutes before the end of the task the research assistant went to each team and gave them two paper clips and instructed them to put their design notes in one clip and their design artifact plus a one-page description in the second paper clip. Otherwise, it was often impossible to differentiate between discarded or partial solutions on the one hand, and the final solution, on the other.

**DESIGN NOTES**

We wanted to be able to document students’ use of design process while they were completing the design task. We took two steps to make the design process visible. First, we had the students work in teams and videotaped them. Research assistants prompted them to discuss what they were thinking if no one spoke for several minutes. Thus we could conduct verbal protocol analysis. Second, we wanted to use their design notes. Because different students would be writing on different pieces of papers at different times, the research assistants had the teams change pen colors every 15 minutes. Thus, it was possible to reassemble the students’ design notes after the fact and even chart iterations, since second and third passes were notated in a different pen color.

**VIDEOTAPES AND ASSESSING DESIGN PROCESS**

Our original intent had been to use the videotapes and design notes to create protocols that we could submit to coding and analysis. Inspection of the first set of videotapes disabused us of that notion. The students, both sophomores and seniors, said very little that could be assessed as design process. Likewise, their notes contained little that could be considered indications of a consideration of design process.

**TEAM VS. INDIVIDUAL ASSESSMENT**

Initially our major assessment measures on the design task were team-based—the final product and the design notes. The videotapes revealed, that, as is often the case, some team members said very little. We had no rules about who should write notes, so team members were free not to contribute. This necessarily results in a skewed evaluation of learning. Team performance may reflect one person, a subset of the team, or all members. Thus, if one applies the 20% rule of continuous improvement and determines, for example, that the class adequately addresses issues of feasibility, one may be in actuality assessing only one-quarter of the class (if teams have four members each).

We thus introduced into the design task two individual assessments. First, we had every team member write a one-page description of the final design. Second, we had everyone complete a design process postmortem, to be described later in the paper.

In the second year of the grant (2003-2004), having decided that the videotapes were not worth analyzing, we tried running the whole design task individually. Students were given packets and initial instructions. Every 15 minutes, the assistant gave them different colored pens, so that the time flow of their design notes could be analyzed. This procedure was successful and can be done in a course-based improvement context.

**RUBRIC FOR DESIGN PERFORMANCE**

The scoring rubric for the product developed in the design task went through four iterations. Each version was the result of the joint energies of our team, which included an electrical engineer, a computer engineer, a computer scientist, and a psychologist. We developed a prototype, tried it with either sophomore or senior students, and then revised it. In the fourth version, reproduced in Table 1, the first four criteria were used by all and the last criterion, coupling, only for software designs. The scores can be used in two ways. First, performance on each criterion can be evaluated for continuous improvement, for example, identifying those criteria for which more than 20% of the class got scores of 1 or 2. Second, scores on each criterion in Table 1 can be summed for an overall score, which can be used to grade students.

Because teams who had no members with prior knowledge about the solution of the tasks were at a disadvantage, we added a criterion so that faculty could judge the degree of knowledge and those designs that were influenced by prior knowledge could be assessed separately.

Our participating faculty disagreed about whether to include design notes or to assess solely on the basis of the final artifact and descriptions. Those who used both felt that the design notes sometimes clarified the final designs. Those who used just the final design wanted to assess the quality of the product separately from the quality of the process, because the two are separable [2], [5]-[7].

When we instituted an individual one-page description, we added a rubric to evaluate them (see Table 3). Having two criteria allows faculty to separate the issues of whether students understand the design from their communication skills.
Collecting Additional Design Process Data

Because neither the videotapes nor the design notes yielded enough information to judge design process, we tried three different alternative strategies of collecting additional data about students’ understanding of design process during the task.

Postmortem to Videotapes

Shortly after they completed their design tasks, 50 students in a software design task were asked to review their videotapes and check which of 27 activities they did during each of the six time periods marked by each pen color (see Table 4 for examples of those activities).
the self-assessment yields a percent of students who claimed to engage in that activity for each of six 15-minute periods. To give a flavor of the kind of data this provided, Table 4 displays the percentage of students who checked each of five of the activities for each of the periods.

