The "Prototype Walkthrough": A Studio-Based Learning Activity for Human-Computer Interaction Courses

Christopher D. Hundhausen
School of Elec. Eng. and Comp. Sci.
Washington State University
Pullman, WA 99164
+1-509-335-4590
hundhaus@wsu.edu

Dana Fairbrother
College of Education
Washington State University
Pullman, WA 99164
+1-509-335-1738
danafairbrother@wsu.edu

Marian Petre
Faculty of Mathematics & Computing
The Open University
Walton Hall, Milton Keynes
+44-1908-65-33-73
m.petre@open.ac.uk

ABSTRACT
For over a century, *studio-based* instruction has served as an effective pedagogical model in architecture and fine arts education. Because of its design orientation, human-computer interaction (HCI) education is an excellent venue for studio-based instruction. In an HCI course, we have been exploring a studio-based learning activity called the *prototype walkthrough*, in which a student project team simulates its evolving user interface prototype while a student audience member acts as a test user. The audience is encouraged to ask questions and provide feedback. We have observed that prototype walkthroughs create excellent conditions for learning about user interface design. In order to better understand the educational value of the activity, we performed a content analysis of a video corpus of 16 prototype walkthroughs held in two undergraduate/graduate HCI courses. We found that the prototype walkthrough discussions were dominated by relevant design issues. Moreover, mirroring the justification behavior of the expert instructor, students justified over 80 percent of their design statements and critiques, with nearly one-quarter of those justifications having a theoretical or empirical basis. These results suggest that prototype walkthroughs can be useful not only in helping to teach HCI design, but also in helping to gauge students' evolving design knowledge.

Categories and Subject Descriptors
K.3.2 [**Computer and Information Science Education**]: *Computer science education, Curriculum.*

General Terms
Design, Experimentation, Human Factors.

Keywords
Studio-based learning and instruction, prototype walkthrough, design crit, HCI, user interface design

1. INTRODUCTION
For over a century, studio-based instruction has served as an effective pedagogical model in architecture and fine arts education. In this model, students iteratively develop solutions to a series of design problems. In periodic "design crits" (design critiques), students present their evolving solutions to their peers and instructors for feedback and discussion.

User interface design is a central skill taught in an upper-division computer science course on human-computer interaction (HCI). In such a course, students often undertake a capstone design project that takes them through all phases of the user-centered design process [15], including initial data gathering, user interface prototyping, and usability testing. Because of its focus on design, a human-computer interaction (HCI) course has been identified as an excellent candidate for studio-based instruction (e.g, [1, 11, 16]).

Within the context of a multi-institutional research project in which we are adapting and refining the studio-based instructional model for computing education [8], we have been exploring a new kind of studio-based learning activity—the *prototype walkthrough (PW)*—in our undergraduate/graduate HCI course at Washington State University. In preparation for PWs, student capstone project teams develop low fidelity user interface prototypes of their evolving project designs, and a set of five core tasks to be completed with their prototypes. In PW sessions lasting approximately 20 minutes each, project teams simulate their low fidelity prototypes in front of the class. A student from the audience serves as the test user by interacting with the prototype, and thinking aloud in the process. At any point, the audience can jump in with questions, comments, or feedback. After the five tasks have been completed, the instructor invites the class to engage in a reflective design discussion intended to help the project team improve its design.

Our exploration of the PW activity raises a pair of basic research questions regarding its educational value as part of an HCI course:

RQ1:	*To what degree does the PW promote discussions that are relevant to user interface design?*

RQ2:	*To what degree do students participate in those discussions?*

Resonant with situated learning theory [12], the PW activity is designed to provide opportunities for varying levels of participation in a community of practice. In particular, it provides experts (HCI instructors) with opportunities to model the application of the design knowledge explored in an HCI course, while providing learners (HCI students) with opportunities to practice applying their emerging design knowledge. In design

ICER'11, August 8–9, 2011, Providence, Rhode Island, USA.

discussions, such design knowledge manifests itself most readily in the ways in which, and extent to which, design critiques and suggestions are justified. This observation leads to an additional research question regarding the educational value of PWs:

RQ3:	*How are design critiques and suggestions justified?*

This paper addresses these questions by presenting a detailed content analysis of a video corpus of 16 PWs, which were run within successive offerings of the conjoint undergraduate/graduate human-computer interaction course at Washington State University. In furnishing the first-ever detailed video analysis of design discussions within an HCI course, our study makes three key contributions to the computing education literature:

1. It introduces the PW as a novel studio-based learning activity for HCI education.
2. It presents a rigorous content coding scheme that can be used to analyze critical discussions about user interface design.
3. It provides a rich descriptive account and analysis of the design discussions promoted by PWs, thus providing evidence of their educational value.

The remainder of this paper is organized as follows. In Section 2, we present the background and related work on which our study builds. Section 3 details the design of our study. Sections 4 and 5 present and discuss the study's key results. Section 6 presents conclusions and discusses future work.

