3. Discovering SRtID Opportunities in College Lecture Courses

3.1 Introduction

In chapters one and two I argue that design research in education can be more productive by designing, studying and iterating upon systems that operate in real-world, authentic contexts. To design a system to operate successfully in authentic contexts requires a rich understanding of that context. Further, for the system to be adopted requires designing for the factors that drive stakeholders to consider and use the system. Developing this understanding of stakeholders helps to identify opportunities for which to design that will be adopted easily, advancing the practice of education and the resources with which to advance the theory. I describe this as a sketching phase in a broader learner experience design process.

In this chapter, I describe a case of adapting the design sketching methods common in HCI to the theories and methods of learning science. The outcome is a lightweight local theory of the target context. This local theory makes predictions about how different design interventions would affect different desired outcomes in the context. These design intervention hypotheses can, and often are, developed directly from literature or casual observation, but I will argue that the rich tacit knowledge gained through this process improves the designs and thus the interventions and their assessment. Further, I describe a subsequent process for combining local theory with general theory to evaluate contextual factors in the implementation of general principles and opportunities for general principles to address issues observed in context. With this map of the opportunity space, I describe a process of ideating solutions to the most opportune issues and sketching out designable experiences that highlight those issues. Finally, I return to the stakeholders to discuss the sketched learning experiences, refining the opportunity map and my understanding of the learner needs within it.

3.2 Methods for Discovering Opportunities

Design requires first articulating a current state and a preferred state (Simon, 1969). Education researchers often have current and preferred states in mind, as they are well familiar with the challenges and goals of education. Successfully designing systems that change the world from the current to prefer state also requires understanding the context.

For a technology to affect change in a context, it must work through the culture of that context. (Culture is meant here broadly as the influences on a person's behaviors, e.g. from institutions, authorities, peers and self.) A technology artifact with a perfect cultural fit would be adopted quickly and effortlessly. Consider the iPhone. What prospects would an entirely new computer platform have in a saturated marketplace? The iPhone was adopted because, although it is a general-purpose computer, it fit the cultural role of phone. There are contexts it did not fit to initially, such as enterprise environments, which often require top-down administration. Once it met those contextual needs, the iPhone was rapidly adopted.

Research in education must also have cultural fit at some level. If an artifact does not fit the culture now, why expect that it would later? If the goal is to study a principle and not an artifact, why expect that the knowledge gained could be applied? There can be valid answers to these questions but they must be asked. A further advantage of cultural fit is it increases the rate at which research questions can be asked and answered. A widely adopted artifact can provide a large and varied data source in which to compare quasi-experimentally across differences of learners or context, or to compare experimentally across theoretically driven variations.

Design researchers in education have developed methods for understanding context, such as the Informed Exploration phase of the Integrative Learning Design Framework (Bannan-Ritland, 2003). Many variants of design-based research practices are used in research on technology-enhanced learning (Wang & Hannafin, 2005). Figure 3-1 presents a simple framework, orthogonal to those, to describe the process of producing software artifacts used in research and practice. The vertical axis indicates the steps in developing a software artifact from initial motivations in a context all the way to a working system. The horizontal axis indicates the setting of the research activity and progresses from understanding the present state to imagining preferred states to testing them through simulation and finally to enactment in the naturalistic context.