Initial examination of the results in Table 4 suggest that the students are engaging in all these design process activities, because more than 80% of students checked each of these in at least one period, which passes the standard cutoff for continuous improvement. Nonetheless, examination of the patterns suggests some fuzziness in their thinking. For example, 36% claim to be revisiting their earlier design decisions in the second segment and 42% claimed they were still comparing alternative solutions in the second to last time segment. In fact, 8% of the students claimed to analyze feasibility in every segment and 6% considered alternative solutions in every segment. There appears to be a fair degree of confusion about the precise meaning of these various constructs.

This impression is strengthened by examination of Table 5. Here the students on average claimed to have spent less than 8% of their time examining feasibility and 10% of their time reaching consensus on a single design.

### Table 5
**Students’ Assignment of Proportion Time Spend in Following Activities**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understanding the needs of customers and users</td>
<td>20.8</td>
<td>10.4</td>
</tr>
<tr>
<td>Establishing the program requirements</td>
<td>13.1</td>
<td>8.0</td>
</tr>
<tr>
<td>Generating alternative ideas</td>
<td>10.0</td>
<td>8.3</td>
</tr>
<tr>
<td>Evaluating feasibility of alternative ideas</td>
<td>7.7</td>
<td>4.9</td>
</tr>
<tr>
<td>Evaluating tradeoffs of alternative ideas</td>
<td>6.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Reaching consensus on one design</td>
<td>10.2</td>
<td>5.9</td>
</tr>
<tr>
<td>Creating and testing prototypes</td>
<td>5.2</td>
<td>6.7</td>
</tr>
<tr>
<td>Generating components</td>
<td>13.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Checking design against requirements</td>
<td>11.2</td>
<td>7.6</td>
</tr>
<tr>
<td>Testing the design</td>
<td>1.8</td>
<td>3.0</td>
</tr>
</tbody>
</table>

We felt that this procedure might have substantial merit if given repeatedly within a continuous improvement process, because students could have refined their answers. It is not, however, practical in the typical classroom, because of the time and expense of videotaping.

**Postmortem to Design Task**

The second strategy we used to get more data about students’ design process was to institute a postmortem immediately after the students completed their design task. The questions asked and the rubric designed to assess the students’ responses are listed in the left column of Table 6. We also asked students to describe their prior knowledge concerning the task. Table 7 presents the rubric for prior knowledge.

### Table 6
**Rubric for Design Process Postmortem 2**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Distinguished (5)</th>
<th>Proficient (3)</th>
<th>Emerging (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>What design alternative(s) did you consider?</td>
<td>At least two substantive designs are described.</td>
<td>More than one design is described, but at least one of those designs is inadequate or poorly specified.</td>
<td>Only one design is described.</td>
</tr>
<tr>
<td>What trade-offs did you think about?</td>
<td>Description includes at least two trade-offs that truly differentiate solutions.</td>
<td>At least two trade-offs are mentioned but no more than one clearly differentiates between solutions.</td>
<td>Student’s description of trade-offs does not tie it (them) to the solutions considered.</td>
</tr>
<tr>
<td>How did you address feasibility?</td>
<td>Consideration of at least two issues (e. g., cost of parts, reliability, ease of manufacture) tied directly to solution(s).</td>
<td>At least one feasibility issue tied to solution or a good list of issues not directly tied to solution.</td>
<td>Description of feasibility issues vague and scattered.</td>
</tr>
<tr>
<td>How did you check your design? (against requirements)</td>
<td>Description of systematic check of design against requirements.</td>
<td>Mentions requirements, but does not describe going back to list.</td>
<td>No mention of requirements.</td>
</tr>
<tr>
<td>How did you check your design? (prototype)</td>
<td>Description of systematic testing procedures such as prototype, scenarios etc.</td>
<td>Describes consideration of accuracy and completeness, but does not describe procedure.</td>
<td>No clear description of checking for accuracy and completeness.</td>
</tr>
<tr>
<td>What kinds of sketches/drawings did you make and why did you make them?</td>
<td>Describes drawing as analytic tool, including description of decomposition or checking interrelations of components.</td>
<td>Describes drawing as way of specification, but not in a way that makes it clear that this is a tool for abstraction/decomposition.</td>
<td>Either no drawings or very vague description of drawings.</td>
</tr>
</tbody>
</table>