2. BACKGROUND AND RELATED WORK

A form of "design crit" in the *studio-based* instructional model, the PW activity explored in our study engages students in discussions with experts about their user interface designs and how to improve them. A rich legacy of empirical work, nicely synthesized by Cross [3], has explored the behaviors, activities, and processes of both novice and expert designers. In a similar vein, the literature on computer-supported collaborative learning is replete with content analyses of discussions that take place during learning activities, with a focus on how representations serve to mediate those discussions (e.g., [19]).

Within computing education, we have previously performed such analyses of "design crits" anchored in visual representations of algorithms in both an upper-division algorithms course [5] and a CS 1 course [6]. The study presented here contributes to all of these lines of work by performing the first detailed content analysis of critical discussions about user interface design within a course on human-computer interaction design.

Kehoe [11] calls the kind of "design crit" on which our study focuses *critical design dialog*, and points out that it differs from other forms of learning discussions in that it is directed toward critiquing students' work in a public forum, with the dual-aim of (a) influencing the trajectory of the work, and (b) providing opportunities for students to learn from each other's design work and feedback. Kehoe [11] (see also [16]) makes a strong case for the educational value of critical design dialog as a means of learning about HCI design. In brief, she argues that the kinds of design problems that are common in HCI are fuzzy and have no clear-cut solutions. Design principles and heuristics that might guide one to solutions are necessarily vague; learners often find them to be unclear and overly ambiguous [18], leading to their getting stuck during the design process [17]. Learners, she argues, can therefore best develop design competence when they (a) receive feedback on their own designs that is also connected to more general design principles and heuristics, and (b) observe how experts think about design. Critical design dialog provides ideal conditions for both.

In addition to Kehoe's arguments in favor of critical design dialog as a valuable HCI learning activity, the activity has a strong foundation in situated learning theory [12]. According to this theory, one gains competence within a community of practice by having opportunities to participate, in increasingly central ways, in the practices of the community. Critical design dialog, as manifested in the PW, provides such multi-faceted opportunities for participation. In PWs, students can observe expert critiques of design, remaining on the periphery of the discussions as audience members. As they become more comfortable, they can gradually explore opportunities to offer their own critiques and suggestions. As design team members, students are placed in the position of presenting, justifying, and defending their own designs. This constitutes more central participation in design practice; it is akin to the situation of an expert designer at a real-world company.

Computing educators have explored the use of studio-based learning in individual computing courses (e.g., [10, 13]) and even in entire degree programs (e.g., [4]). In one of the few published studies of studio-based learning in HCI education, Reimer and Douglas [16] describe their implementation of an undergraduate HCI course based on the studio model. The course included weekly design crits that were similar in spirit to the PWs described here. The key difference was that, in the design crits, the design teams themselves simulated their user interfaces for demonstrational purposes, rather than enlisting a student audience member as a test user. While Reimer and Douglas did not present a video analysis of their design crits, their observation that the design crits "fostered a highly interactive and constructive learning atmosphere" ([16], p. 201) well resonates with the findings presented here.

In a similar vein, Cennamo et al. [2] performed a detailed qualitative comparison of design studios in both industrial design and human-computer interaction courses, gleaning insights into how these studios promoted the generation and analysis of design ideas. Likewise, Arvola and Artman [1] compared HCI students' studio work in a traditional space against that in a computer-augmented space. While their study focused on studio activities that were far less structured than the PWs we studied, it is similar to our study in that it extensively analyzed video footage.

3. EMPIRICAL STUDY

We conducted our empirical study in conjunction with the spring 2007 and spring 2008 offerings of CptS 443/543 ("Human-Computer Interaction"), the conjoint undergraduate/graduate HCI course at Washington State University taught by the first author. Using a mix of lecture and small group activities, and a pair of textbooks [14, 15], the course explored the application of relevant theories, principles, and processes to the design of interactive software. A focal point of the course was a capstone user interface design project, which students were required to complete in teams of two to three. Student teams could choose the focus of their projects, or they could take on a project suggested by the instructor. During the tenth week of the 15-week semester, project teams presented prototypes of their evolving designs to the class within PW sessions scheduled during regular course lecture periods. These were the focus of this study, which we describe in further detail below.

3.1 Participants and their Design Projects

The spring 2007 course offering enrolled 13 upper-division undergraduate students and two graduate students, while the spring 2008 course offering enrolled 13 upper-division undergraduate and ten graduate students. All but four of these students were computer science or computer engineering majors who had minimally completed a sequence of core courses in software design. The other four students came from a mix of majors, including geology and management information systems. None had taken a prior course in HCI.

Our study considered the PWs of all seven project teams in the 2007 course offering, and nine of the 11 project teams in the 2008 course offering (two were not recorded because of technical difficulties with the video equipment). Table 1 presents the key attributes of the 16 project teams whose PWs were considered in the study. As can be seen, the projects on which they focused were diverse. Moreover, whereas project teams constructed their prototypes mostly out of simple art supplies (pen, paper, transparencies) in the 2007 course offering, most project teams in the 2008 course offering constructed their prototypes using WOZ Pro [9], a computer-based low fidelity prototyping tool we have developed specifically for this purpose.