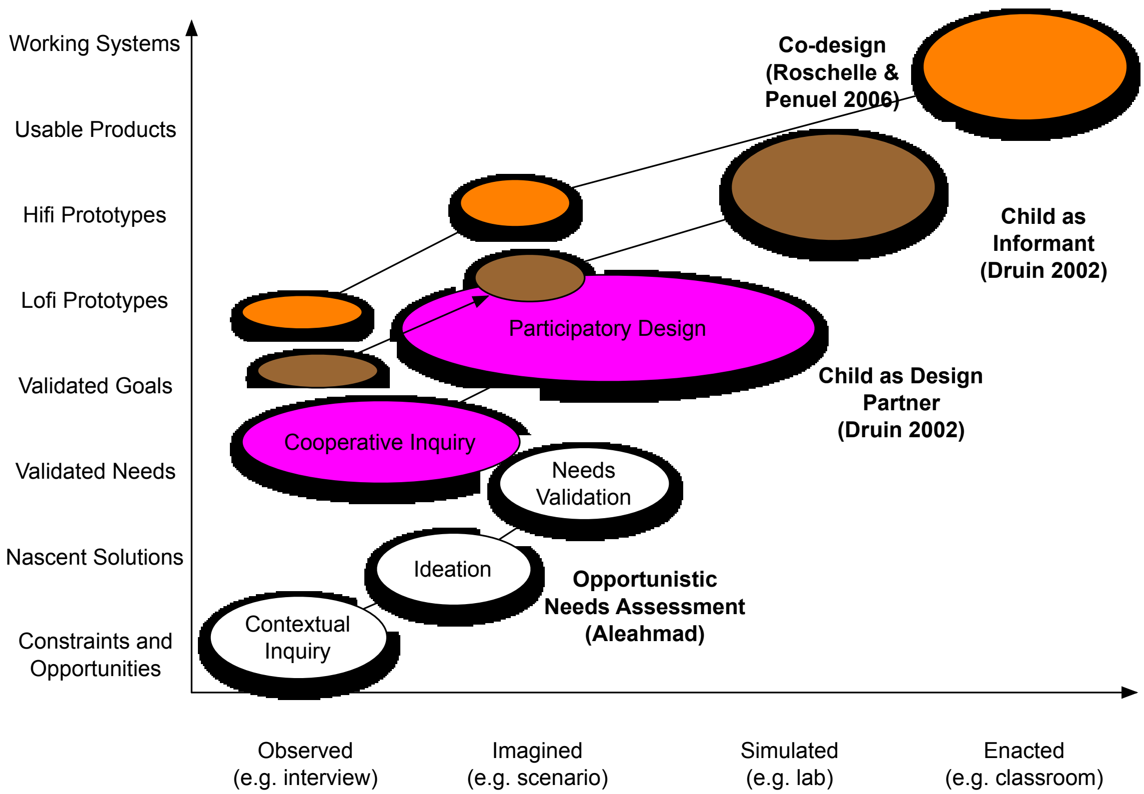

Figure 3-1 Methods in education technology design to support contextual fit

To design systems for user-driven adoption requires one more consideration: what are the current and preferred states as perceived by the stakeholders in the context? Sometimes you can simply ask them, but as Henry Ford is (apocryphally) alleged to have said, "If I had asked people what they wanted, they would have said faster horses" (O'Toole, 2011). Co-design (Roschelle & Penuel, 2006) and Child as Informant and Design Partner (Druin, 2002) are two examples of working with the stakeholders to design for their needs. The field of HCI, as a discipline largely dedicated to designing new technological artifacts that are complex and fit to people (Fallman, 2004), offers methods that are more widely practiced and validated. I contend that an earlier inquiry into context, by adapting HCI methods to an opportunistic assessment of needs, can help design researchers in education to develop systems that fit better into real world use. Systems that tap into a need deeply felt by stakeholders may be adopted easily and become more productive tools for research and design.

The goals of this inquiry were two-fold. One was to discover opportunities for which I could build an operant probe and conduct a scientific evaluation within the time constraints of this dissertation (1 year). The design process steps to select the needs for which to design, i.e. the opportunities, are depicted in Figure 3-2. An additional goal of this inquiry was to describe the culture of undergraduate lecture courses in such a way that other researchers and technologists can find new opportunities for their own work. The *Theory in Context* section below offers this type of description.

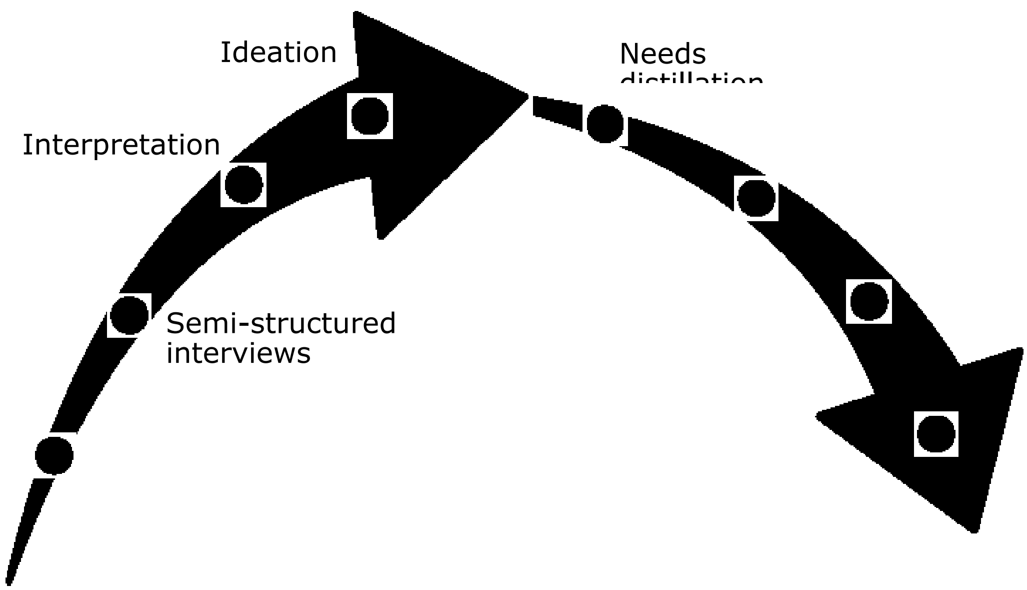

Figure 3-2 Steps of design inquiry process

3.3 Focus and context selection

For this design process, I began by selecting a context that could benefit from the scalability afforded by computing technology. Large college-level lecture courses were chosen because they fit this criterion, are relatively convenient to study in a university, and college instructors have more freedom to try experimental technologies and deploy ones demonstrated to work.

A successful design that fits easily into college lecture courses could improve the learning outcomes of tens of millions of students around the world. In 2007, over

150 million students were enrolled in college ("Data Points: More College Students Around the World," 2009) and 18 million in the US alone (Snyder & Dillow, 2011). The numbers of enrollments is growing rapidly, and most of these students are in the familiar lecture type course. The learner experience in these courses is based in centuries of tradition and ripe for innovation.

3.4 Semi-structured interviews

To investigate the context I use HCI methods and adapt them to the values of the operant probe paradigm. The first set of methods come from Contextual Design, a methodology that has been used in hundreds of software products (Beyer & Holtzblatt, 1997). The first step of contextual design is Contextual Inquiry, collecting data through interviews and observations in the actual context of potential use. The second is Interpretation of the data into models of perspectives on the context (e.g. cultural, physical) for each informant. The third step is Data Consolidation of the models across informants to create a single comprehensive model for each perspective. I also used Affinity grouping, another approach to Data Consolidation, to find themes among the side notes I took along the way.

To conduct the contextual inquiry interviews, I first began by focus setting, a process for making explicit key goals of the inquiry. Aspects of the context outside the focus can come up in the interviews but the focus helps keep the dialogue on a productive course. To choose the foci, I first wrote on sticky notes a few dozen ideas I had about aspects of the context that technology could help improve. I then clustered these together on a whiteboard in an affinity diagram to find distinct themes. From these several themes, I chose two. The "material use" focus encompassed the production and management of materials such as syllabi, lessons, quizzes, class notes, etc. The "social context" focus encompassed how student peer interactions, discussion boards, sharing of materials among instructors, asking questions, etc.

3.4.1 Participants

Contextual Inquiry requires interviews with participants in each role that will be impacted by the technology. For our chosen context of the large lecture course, I included students (10), faculty instructors (6) and graduate teaching assistants (3). The faculty instructors were drawn from two universities in the area by emailing a list of candidates. The list of candidates was made from personal recommendations and the course schedule. All faculty were from psychology except one from biology. The teaching assistants were drawn from the course schedule for the summer term in which the research was conducted and were all from psychology. Students were eligible by virtue of a participating instructor and were solicited by announcements in class or by email.