### Table 7
**Rubric for Prior Knowledge Postmortem Question**

<table>
<thead>
<tr>
<th>Specific Knowledge (Sp)</th>
<th>Generic Knowledge (Gn)</th>
<th>No prior experiences (NK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>What prior experiences have you had that helped you create this design?</td>
<td>Prior knowledge of particular solution (e. g., microcomputers or acoustic engineering in windshield wiper task)</td>
<td>Describes prior experiences relevant to design, e. g., took circuits, know C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Either left blank or description irrelevant or too vague to assess its relevance (e. g., took physics).</td>
</tr>
</tbody>
</table>

The rubric was applied to the postmortems of two software engineering classes that had both computer science and computer engineering students. One was given the reminder problem (n=60) and the other the polynomial grapher (n = 50). The rubric proved feasible in the sense that it could be applied, but the performance was very low. In all but two questions the majority of students were given a score of 1 and only on the sketches question on the polynomial grapher did a

---

October 19 – 22, 2005, Indianapolis, IN
sizeable number of students (17) get the score of “distinguished. In both designs virtually all students (60 and 49, respectively) got a 1 on feasibility, suggesting probably that they did not think to address issues of feasibility.

The design postmortem does not assess whether students understand constructs such as feasibility, but just whether they address them while designing. If students do not understand these constructs, they obviously can not answer the postmortem questions. Thus, using this postmortem is most appropriate after students have demonstrated that they understand the constructs themselves. We have developed a Design Process Knowledge Test to assess understanding [12].

Our experience suggests that a two-hour design task can not address all of the constructs assessed in the design postmortem. It appeared from the students’ answers that the reminder task was so complex students had little time to consider issues such as feasibility, trade-offs, and alternative designs whereas the polynomial grapher did not have sufficient substance for these issues. Rather than searching for the perfect task in which all of these issues can be assessed, it might be better to develop part-design assessments, in which, for example, a set of design alternatives are presented and students are asked to decide among them (using, we hope, feasibility and trade-offs in that consideration).

In summary, the scoring of the process postmortem using the rubric is a feasible approach to assessing design process. Although the scores of the software engineering class may have been an underestimate of their ability because the design task was too difficult, it still provides pointers toward improvement. The overall performance of the class indicates, for example, that students either did not appreciate the importance of considering alternatives in light of trade-offs and feasibility or did not know how to go about such comparisons. The instructor needs to figure out activities that will help students learn to do so.

Pretask

The third strategy we used was to assess students’ design process knowledge before and during the design task. The purpose of the pretask was to get the design process activities in the forefront of students’ minds so that they could identify these activities while they completed the design task. When the students arrived to complete the design task, they first were presented an Excel file with the following list of design process activities, in alphabetical order, in Column A.

- Brainstorm
- Check against requirements
- Check cohesion
- Check coupling
- Consider alternatives
- Consider feasibility
- Consider tradeoffs
- Create a sketch
- Create a user scenario
- Find information
- Identify components
- Identify relations among components
- Meet with customer
- Prototype
- Revise and expand problem statement
- Run a scenario
- Test
- Think through hidden functions
- Write requirements

The students were asked to copy and paste the activities that they would use in the order they would use them. After they completed this task, they were asked to insert into their design notes headings corresponding to the activities they chose, which were available for reference throughout the design task.