3.2 Prototype Walkthrough Procedure

Prior to participating in the PWs, project teams were required (a) to perform at least two early data gathering activities (e.g., interviews, questionnaires, field observation) in order to establish the functional, usability, and user experience requirements for their project, (b) to develop a low fidelity user interface prototype based on those requirements; and (c) to formulate a set of five core tasks that their prototype had to support. Project teams brought the prototype and set of tasks to the PW sessions, which took place in a small classroom during the two 75-minute lecture periods of the tenth week of the fifteenth week semester.

All students were required to attend and participate in the PWs. Each project team was assigned a 20-minute slot for their walkthrough; students whose team was not immediately presenting were required to observe the walkthroughs, and to fill out a structured evaluation form intended to provide feedback for the presenting project team. Each PW began with the instructor calling a project team to the front of the room. The project team randomly chose a member of the audience to serve as the "test user" for the PW. The team proceeded to provide a brief description of the prototype interface they had designed, along with a general task scenario. At this point, the project team handed the test user a written set of task instructions, and projected their prototype onto a large screen at the front of the room. Depending upon the prototyping technology used, either an overhead projector or LCD projector was used for this purpose.

The test user proceeded to complete the set of tasks as the project team simulated their prototype's user interface. The test user was instructed to read each task aloud prior to performing it, and to think aloud while performing each task. In order to perform tasks, the test user directly pointed at and manipulated elements of the image projected on the large screen, describing his or her actions along the way (see Figure 1). Audience members and the instructor were welcome to interrupt the walkthrough at any time with questions or comments. The walkthrough ended when the test user completed all five tasks, or the 20-minute time limit had been reached, at which point the instructor initiated a round of applause for the presenting project team and called on the next scheduled project team.

Table 1. Key Attributes of the project teams studied

Team	Size	Project Focus	Prototyping Technology Used
Sp07-1	3	Distributed team problem management	Art supplies
Sp07-2	2	Smart home event scheduler	Art supplies
Sp07-3	2	Campus map route finder	HTML/Javascript
Sp07-4	2	Personal travel blog site	Art supplies
Sp07-5	2	N-body simulator	Art supplies
Sp07-6	2	DVR Remote Control	Art supplies
Sp07-7	2	Power utility mapping software	Art supplies
Sp08-1	2	Smart home control system	WOZ
Sp08-2	2	Low Fidelity UI Prototyping Tool	WOZ
Sp08-3	2	Custom grid-based game builder	HTML
Sp08-4	2	Poker Game	Power Point
Sp08-5	3	Online code review environment	WOZ
Sp08-6	2	Campus map route finder	Power Point
Sp08-7	3	Recipe management software	WOZ
Sp08-8	2	Custom Game Builder	WOZ
Sp08-9	2	Group Collaboration Tool	HTML/Javascript

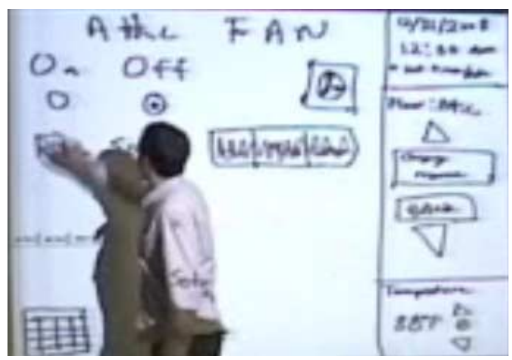

Figure 1. A test user interacting with a prototype within a prototype user interface within a PW

3.3 Data Collection and Analysis Method

Using a video camera positioned near the middle of the classroom and focused on the projected screen, we obtained 4.91 hours of high-quality video footage of the 16 PWs. In order to analyze the content of the talk that took place within our video corpus (RQ1), we began by partitioning the talk into *segments*, where a segment was defined as a single thought or idea uttered by a single participant. We then iteratively developed the coding schemes described below by watching a subset of the walkthrough sessions and adding and refining categories until no new ones emerged. As we did this, we composed a coding manual with detailed categorical descriptions, rich examples of how to distinguish among categories, and step-by-step instructions for coding. Those interested in using or adapting our coding schemes should consult this manual, which we have made available online [7].

Table 2 presents and briefly describes the nine top-level categories in our content coding scheme. Because of its perceived relevance to the HCI course, Design Talk was of particular interest in this study. Table 3 presents a more detailed look at Design Talk in terms of its six subcategories. While they are intended to provide an overall feel for the categories, we emphasize that the descriptions provided in these tables are necessarily terse, and lack sufficient detail and examples for one to make reliable distinctions. We refer interested readers to the coding manual cited earlier [7] for more detailed descriptions.

We also note that the categories in these tables are listed in order of decreasing priority. In cases in which, despite our detailed categorical definitions, we felt a given segment could be coded into multiple categories, we always coded the segment into the category with the *highest* priority.