3.4.2 Data Collection

The interviews were conducted in the contexts in which the subjects engage in activities for the course. For the faculty and teaching assistants these were their offices. For the students, these were the library, their dorm rooms, and cafés. Each interview spanned 1-½ hours and was audio recorded. For interviews in which the subject mentioned a tangible artifact, such as their class notes or a study guide, these were documented by taking a photo.

3.5 Interpretation

After the interviews, I listened to all the recordings and recorded observations relevant to the models specified by the Contextual Design methodology: Cultural, Flow and Sequence. I also noted "Design idea" observations that subjects had offered, that I had thought of out loud in the course of the interview, and that occurred to me while listening. For the affinity grouping technique of Contextual Design, I also noted "Affinity" observations, a catchall for observations to continue to process but which weren't design ideas or informative to the models. The design idea and affinity type were used later for the ideation phase.

Rather than a strict transcription, I typed paraphrases that were time-stamped to the audio with Transcriva, a Mac application. I exported the observations from each interview and concatenated them all in a spreadsheet of 1,014 observations. Each row contained the type, the subject, the time in the audio recording, and a note with the observation. Figure 3-3 shows the distribution of observation types. Note that the Cultural model observations were much more numerous than the Flow and Sequence observations.

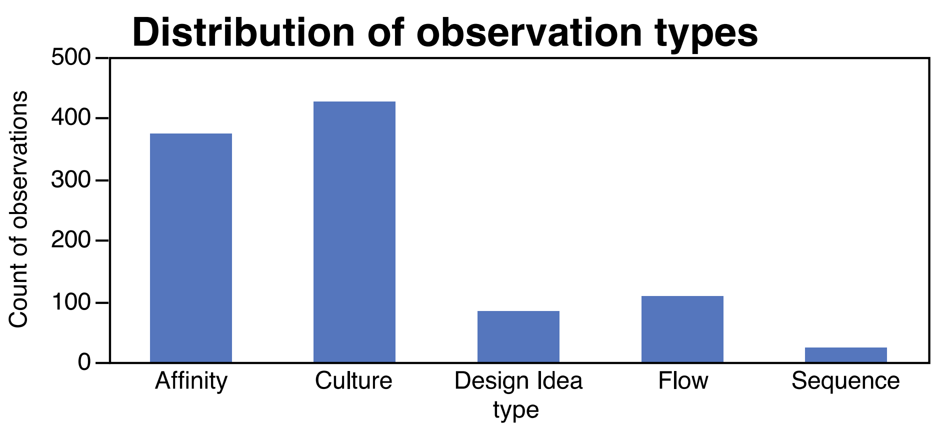

Figure 3-3 Counts of each type of observation recorded

To visualize the cultural forces influencing each stakeholder's behaviors, I created graphical models of the data in each interview. I used a subset of the models in Contextual Design, including a cultural model for every interview, flow models for most of them, and a few sequence, artifact and physical models. Only the cultural model revealed insights not obvious in this context, and it is the only type for which I generated a consolidated model.

3.5.1 Consolidation

To reveal patterns across the interviews, I combined all 19 models into a single consolidated model. Because I had come to focus on the cultural aspects, I did this only for the cultural model. The observed cultural factors that influence behavior of students and teachers are in Tables 3-1 and 3-2. Each *Factor* is an abstraction of multiple statements recorded in the interviews. The *Influencer* is either the actor themselves (e.g. students behaviors are influenced by their own goal of doing the minimum to get the grade they want) or another actor in the context (e.g. teachers influence student behaviors through culturally communicating the idea that they reward students for short-term, not long-term learning.) This consolidated cultural model presented a coherent view of all the roles in the cultural context and the cultural forces acting internally and between the different roles. Because it combines the models from each informant, it reveals tensions, breakdowns, and design opportunities among the cultural elements.

Table 3-1 Cultural influences on student behavior

Influencer	Factor
Self	I do the minimum I must to get the grade I want
Self	I procrastinate on everything
Self	I need to do well in my classes to feel good (otherwise feel stressed)
Self	I am distracted easily
Self	Organization reduces my stress
Self	I don't retain information
Self	I feel bad about my poor study habits
Self	My effort is based on how interested I am
Self	I want to know what I know and what I don't know
Self	I am unmotivated without personal interaction
Self	I can never be good at certain parts of this class
Self	I doubt my value to my peers' learning
Self	I enjoy having room for my own perspective
Self	I want immediate feedback on my reasoning
Self	It's painful to put in effort and not see results
Self	I am motivated by work that will help me succeed later
Self	I can't bear the thought of being wrong
Instructor	When we care, you care more

Instructor	I reward you for short-term, not long term retention
Instructor	I let you get away with bullsh*tting
Instructor	I grade you on things that you don't care about

Table 3-2 Cultural influences on instructor behavior

Influencer	Factor
Self	If students don't do well, I haven't done a good job
Self	It's important to connect class to students' lives (prior knowledge and real world experiences)
Self	I enjoy teaching depth more than breadth
Self	I care deeply about the quality of my assessments
Self	Stimulating learning is more important that accurately measuring it
Self	I must train you to learn on your own
Self	It's my responsibility that you value understanding this domain
Self	I want to install ideas that will shape your life
Self	The best learning occurs when it's fun
Self	I want each student to succeed in their goals
Publisher	Use shallow instruction and assessment we are selling
Department	You must turn out creative smart people
Department	We reward you for time spent researching, not teaching better
Students	When we care, you care more
Students	You should be able to make learning easy

3.5.2 Results of cultural modeling

The consolidated cultural model brought to light several conflicts between the cultural forces shaping the behavior of participants in different roles.

Misaligned goals: Student comments indicate an almost exclusive focus on getting a good enough course grade as their primary goal. Faculty, on the other hand, talk almost exclusively about learning. They lament that high quality evidence of learning cannot be measured in class. What can be measured easily, such as recall, is not important. What is important, such as application of knowledge, is laborious. What is most important, that the learning help students achieve their life goals, is intractable. (e.g., a professor quoting Herb Simon: "I never gave an exam I liked that was gradable.")

Misaligned motivations: Students are motivated to learn things they care about, but feel they are graded on content they do not care about. Consequently, students feel the system rewards those who focus on grades and not those focused on their interests or what is useful to know.