This procedure was tried in a sophomore electrical engineering system design task of 25 students. Although the students used most of the Column A entries in the pretask the only headings used during the design by more than 5 students were brainstorm \( (n = 12) \), create a sketch \( (n = 6) \), find information \( (n = 8) \), identify components \( (n = 7) \), and write requirements \( (n = 10) \). The average number of headings was 9 in the design notes and 11 in the design artifacts. The sparseness of use of headings should not be construed as evidence that the strategy failed. First, the students did clearly differentiate the activities they saw as related to design. That students did not include headings such as “Consider feasibility” and “Consider tradeoffs” suggests that they were not focused on those aspects of process and can be considered pointers toward improvement. Second, we made no attempt to be sure that they had complied with the instruction to annotate their design notes with headings. It would be a simple matter to check at the end whether headings had been used and to ask students to make sure that they had included headings for all their notes. The activities that were never used as headings could be a focus for continuous improvement.

CONCLUSIONS

Students’ Understanding of Design Process

This project has demonstrated that students are not aware of their own design processes, a phenomenon that has been found in design [8] as well as in learning physics and other cognitive tasks [9-11]. Smith and Tjandra [8] videotaped nine groups of industrial engineering students enrolled in a senior design course creating a design. There was very little discussion about process and very little process apparent. This was in contrast to an earlier study with practicing engineers who spent significantly more effort discussing process. Our seniors in design courses likewise did not spontaneously address design process issues when asked to design. Even when asked explicitly, they demonstrated little awareness of design process.

Assessment Tools

When we began this project, we conceptualized videotaping students completing a design task as a criterion by which we would gather validity data for our other assessments. This was
infeasible because students exhibited so little design process in either their verbalizations or their design notes.

What was our psychometric loss, however, became a gain, because we were able to develop a set of alternative assessment techniques that can be given in the context of a design task and that provide pointers for continuous improvement.

The review of videotapes, although feasible only in classrooms that can manage to create the videotapes, has the most face validity, because it would appear that the videotapes would keep the students honest. Comparable use in verbal protocol analysis in research studies has shown that additional and useful information is gathered when participants review their videotapes immediately after doing a task. Nonetheless, it appears that the task had such strong demand characteristics that the students rather indiscriminately claimed to be engaging in a variety of design activities even though those claims were inconsistent with their final claims of the proportion of time spent in various design activities. Nonetheless, this technique is the only one of our alternatives that permits the instructor to discover whether students are considering particular aspects of the design at specific times during the process.

The postmortem task is a more feasible task. It can be given after any design task instructors choose, although the characteristics of the design task clearly influence the quality of answers. Because it is done after the design task, it can not influence how students approach the design task. Thus, it would be a good choice to use at the end of a semester or as a pre-post test as an indication of how much students have learned. It can also be used as a learning tool during a semester. If a series of design tasks with postmortems are given during a term, instructors and/or students can evaluate their strengths and weaknesses by analyzing the postmortem questions and using them to find pointers toward improvement.

The pretask/heading strategy would be effective when instructors wish to encourage improvement within a task and are not concerned with the influence of the assessment on performance. The instructor could first give the pretest, then discuss the activities she or he wishes the students to focus on during the design process, then have the students add the headings to their design notes at the end of the task. This would serve as a quick self-check to the students and the class could have a discussion on the spot of when and how to engage more effectively in the targeted activities.

Thus, the assessments we developed provide several different avenues by which to assess students’ strengths and weaknesses in the application of design process. Their use can make students aware of design process and provide a basis for discussions of how to improve their design processes.

ACKNOWLEDGMENT

This research was made possible by a grant from the Assessment of Student Achievement program of the National Science Foundation (DUE-0127751). The authors wish to acknowledge the insight of Tesfay Meressi in developing the postmortem, the work of Plamen Trichov in developing the pretask and Karen Baumgarner, Samantha Stokes, Rebecca VonderHeide and Madeleine Page in carrying out the research.

REFERENCES