In order to gauge the extent to which students and the instructor participated in PW discussions (RQ2), we additionally classified each segment according to role of the participant who uttered it:

- *instructor*—the first author of this paper, an HCI expert with two years of industrial experience who taught the course and moderated each PW;
- *design team member*—a member of the two or three-person student team whose prototype was being tested;
- *audience member*—a member of the student audience;
- *test user*—the student who acted as the test user; and
- *class*—at least two speakers in any of the previous speaker categories (reserved only for segments coded as *Laugh*).

Table 2. Top-level content coding categories

Category	Description	Example
Design Talk	Talk focused on design, including justifications, critiques, suggestions, issues, and strategies	See Table 3.
User Interface Talk	Talk focused on the functionality and appearance of the user interface being tested.	"There's a button at the bottom of the screen." "Show me how this works."
Task Description Talk	Talk focused on the task being performed in the PW	"Was I supposed to do Task 1?"
Task Execution Talk	Talk focused on what the test user is doing or thinking as s/he performs tasks	"I'm clicking here, and I want to change the date."
Activity Talk	Talk directed toward running the PW activity	"It's your turn now." "Any questions?"
Project Talk	Talk focused on the scope or focus of the team's project	"Our project is building a smart home interface."
Took Talk	Talk focused on the prototyping technology being used in the PW	"These sticky notes are awkward." "Go to "Run" mode in WOZ Pro."
Laugh	Laughter uttered by at least two people	[Laughter]
Off Task Talk	Talk unrelated to the PW activity	"My job interview went well!"

Recall that RQ3 focuses on exploring justifications of design critiques and suggestions. To that end, we developed a scheme for classifying design justification statements according to the *basis* of the justification. Table 4 describes the twelve justification basis categories in this scheme. These categories are listed approximately from strongest to weakest, based upon our perception of what an HCI expert would take to be a good justification. The top four categories are rooted in either established principles (e.g., those described by Norman in [14]) or empirical evidence. Categories that appear further down the table have more to do with personal experience, intuition, or practical concerns. The last category in the table accounts for justifications with no apparent basis.

In order to verify the reliability of our coding schemes, the first and second authors independently coded a 20 percent sample of the video corpus with respect to both the content and the justification basis schemes. We attained a level of agreement of 84 percent (0.82 kappa). Having reached a high level of inter-rater reliability, we had the second author code the remainder of the video corpus.

4. Results

Table 5 presents key summary data on the 16 PWs in our corpus. On average, a PW session lasted 18.4 minutes (SD = 7.4), and contained 176.9 coded segments (SD = 61.1), including 54.2 design talk segments (SD = 36.8) and 16.5 justification segments (SD = 11.6). Not surprisingly, session length was strongly correlated with the number of segments in the session ($r = 0.687$, $p = 0.003$). Interestingly, session length was also strongly correlated with the number design talk segments in the session ($r = 0.647$, $p = 0.007$) and the number of justification statements in the session ($r = 0.698$, $p = 0.003$).

Table 3. Design Talk subcategories

Category	Description	Example
Justification	Justifies statements about design (critiques and suggestions) or the design of the interface under test	See Table 4.
Critique	Makes a statement about the goodness of the design	"I don't think that design will work."
Suggestion	Suggests an alternative design	"I think you should re-label the button."
Issues and Strategies	Discusses design issues, assumptions, strategies for arriving at new designs, and tradeoffs among design alternatives	"You might need to iterate again on this."
Meta-talk	Design talk that transcends the specific design under consideration, including comparisons with other designs and general conceptual design issues	"It's a question of how to do good layout." "The issue really is, what is a screen link?"
Encouragement	Congratulates the designers or encourages them to continue their work	"You'll get there." "Keep at it."
Response	Responds to design talk or expresses understanding of design.	"Good point." "That's a tough call." "I see where you're coming from."

Table 4. Justification basis subcategories

Category	Description	Example
Design principle	Appeals to, or implicitly enlists, an established design principle.	"That's a poor natural mapping."
Test user behavior	Based on what the test user actually did, thought, or expected during the prototype walkthrough.	"The user stumbled when he saw it."
Past user behavior	Based on the behavior of a user in a past user study (out prior to the PW).	"Users had trouble with this in a previous study."
Other computer software	Appeals to the design of other similar computer software.	"I think the way Photoshop does it is better."
Limitations of prototyping technology	Based on perceived limitations of the prototyping software used for the walkthrough.	"Art supplies made it difficult to create polished buttons."
Implementation Difficulty	Based on the perceived difficulty of implementing the design in a given way.	"Ideally, you could freeform draw it, but that would be hard to implement."
Limitations of PW Activity	Based on limitations of the PW activity, including the tasks and their ordering.	"The fact that he did these tasks in a certain order gave him an advantage."
Personal Experience	Based on the speaker's personal experience.	"When I've done this in the past, I've always had trouble with making tables."
Hypothetical user	Appeals to what a hypothetical user might do or think in given set of circumstances	"A user wouldn't know how to interpret that."
Logic or common sense	Based on logic, reason, or common sense, or the need to accommodate "real world" scenarios.	"The label needs to be changed because the current label doesn't make sense."
No basis	Justification that has no apparent basis	"It's the best we could come up with."