Tension in scope: Students want to know exactly what they have to learn. Professors want students to learn things that they do not specify or measure. Grades are the strongest signal to students and what is not assessed they will ignore. "Is it on the test?"

Depth versus breadth: Instructors enjoy teaching depth more than breadth. Students are more comfortable with breadth. Textbooks are broad and shallow to lower costs and meet majority preferences of the market. Publishers perpetuate a feedback cycle incentivizing breadth over depth.

Tension in support: Students have less stress when the class is organized. Faculty believe students should be learning how to learn so do not help "too much". Faculty sees their role as teaching a domain and do not provide instruction or assessment of how to learn.

3.6 Theory in Context

3.6.1 Method

To find the challenges and opportunities posed by the context to operationalize and contribute to educational theory, I identified how different theoretical models and recommendations were active in the context and barriers to activating them. To begin, I selected a scope of theory and practice recommendations to consult. For the recommendations I draw from three theory-based Practice Guides published by the IES What Works Clearinghouse: Organizing Instruction and Study to Improve Student Learning (Pashler et al., 2007), Using Student Achievement Data to Support Instructional Decision Making (Hamilton et al., 2009), and Structuring Out-of-School Time to Improve Academic Achievement (Beckett, Borman, & Capizzano, 2009). I refer to these, respectively, as *OI&S*, *USAD* and SOST.

To orient to how these theory-based recommendations are active in the context, I coded each theory-based recommendation by the evidence for its use in the context. For each recommendation, I coded what data were presented in the interviews for its application and perceived need in the context. Levels of None, Low, Medium and High emerged, which I coded 0-3. When there was evidence against the application or perceived need, it was coded as negative. The two factors were also split by the two key stakeholders, students and instructors. When there was a range among informants within the role, the value in the range was chosen that most highlighted an opportunity for design.

In the spreadsheet I then added calculated columns: *Needs Spread* for the difference between the perception of students and teachers (three were >=3), *SN vs. TP* for the difference in student perceived need versus the teacher practice (two were >=3), *TN vs. SP* conversely (five were >=3), and *Sum Practice* adding student practice and teacher practice codes (three at -4 and four at -3). I discuss these below situated in the cultural model.

3.6.2 Cultural barriers to implementing educational theory

The consolidated cultural model reveals several cultural barriers that may hamper the use of evidence-based recommendations. Below I begin with the recommendations highlighted by the above analysis and situate them among the cultural forces of the actors in the context.

3.6.2.1 1. Quality questions are scarce

Because of the cultural practices of question production and evaluation, several of the recommendations are difficult to practice. One of the recommendations for which there is the most evidence, "Help students build explanations by asking and answering deep questions" (*OI&S* #7) also scored highest on teacher perceived need versus student practice (5). Teachers value deeper conceptual reasoning and particularly enjoy teaching the deeper concepts (e.g., "I'd like to teach a higher level"). Students tire of rote questioning and enjoy questions for which there is not a single "right" answer that they either hit or miss. A strong opposing force is that deep questioning requires more effort on the part of all participants. These efforts can be divided into question generation and answer assessment. For professors, there is a trade-off in time and difficulty between their generation of questions and grading assessment. An open-ended question can be easier to produce than a set of multiple choice questions that measure the same understanding, but the price is paid at grading time when each answer has to be read and fine distinctions have to be inferred as to what levels of understanding are demonstrated. Grading time can be reduced by multiple choice or short answer questions, but to make these "deep" questions requires much more thought and, often, more time pilot testing to produce good foils that represent common student misconceptions.

The instructors in our study often turn to textbooks for their questions because they are plentiful and no additional cost. Unfortunately, the questions in texts are often shallow and rarely valued by the instructor. The textbook publishers sell primarily the textbook and a book's question bank is offered to help sell that book, rather than being a revenue builder itself. An associated design idea is a technology in which publishers' create question banks as a living online resource that instructors can review, collaboratively filter and improve upon.

The efforts of professors to produce questions can also be amplified. They work so hard to make these that they are cautious in sharing them and they are typically shared with only a few colleagues at best. This observation supports a design constraint that there are sufficient security assurances of access only by qualified instructors. With such, there might be a peer market so that faculty can exchange their valuable questions, increasing the supply and thus reducing the cost.

3.6.2.2 1a. Secrecy of questions limits formative assessment

The scarcity of questions compels the teachers to re-use the good questions they do have. The re-use leads to careful guarding of questions limited in use to summative assessments and accessed only by trusted colleagues. The hampers *OI&S* recommendation #5a, "Use pre-questions to introduce a new topic." Practicing this would help address students' expressed needs to be aware of what they do not know and have evidence of their learning. It would also help prevent the perception of some students that teacher's questions are meant to trick them. If there were a sufficient supply of exam-quality questions, they could be used to help guide students' attention to a new topic and familiarize them with the content and style of questions their grade will later depend on.

Another formative use of questions is after they are answered. The OSAD guide recommends to "Teach students to examine their own data and set learning goals" (#2). The guide explains that "instructional strategies such as having students rework incorrect problems can enhance student learning." (Clymer & Wiliam, 2006) Because teachers are compelled to reuse their questions, they cannot allow students to keep their graded exams. (Students would certainly share the questions with other students who had not yet taken the course.) One technology to address this without revealing the exact questions would be to provide detailed formative feedback with each summative assessment. Instead of the single percentage grade from the Scantron™ most students receive now, they could be provided with a report of their performance on different knowledge components (of large grain size if need be) and supplemental instruction to help address their deficiencies. This could help promote a Mastery style of learning and explicit goal setting.

Another technology opportunity is to address the scarcity of questions by the recommended practice of students asking deep questions. With the proper supports for quality in production and filtering, students' questions can be used for formative or even summative assessment. Because students' questions are often shallow, they could also be used to free up the teacher's efforts for deeper questions or as a starting point for a guided inquiry into deeper questioning. This crowd-sourcing technique could be extended to students that are more senior or the entire Internet. With sufficient quality assurance mechanisms and participation, instructors or tutoring systems could draw randomly from large pools without risking repetition.

3.6.2.3 2. Specifying learning goals conflicts with flexibility and adaptivity of the course

Several recommendations and perceived needs hinge on specifying learning goals. The OSAD guide recommends to "Teach students to examine their own data and set learning goals" (#2). "Tools such as rubrics provide students with a clear sense of learning objectives, and data presented in an accessible and descriptive format can illuminate students' strengths and weaknesses (see recommendation 5 for more information on reporting formats) (*Assessment for learning: putting it into practice*, 2003) Students want to know what they know and don't know, which depends on a taxonomy of what they are expected to know. They also want to feel a sense of progress for their efforts, which is difficult without clear targets. Finally, students expressed that organization in the curriculum lowers their stress and some instructors expressed a desire to minimize anxiety for students.

The responsibility for specification of learning goals falls on the instructor, and in them are two forces pushing back. One, it is a lot of work that they see little point in. They believe they know what the goals are even if they have not articulated them. They will teach them so why spend the effort to write them down? This could be addressed by outsourcing that labor somehow. Students or assistants could be involved in writing the specifications. One technology opportunity might be a shared workspace in which the students collaboratively develop and revise what they believe to be the learning goals of their lessons and activities. This would certainly yield gains in metacognition and the instructor could give students the feedback necessary to reach a working set of goals.

The second conflicting force is deeper and a part of the instructor's conception of their role. The instructors see themselves as providing a unique experience based in their particular domain and pedagogical expertise. To follow a highly specified curriculum, especially one they did not create, reduces their creative role to one of executing a program. While detailed rubrics help students plan their learning and understand their progress, they also limit the instructor's sense of freedom in adapting to his students or seizing upon spontaneous opportunities that "will often pop into mind during lecture". In other words, instructors feel that learning goals can be slippery and shift throughout the course. Accordingly, assessments are crafted based on the course as taught, not as planned, with redistributed emphasis based on how much a topic was or was not covered.

These observations lead to a design constraint to preserve instructors' needs for creative input and flexibility with students' desire for structure and predictability. Can we create a design to support explicit goal setting in a fluid manner? Students might be involved in articulating the learning goals based on the professor's teaching and revise the set as it evolves in class. The professor can refine the specification and use it with software to dynamically generate assessments that align with the course.

3.6.2.4 3. Faculty ostensibly teach how to learn but don't assess or instruct it directly

The professors in our study see their most important role as teaching students how to think, not what to think. While they carry the responsibility that students value understanding of the domain, their larger goal is to instill ideas that will shape each student's life. These skills for learning range from the domain-specific metacognitive support to study strategies to general time management. Yet, none of these is taught explicitly, even though the largest courses are introductory courses with students unaccustomed to the rigors of post-secondary education. (e.g., ""by second semester freshman year I was trying to learn how to study, pretty much teaching myself.")

Teaching study strategies has field-based evidence (*OI&S* #6) and little cultural resistance. The guide recommends teaching students techniques to break their "illusion of knowing". One way it recommends is through tests and quizzes, which is widely practiced and does not teach any strategies to the student. The other recommendation is to teach students how to create "judgments of learning" themselves while studying, transitioning from demonstrating in class to using the techniques on their own. Some participants, particularly in cognitive psychology courses, do share techniques like this. What they do not do is assess it. Students make clear that they do the minimum they can to get a desired grade. Whether a technique works or not, students are unlikely to find out because they lack the motivation to try it. For those students, it is ineffectual instruction. Formative assessment helps detect ineffectual instruction, but the instructors avoid explicit assessments of strategies because while they wish students in their course learn how to learn, they are not teaching a course on "how to learn" per se. They do not want to award points in a psychology course for general study skills. Moreover, these skills are implicitly assessed through the student's achievements in the domain.

The technological opportunity here is to develop a study tutor that motivates use either intrinsically or extrinsically besides a grade. The system might scaffold students in acquiring the skills and dispositions of effective study strategies and fade as they internalize them. Ideally, it would assess student use of target strategies and provide formative feedback to the instructor on how the students are learning to learn. Moreover, it would provide more data by which instructors can understand students' achievement in the domain per se.

With sufficient data and interactivity, computer-based systems could support mastery of fine-grained metacognitive skills. An intelligent tutor for metacognitive skills might help students to see their progress and provide the immediate feedback that they desire. A fine-grained model of metacognitive skills in the domain would support a Mastery orientation to the skills, motivating students who believe their abilities are fixed. Of course, this is a grand challenge but the results of our inquiry suggest it is a worthy one.

Another decomposition that technology can support is at the large grain size: completing the course. Students have difficulty breaking their goals into actionable chunks and scheduling them optimally. To the extent that the course is not meant to assess this skill, the entire course could be broken down to a hierarchical checklist with target dates. Software guides through the checklist would provide students the organization they seek and lower their stress, allowing them to focus on learning. To the extent that instructors wish to train or assess these skills, software could support student decomposition of the course syllabus and provide feedback on the chunks and on the schedule students have set. By making planning explicit, instructors can choose and convey how much students should be learning and internalizing "how to learn" versus the domain alone. Such approaches would ideally be implemented and applied across courses.

3.6.2.5 4. Improving instruction through student data requires infrastructure and roles that don't exist

Instructors care very much about the achievements of their students. They pay attention to class discussion, assignment submissions and exam scores to form their course, and to the apparent needs of the students. (E.g., "if people are not scoring very well, my presumption is that I didn't do a good job.") There is much evidence for the importance of these practices, so much that a whole IES practice guide it dedicated to using student achievement data to form instruction. The guide emphasizes that data practices require a culture and infrastructure. One recommendation "Provide supports that foster a data-driven culture within the school", includes steps: 1) Designate a school-based facilitator who meets with teacher teams to discuss data, 2) Dedicate structured time for staff collaboration, 3) Provide targeted professional development regularly. I did not find any evidence of these practices.