In this section, we treat the *individual PW session* as the unit of analysis. Hence, the percentages we present and analyze reflect the *mean* percentages of categorized talk across the 16 PW sessions, not the *overall* percentages of categorized talk in the 16 PW sessions combined. Analyzing the data in this way gives equal weight to each PW session, rather than weighting each session by its length.

4.1 PW Content and Contributions

We first explore our data relevant to RQ1 and RQ2. Figure 2 presents the mean percentage of talk dedicated to each of the high-level content categories within a PW session. Within each category, the talk is broken down further by participant type. As Figure 2 indicates, three categories of talk dominated the PW discussions:

- Design Talk (*M* = 27.8%, *SD* = 13.5%), which focused on actual user interface design issues;
- User Interface Talk (*M* = 24.7%, *SD* = 11.6%), which focused on helping PW participants better understand the user interface being tested; and

Table 5. Summary data on PW sessions

Session	Dur. (Min)	Total Segments	Design Talk Segments	Justification Segments
Sp07-1	21.5	143	58	12
Sp07-2	24.0	160	34	5
Sp07-3	8.1	100	3	0
Sp07-4	15.6	101	32	12
Sp07-5	34.1	216	85	39
Sp07-6	24.1	195	77	36
Sp07-7	13.4	126	7	4
Sp08-1	8.8	100	24	13
Sp08-2	18.0	241	102	34
Sp08-3	19.1	190	44	20
Sp08-4	17.5	236	108	23
Sp08-5	27.5	259	75	23
Sp08-6	21.5	238	65	30
Sp08-7	22.4	270	116	32
Sp08-8	8.5	108	18	7
Sp08-9	11.0	148	19	8

- Task Execution Talk (*M* = 23.0%, *SD* = 10.1%)—the test user's think aloud protocol, which provided a basis for evaluating the strengths and weaknesses of the user interface being evaluated.

Inspection of Figure 2 suggests that each participant type contributed in different quantities to the PW discussions. Figure 3 brings this into sharper focus by presenting the mean percent contribution of each participant type. As Figure 3 illustrates, members of the design team who were simulating their interfaces contributed roughly one-third of the discussion content—the most of any participant type. Not far behind were the test user, who thought aloud while completing tasks with the design team's prototype interface, and the course instructor, who facilitated the PW sessions; both contributed roughly one-quarter of the discussion content on average. Audience members were not as extensively involved, contributing 10 percent of the talk. The

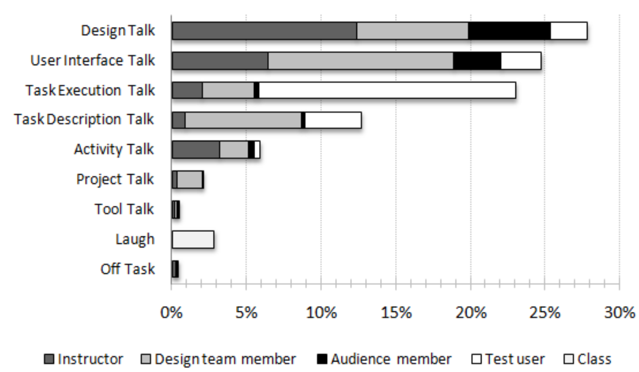

Figure 2. Mean PW session content classified by content category (see Table 2) and participant type. Note that the "Class" participant type applies only to "Laugh" content.

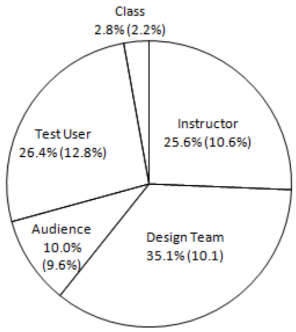

Figure 3. Mean contribution of participant types to all talk (standard deviations in parentheses)

"class" speaker type, used only in conjunction with Laugh segments, contributed just under three percent, reflecting the fact that, on average, roughly three percent of PW discussion content consisted of laughter.

In examining Figure 2, one also sees that each participant type contributed different types of talk to the PW sessions. According to a chi-squared test of homogeneity, the distribution of talk across our high-level content categories varied significantly by participant type, $\chi^2(18, N = 2741) = 1370.0$, $p = < 0.0001$.[1]

Figure 4 takes a closer look at Design Talk, breaking it down both by the subcategories described in Table 3, and by participant type. As can be seen, roughly one-third of Design Talk statements consisted of critiques of the user interfaces being presented in the PWs (*M* = 12.7%, *SD* = 23.8%), or suggestions for improvement (*M* = 20.5%, *SD* = 9.7%). Roughly another third of Design talk statements either justified those critiques and suggestions, or justified the design of the user interfaces being considered in the PWs (*M* = 35.1%, *SD* = 15.3%). The remaining third of Design Talk was dominated by discussion of issues and strategies, and direct responses to other Design Talk.