An opportunity this presents to technologists is to support a bottom-up culture of data-driven instruction. For several professors in our study, exam scoring is outsourced to an office of testing, which returns the student answers and simple percentages. One instructor checks each question and each student test for irregularities. It requires hours of manual work that could be automated and other instructors do not have the time. If various reports were instantaneous, instructors could learn more about the quality of their assessments. If they were linked to a question bank, instructors could compare student performance, and their instruction, from semester to semester. One constraint would be to include the testing office in the design process and solution. Otherwise, they may fight what they perceive as a threat.

There are many other data relevant to student achievement that are not available in a data analytic form. For example, textual responses are often hand-written, preventing use of computer-based language analysis. Other modes of data could be collected, such as non-traditional assessments like question authoring, the metacognitive activities above, how students use their time, or simply class questionnaires. Using technology to provide new measures of student activity would allow instructors to explore and potentially develop their own hypotheses about how to improve their courses. Detailed student data could also help instructors differentiate instruction. "One of my biggest challenges is the range of abilities of students coming into the class," said one instructor.

3.6.2.6 5. Instructors have precious little time to integrate new technologies

A barrier impedes the implementation of all any technological system for implementing the practice recommendations is that instructors generally have very little time to spend integrating them. Faculty #3 lamented on the use of PowerPoint, "I spend an amazing amount of time with the apparatus of instruction rather than what I really want to be doing, which is thinking." For most professors and teaching assistants, student learning is not their primary responsibility. Once a curriculum is developed, they have little incentive to change it. No wonder that many instructors reported beginning new courses with a colleague's or publisher's materials and adapting them over time.

Information technologies that are meant to help can end up taking more time. Faculty #2 tried using online discussions and retreated. "I did feel like more people participated, but it was easy to consume a much larger number of hours [of] everybody's [attention]. […] I remember thinking, wow, if I was doing this all the time I would end up spending an inordinate amount of time."

The instructors I spoke to were interested though in trying new ways to improve their courses. For example faculty #2 continued, "The things that make me change things are more… sort of new things that come my way, because I get bored. You know what I mean? Doing the same thing." To be accepted, though, new technologies must be easy to integrate into courses and not require much time or curricular changes from the instructor.

3.7 Ideation

The aim of the fieldwork was not merely to describe the present but to look to the future to what can be designed to improve on the present. To take a step towards the concrete, in the next phase I generated many distinct ideas for systems that could be built. I began this process by reviewing the observations and organically grouping them per the affinity grouping method of Contextual Design.

The numerous observations and ideas required a different approach than the traditional whiteboard of sticky notes. To create the paper notes, I wrote a Python application in the Django web framework to format the 457 affinity and design observations into a page template that I then print and cut into small squares. Instead of a whiteboard, I laid them out on the floor. Together with a design assistant, we put related observations closer together and gradually honed in on 14 parts of the context for which to design (Table 3-3).

Table 3-3 Clusters of design ideas and observations

Class logistics	Knowing what I know	Study habits	Professor knowing each student
Assessment production	Help seeking	Motivation	Learning goals
Attention in lecture	Encoding lectures	Community	Instruction production
Reëncoding	Pear learning		

Using the data on the context, its interface with theory, the tacit understanding developed through extensive interviews, and the organization of the affinity groups and process, I began to generate ideas for systems I could build. This ideation was a relatively fast process and after a few days I had generated 64 distinct ideas, some of which are listed in Table 3-4 below and all of which are listed in Appendix B. Many of the ideas were carried over from my spontaneous thoughts while interviewing and later listening to the interviews.

Table 3-4 Sample of ideation

Class study partner pairing system.
TA review session voting system (submit questions, everyone votes and popular ones first)
Contributions that do not require being right or wrong. (E.g. cog psych scenarios, provocative questions)
Analytics on how much students are working and how. (Anonymous logging and reporting.)
Study behaviors tutor, tied to real data from learning activities and outcomes.
Big ideas database to find concepts that cut across findings
Versioning system for teaching materials with in-class annotations on each version.
Micro-experiment tracking system for educators. Quick pre/post assessments around a small treatment.
Recording questions keyed to time code and point in the slides. Embed student experience in materials (for future self, and others)
Crowd-source the content of a learning game (authoring/use class/library/bus/home)
Algorithm to distribute students throughout the hall for break-out with different groups

3.8 Needs Distillation

Each of the ideas was potentially good, but which were worth trying to develop into successful operant probes? Which ideas would be accepted by users and also contribute to scientific understanding? Instead of selecting directly from the ideas, I used them as part of a process to better understand the user needs and constraints, and the relevant science. The work following Ideation contracted the search space and helped to frame the problem, articulating the current and preferred states for which to design.

Many different needs were evident in the observations from the interviews, but for which needs would designing lead to a successful operant probe? The ideation phase helped to think creatively about the ways that technology could help with those needs. Working back to the user, I distilled these numerous technological ideas into the distinct user needs that motivated them. In doing so I focused on the subset of all observed needs that technology could be used to address. E.g. "Motivating interest, nourishing curiosity". The full set with initial solution ideas is list in Appendix B.

3.9 Scientific Impact Evaluation

These distilled needs helped me navigate the design space to technological systems that would likely be accepted, but for scientific research through interaction design, I had a second goal of operationalizing scientific theory through designing for the need. This is where the design of operant probes deviates from traditional HCI practice. I conducted a literature review and for each for each candidate need, I cited lab-based results relevant to that need, and annotated it with three factors in its potential for impact: 1) That learning effects are predicted; 2) That the design and use of the artifact would inform other applications of those results; 3) Its fit to the research team. Could we capitalize on this research opportunity?

Through this new design process technique, Scientific Impact Evaluation, I identified which needs would more likely have scientific impact if designed for. Appendix C shows the evolved description of needs and the relevant learning science principles. The principles were drawn mostly from the IES Practice Guide for "organizing instruction and study to improve student learning" (Pashler et al., 2007). For example both students and teachers felt the need to support students' sense of what they've learned. The space of solutions to this need related to the lab-based research on "help[ing] students to allocate study time efficiently" (#6) and the studies behind the two sub-principles to "Teach students how to use delayed judgment of learning techniques to identify concepts that need further study" (#6a) and "Use tests and quizzes to identify content that needs to be learned" (#6b). Further it was expected that a successful operant probe designed for this need would contribute to the scientific understanding of students' motivations.

3.10 Needs Validation

After the distillation and scientific impact evaluation, I had 17 needs that I had observed. Because my goal was to design for technologies that would be accepted, I also had to validate whether the stakeholders themselves felt these. Needs Validation is a design method to assess whether what needs the stakeholders themselves perceive and what solutions they are likely accept (Davidoff, Lee, Dey, & Zimmerman, 2007). This is no guarantee that the system built would be accepted, but it helps point the designer in the direction of acceptable systems and improves understanding of each need in order to properly frame the problem.

I elaborated the 17 needs for use in stakeholder interviews. Each scenario was refined and illustrated into a storyboard, to activate the user's memories and feelings in a situation and draw out their perceptions of the plausibility of the situation, the character's behavior, and their perceptions of the technological artifact. This can reveal the needs that a stakeholder perceives for their own role in the context. Appendix C shows all the needs annotated with learning science principles and descriptive scenarios to probe on the need. Figure 3-4 shows an example illustration for the scenario in need #10 and Appendix D shows them all.

Scenario 10

Thursday night, Peter realizes he has a big essay due Monday morning and he'll have to work hard on it over the weekend.

At 9pm on Friday night Peter is tired of reading and decides it's time to go out. He records on his phone that he's switched from Studying to Partying.

Saturday afternoon Peter picks up his essay again and records Studying on his phone. After an hour he records he'll check Facebook for just 10 minutes. He loses track of time until the phone beeps to remind him and he returns to studying.

Sunday evening Peter has completed his final draft and submits it early. With satisfaction he records Watching TV.

Figure 3-4 Sample scenario for needs validation interviews

With these elaborated scenarios and storyboards in hand, I conducted interviews. For each scenario I articulated my idea of what need the scenarios would recall for the student and teacher, along with a lead sentence for the interview to frame the conversation. I interview 7 students (2 as a pair), a group of 3 TAs, and 3 faculty (2 as a pair).

The needs validation process informed both whether stakeholders perceived the hypothesized needs and the cultural and practical constraints on my leading solutions to those needs. For example, I observed that some students resented school for reducing their connections with their friends. I perceived this as evidence of a need to involve students' wider social worlds in their scholastic activities and presented a scenario in which students share their grades and study activities with their friends and family (need #6, LearnShare). Both students and faculty rejected this. One faculty pointed out that during college young people are trying to develop their individual identity, breaking such support ties and finding their own way. While such a system may be desired and helpful in secondary education or even adult learners, the college student context would not accept such a design.

Another notable rejected need (or want) was to be more motivated in class through competition. This was based in my observations that some students already perceived class as a competition for good grades and wanted to be recognized for other achievements, such as writing the best study questions for one's peers (need #5, PeerQuiz). Students and faculty had no interest in promoting competition. While some students may thrive with a competitive technology, such a system would not be accepted by the faculty who are the gatekeepers to their courses. This is an important constraint on the move towards gamification in education.

As for the idea of students writing exam questions, faculty did like it. One instructor said, "This I would use. I'm going to do this actually." However students did not. They didn't trust that the questions would be helpful for the exams and might even lead them astray. Getting this design right would entangle complex social issues and require the active participation of the instructor. Indeed, systems such as PeerWise do implement such a system and do. Considering this system helped me to hone one of my design constraints: that the systems be scalable broadly without requiring the instructor to spend much time or to change their course instruction or assessment.

An interesting constraint of the traditional institution of higher education as a whole was evident in the rejection of the long-term retention scenario. In interviews students expressed frustration at forgetting what they had learned. This could be triggered, for example, by trying to help a friend through a class they had taken earlier and no longer being able to answer the questions, or simply by taking a subsequent course in a series and forgetting the prerequisite knowledge. While students and faculty both agreed it would be best for students to retain what they learn, and the OI&S first principle is spacing of practice, no one was willing to accept OlderCheck, a system in which receive an electronic follow-up quiz months after they complete a course. While an instructor stated plainly, “They forget most of what they learned after the exam,” he could not imagine a scenario in which students would do this voluntarily or could be incentivized to do this. This design constraint is imposed by the current structure of university courses and credits. Competitors to the university model may break free from this constraint. However the goal of this fieldwork was to identify improvements to the college courses of today. This issue highlights the potential for new institutional structures in the future and through this I adopted the design goal to design for back-porting the features from this better future into the structures of today.

Some of the needs were strongly felt but did not provide much opportunity to engage scientific research. An example of this is the students need to ask questions in lecture when they are afraid of seeming ignorant or slow. Need 13 in Appendix C and illustration 13 in Appendix D present a scenario for a system whereby students can ask questions asynchronously on their mobile phones during lecture? Students were eager for a system like this and teachers also supported it (if it didn’t disrupt the flow of their lectures or take too much time). However, based on the scientific impact evaluation, I saw little opportunity to advance science through designing for this need. Of course, other researchers or teams may have more relevant expertise towards a scientific understanding of this need and acceptable solutions.

3.11 Needs Selection

The ultimate goal of the field study was to identify opportunities for new designs that would be accepted in college lecture courses. The needs validation affirmed two needs that pointed to solutions that would be informed by the science-based practice recommendations. In other words, two needs were felt (without satisfaction) and had solutions that were acceptable and had scientific evidence in to guide their designs. These features support the opportunity for an operant probe that could test the contextual design knowledge, the general scientific knowledge, and the effects of integrating them in a concrete scalable form.

3.11.1 Time Management

The first need was to support students' time management. All faculty and teaching assistants perceived this as a problem. Many students personally grappled with this and those who did not have this problem affirmed that many of their peers did. Students conceived of this primarily as procrastination. Probed on her study habits, Student 3 said sheepishly "I'm kind of a procrastinator. It's not good!" This student was ashamed of her time use for school ("I should put a lot more effort in all my classes. I've been a real slacker.") even though she falls asleep studying late into the night to wake up for work in the morning. Student 5 felt similarly guilty, "I have very bad study habit. I don't prereading." What students conceptualize as their lack of willpower or discipline could instead be the failure of the educational offering to support them. For example Student 7 reported,

When I got to college, my first semester freshman year was not the best. I didn't know how to study, and still don't study well. And I openly know that. Um, didn't know how to study, didn't know how to take notes. Didn't really have to in high school, you know? You can kind of get through high school without doing anything too… extensive. So that was tough. By second semester freshman year, I was then learning out to take notes and how to study, pretty much teaching myself.