Figure 4 suggests that participant types contributed in different quantities to Design Talk. Figure 5 illuminates these differences by presenting the mean contribution of each participant type. As Figure 5 shows, over 40 percent of Design Talk statements came

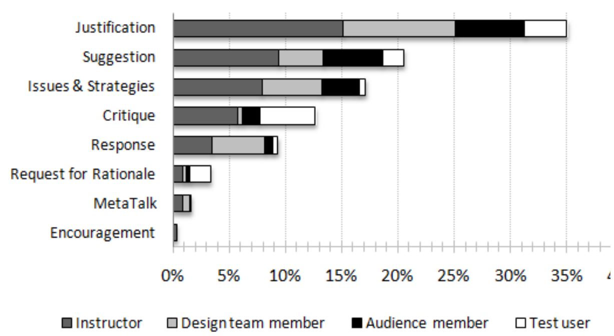

Figure 4. Design talk content classified by Design Talk subcategory (see Table 3) and participant type

[1] Because chi-squared tests of homogeneity test categorical frequencies, they cannot be applied to session means. Hence, this test was applied to the corpus as a whole. When we performed chi-squared tests on each of the 16 PW sessions individually, we obtained similar statistically significant results.

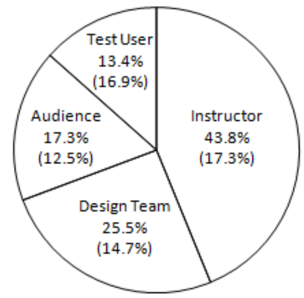

Figure 5. Mean contribution of participant types to Design Talk (standard deviations in parentheses)

from the instructor, with design team members contributing roughly one quarter of the statements, and test users and audience members each contributing less than one-fifth of the statements. Interestingly, the four participant types' contributions differ from their contributions to overall talk: Whereas the instructor and audience members contributed a *greater* percentage to Design Talk than to overall talk, design team members and the test user contributed a *smaller* percentage.

Figure 4 also indicates that participant types contributed different types of Design talk. A chi-squared test of homogeneity confirms that the distribution of talk across Design Talk subcategories differed significantly by participant type, $\chi^2(12, N = 871) = 68.3$, $p < 0.0001$. This is consistent with the findings for overall talk, and reflects the differing roles that participants played in the PW activity.

4.2 How Design Statements Were Justified

We now shift to an exploration of data relevant to RQ3. On average, 9.7% (*SD* = 5.2%) of the segments of each PW session were coded into the Design Justification category. Figure 6 breaks these segments down according to the taxonomy of justification bases presented in Table 4, For each justification basis, a stacked bar additionally indicates the contribution of each participant type.

As the chart indicates, an average of 30.0% (*SD* = 18.6%) of justifications were rooted in either empirical evidence (test user behavior, past user behavior), or the design principles taught in the course. Of the remaining justifications, appeals to common sense (M = 22.9%, SD = 12.5%), a hypothetical user (M = 20.6% SD = 17.2%), other software (M = 8.0%, SD = 7.4%), and personal experience (M = 5.2%, SD = 7.0%) were most common. Practical concerns, including perceived difficulties in implementing a given design (M = 3.59%, SD = 4.1%), limitations of the prototyping technology 3.3%, SD = 5.5%), and limitations of the PW activity itself (M = 2.9%, SD = 5.9%), were less common. Just 3.6% (SD = 4.7%) of justifications had no basis whatsoever.

We believe empirical evidence and design principles form the strongest basis for critiques and suggestions regarding user interface design. These are the "good" kinds of justifications that HCI instructors would like to model, and that HCI students would ideally learn to enlist within an HCI course. Given this, we wondered whether the instructor (an HCI expert with two years of industrial experience) enlisted significantly more "good" justifications than the students. To explore this, we pooled (a) design principle, test user behavior, and past user behavior into one category ("good" justifications), and (b) the audience, test

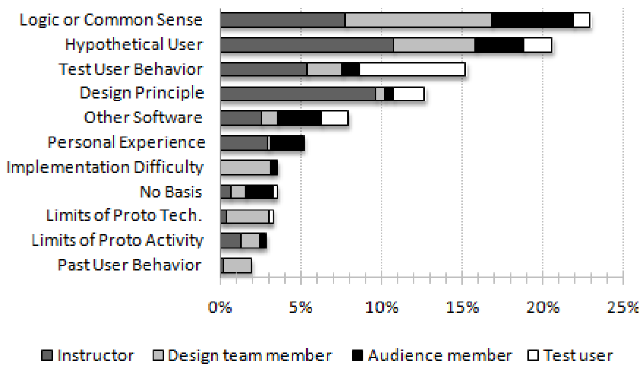

Figure 6. Justifications classified by basis (see Table 4) and participant type

user, and design team into one category ("all students"). After partitioning our data in this way, we found that, on average, 32.5% (*SD* = 30.0%) of the instructor's justifications, and 24.7% (*SD* = 26.8%) of all students' justifications were "good." According to a non-parametric Kruskal-Wallis test, the difference was not statistically significant, ($df = 1$, $H = 0.25$, $p = 0.62$). In other words, we could find no distinguishable difference between the instructor and students with respect to the goodness of their design justifications.

5. DISCUSSION

The results just presented provide a rich descriptive account of the PW activity, including the content of the discussions it promoted, the degree to which people participated in those discussions, and the strategies people used to justify design. In light of these results, we now reconsider the three research questions we posed for this study.

5.1 Relevance of PW Discussions

In order for the PW activity to be a valuable learning activity in an HCI course, it needs to promote discussions that are *relevant* to the course. Accordingly, our first research question focused on the degree to which the PW activity promoted discussions that are relevant to user interface design issues.

Our results provide strong evidence that the discussions were, in fact, dominated by relevant issues. Indeed, on average, Design Talk consumed nearly 28 percent of PW discussions. In addition, User Interface Talk, which is arguably also highly relevant to user interface design because it considers the functionality and presentation of the user interfaces under test, constituted nearly 25 percent of PW discussions. Taken together, Design Talk and User Interface Talk constituted over half of PW discussions on average. Nearly all of the other talk, while not directly related to user interface design, was at least related to the PW activity. Less than one percent of PW discussion talk was off-task.

5.2 Student Participation in PW Discussions

Given our theoretical framework (situated learning theory [12]), which holds that learning takes place through increasingly central participation in community practices, our second research question considered the degree to which students participated in the PW activity. Our results provide solid evidence not only that students participated extensively in the PW activity, but also that the activity promoted levels of participation that differed both quantitatively and qualitatively according to the role that students played in the activity.

At the periphery of the PW activity were student audience members, who contributed the least to the discussions (10 percent). In this role, students mainly observed the activity; however, when they did contribute, their contributions were most likely to be on the topics that were most relevant to the course: Design Talk and User Interface Talk. We speculate that, in their roles as somewhat detached observers, audience members were in a good position to focus and reflect intently on user interface and design issues, without being distracted by the procedural details of the activity.

More centrally involved in the PW activity were the test users, who completed tasks with the interface. One of the key skills to be developed in an HCI course, especially one that consists mainly of computer scientists (as was the case in the courses we studied), is the ability to step away from one's interest in technology development, and into the shoes of users of the technology [14]. The PW activity provided students with valuable opportunities to do just that. Test users were actively involved in the activity, contributing about one quarter of the overall talk task. Owing to the nature of the role, most of test users' contributions were Task Description and Task Execution segments, although they also contributed modestly to Design Talk and User Interface Talk.

Most centrally involved in the PW activity were design team members, who were charged with describing tasks, simulating their user interface, and ultimately explaining and defending their designs. In this role, students had valuable opportunities to engage in two authentic practices of the software industry. First, they got a taste of what it might be like to run a low fidelity prototype test—an important early evaluation activity. Second, they got a taste of what it might be like to present a preliminary design to a software team with an especially critical eye.

Because they were responsible both for describing the tasks to be performed, and for simulating their interface for those tasks, design team members contributed more Task Description Talk and User Interface Talk than any other participant role. They also contributed the most Project Talk, which focused on their overall interface design projects, including its background and history.

5.3 Approaches to Justifying Design

We believe a hallmark of HCI expertise is the ability to make statements about design that are firmly grounded in empirical evidence, established theories, and established design principles. Accordingly, our third research question focused on the ways in which design statements were justified. We found that, on average, just 3.6 percent of design statements had no justification, whereas 30 percent of design statements were rooted in empirical evidence or an established theory or principle. We might have hoped that more than 30 percent of design statements would be grounded in evidence or theory; however, there were no significant differences detected between students (*M* = 24.7%, *SD* = 26.8%) and the instructor, an HCI expert (*M* = 32.5%, *SD* = 30.0%). Moreover, without data from similar empirical studies of design discussions in the software industry, we have no way of knowing whether our results differ from those in the software industry.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced the PW, a studio-based learning activity for HCI education. We have presented a detailed content analysis of 16 PWs that took place in a conjoint undergraduate/graduate HCI course. For HCI educators considering the use of the PW in their own courses, our study furnishes at least three key pieces of empirical evidence:

- PWs promote pedagogically-relevant discussions and active student participation.
- PWs provide students with opportunities to apply their emerging HCI design knowledge by grounding their design statements in empirical evidence and established theories and principles.
- PWs provide opportunities for increasingly central participation in a community of HCI practice [12].

Our findings must be interpreted with some caution. Although they are based on data from two different classes, the same instructor taught both classes at the same university. Hence, the impact of the instructor and the local university culture of the cannot be overlooked. Moreover, the instructor is also the first author of this paper. However, when he taught the course, he had no knowledge of the study's research questions and coding scheme, which were developed over a year later.

We believe that detailed analyses of in-class *processes* like the ones presented here are an important complement to traditional studies of student *outcomes* in computing education. Not only can they provide evidence of learning *in-context*; they can also give insight into how best to design activities in order to promote learning and engagement

The analyses we have presented represent a "quick tour" of our results. In ongoing work, we are preparing a longer article that presents our results in greater depth. The article will include (a) a more detailed analysis of justifications, with analyses of justification *strength* and a richer account of justification bases than could be provided here; (b) an exploration of session-to-session differences, in order to identify features that promoted productive design discussions; and (c) a qualitative analysis of discussions.

Given that computing instructors have limited class time to accommodate studio-based activities like the PW, we would like to explore, in future work, the possibility of conducting studio activities asynchronously online. To that end, we are developing the Online Studio-Based Learning Environment (OSBLE), a learning management system specifically tailored to support the collaborative critical review of student-constructed artifacts. We plan to use OSBLE as a basis for performing detailed empirical comparisons of face-to-face and asynchronous reviews of user interface designs, computer code, and other key disciplinary artifacts of the computing profession.

7. ACKNOWLEDGMENTS

This research is funded by the National Science Foundation under grant nos. CNS-0721927 and CNS-0939017. Contributions to this work by other members of the research team—Michael Trevisan, N. Hari Narayanan, Dean Hendrix, Martha Crosby, Margaret Ross, and Rita Vick—are gratefully acknowledged. Collaboration between C. Hundhausen and M. Petre was supported by M. Petre's Royal Society Wolfson Research Merit Award.

8. REFERENCES

[1] Arvola, M. and Artman, H. 2008. Studio life: The construction of digital design competence. *Digital Kompetanse*. 3, February (2008), 78-96.

[2] Cennamo, K. et al. 2011. Promoting creativity in the computer science design studio, *Proc. 42nd ACM SIGCSE Symposium*, ACM, New York, 649-654.

[3] Cross, N. 2001. Design cognition: results from protocol and other empirical studies of design activity, *Design Knowing and Learning: Cognition in Design Education*, C.M. Eastman et al., eds., Elsevier Sci., Oxford, 79- 103.

[4] Docherty, M. et al. 2001. An innovative design and studio-based CS degree, *Proc.32nd SIGCSE Symposium*, ACM, New York, 233-237.

[5] Hundhausen, C.D. 2002. Integrating algorithm visualization technology into an undergraduate algorithms course: Ethnographic studies of a social constructivist approach. *Computers & Education*. 39, 3 (2002), 237-260.

[6] Hundhausen, C.D. and Brown, J.L. 2008. Designing, visualizing, and discussing algorithms within a CS 1 studio experience: an empirical study. *Computers & Education*. 50, 1 (2008), 301-326.

[7] Hundhausen, C.D. et al. 2009. Studying Prototype Walkthroughs in an HCI Course: Video Analysis Manual (ver. 20), http://eecs.wsu.edu/~veupl/pub/PW-v20.doc.

[8] Hundhausen, C.D. et al. 2008. Exploring studio-based instructional models for computing education, *Proc. 39th SIGCSE Symposium*, ACM Press, New York, 392-396.

[9] Hundhausen, C.D. et al. 2008. The design and experimental evaluation of a tool to support the construction and wizard-of-oz testing of low fidelity prototypes, *Proc. 2008 IEEE VL/HCC Symposium*, IEEE, Piscataway, NJ, 86-90.

[10] Hundhausen, C. et al. 2010. Does studio-based instruction work in CS 1?: an empirical comparison with a traditional approach, *Proc.41st ACM SIGCSE Symposium* ACM, New York, 500-504.

[11] Kehoe, C.M. 2001. *Supporting critical design dialog*, Unpublished Unpublished Ph.D. Thesis, College of Computing, Georgia Institute of Technology.

[12] Lave, J. and Wenger, E. 1991. *Situated Learning: Legitimate Peripheral Participation*. New York. Cambridge University Press.

[13] Myneni, L. et al. 2008. Studio-based learning in CS2: An experience report, *Proc. 46th ACM Southeast Conference (ACM-SE 2008)*, ACM Press, New York, 253-255.

[14] Norman, D.A. 1990. *The Design of Everyday Things*. New York. Doubleday.

[15] Preece, J. et al. 2002. *Interaction Design: Beyond Human-Computer Interaction*. New York. John Wiley & Sons.

[16] Reimer, Y.J. and Douglas, S.A. 2003. Teaching HCI design with the studio approach. *Computer Science Education*. 13, 3 (2003), 191-205.

[17] Sachs, A. "Stuckness" in the design studio. *Design Studies*. 20, 2, 195-209.

[18] Schön, D. 1987. *Educating the reflective practitioner*. San Francisco. Jossey-Bass Publishers.

[19] Suthers, D. and Hundhausen, C. 2003. An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences*. 12, 2 (2003), 183-219.