Faculty and TAs are sympathetic but don't address this is in their primary instruction. Faculty 3 said, "Students that exercise time control, planfulness and stuff, really are ahead of the game." The instructor subjects take some time to help, but students come in with a huge variance in time management skills and they mostly help with such issues in office hours. However many of the most needy students don't come. TA 3 explained,

My experience as a TA is that most students don't come to office hours. Ones who do come to office hours usually are the ones who are already pretty high achieving students, that have, that probably are high achieving students because they have a lot of good skills. How to attain help appropriately. How to know when you need help. My experience teaching […] has been really eye opening about the varying levels of preparation for college.

Synthesizing from the contextual inquiry, recommendation literature ("help students allocate study time efficiently", OI&S #6), ideation and needs validation, I settled upon a design goal for which a solution would likely both be accepted and contribute to learning science,

- Computer support for students to use their limited time most effectively

To design such a system, the field work made apparent these constraints:

- Require no upfront action by the student in order to benefit
- Require no changes to the instructor's curriculum or schedule
- Require little or no time from the instructor to offer in her course

This goal and set of constraints formed the core design principles for Nudge.

3.11.2 Studying More Effectively

The other evident need was to help students study more effectively. Students validated the need for support in the study process but both students and faculty described the constraints on achieving this. Many students have trouble focusing on studying without action. Student 7 explained, "If I'm not actively doing something then I'm gone, that's it. … If I stop writing, within 5 minutes that's it. You're done. It's gone." Some expressed wanting to study using more conceptual questions and not merely recall, but believed multiple-choice questions were all teachers had time for. When asked, "What's your ideal test?" student 10 rejected the question. "They have to be multiple choice questions these days. You can't expect a teacher to hand grade 100 tests." Faculty 5 hears from students that they want more multiple-choice questions ("One thing they always ask is can we get more multiple choice questions") but she can only provide so many. "The problem is if it's a question I think is really good, I'm saving it for the exam. And if it's a question I don't think is good, how much does that help you?"

While students may appear to crave multiple choice questions, my experience talking to students suggests that what they really want is to prepare for the test and receive immediate feedback on how well they are likely to do. Student 6-valued practice tests and said that in some domains, like math, you cannot study for the exam without practice. The paucity of practice currently may be due mostly to the production of good practice materials. Faculty 5 shared, "It gets more challenging for me every year to ask good, challenging questions."

One of the students' goals in studying and practicing is to focus their time and attention only on what they don't yet know and need to. For example when asked, "How do you choose to study?" student 3 replied, "I don't know. I'll just like go over my notes and anything that I'm like unclear about, I'll study that like more closely." While student 3 likes practice exams, she lamented that she doesn't have time to do them all. Feedback on the scenario for a progress tracking system affirmed that students definitely want help knowing what progress they're making. Faculty said they would love to do this for them but it means they have to write a lot more exam questions. They recognize though that many students "just don't know how to study". For example, one faculty recounted that students come to office hours complaining, "I read the chapter and memorized it and I don't know why I didn't do well." Students who don't do well perceive it as a limitation of themselves that cannot be overcome. Student 6 explained that science "just isn't [my] thing. […] I won't do well no matter how hard I try."

The field evidence points to a clear need to support students' study techniques. The IES practice guide provides some theory-driven guidance: to "help students ask deep question in order to build explanations,"(#7 (Pashler et al., 2007). Combining these points to a solution would likely both be accepted and contribute to learning science:

- Computer support for students to prepare for exams interactively by building explanations

The fieldwork also reveals some constraints for the design of such a system:

- Scaffold effective study techniques for students that work even for students who don't know them
- Be interactive enough that students are engaged
- Help students to accurately assess what they know and don't know
- Be self-paced so that students can go quickly over what they are already confident in
- Map well to course assessments so that students know when they are prepared

Again as for the other needs, there are general constraints of designing any operant probe for college lecture courses,

- Require no upfront action by the student in order to benefit
- Require no changes to the instructor's curriculum or schedule
- Require little or no time from the instructor to offer in her course

This goal and set of constraints formed the core design principles for Examplify.

3.12 Conclusion

The goal of this fieldwork was to identify needs felt among stakeholders in college lecture courses and to understand the constraints on what offered computer-based solutions they would accept. The HCI practices of contextual design and needs validation helped draw a map of opportunities to design solutions that would be accepted. Learning, however, is a complex activity for which evaluation is much more difficult than mere adoption or perceived value. People are poor judges of their own competencies and learning experiences. To help ensure that the designed systems also achieved the goal of student learning, the design ideas were grounded in the scientifically vetted practical recommendations for educational experiences through the IES practice guides. In the next two chapters I detail the designs of the two proposed systems and ground them further in the scientific literature, both to improve their likelihood of achieving the desired outcomes and to provide an operant probe to better understand applications of the theoretical principles.

In closing this chapter, I'd like to take a step back to reflect on the methods used in this fieldwork. Could I have designed the same systems without conducting the fieldwork? Could someone else have who is smarter and better versed in the issues of college lecture courses? I believe this ethnographic-oriented work provided three essential benefits to this design process that would be difficult to arrive at simply through theory or intuition.

The first is the identification of needs and constraints felt by stakeholders. There is no literature (that I have found) saying that new technologies for studying education should not impose on the instructor. In fact, most design-based research in education calls for the active participation of the instructor. This is good for exploring how to improve teaching in the classroom, but there are many other needs students feel that teachers don't pay attention to or don't feel empowered or responsible to address. One of the results of this particular field study was that while some students have an urgent need for support with time management and it is critical to their learning, it is outside the domain the instructor is teaching (and on which they are so focused that they dedicated their life to teaching).

The second is to provide a wealth of tacit knowledge for the innumerable small decisions made throughout the design process, inquiry throughout the user-driven iterations, and the design of the classroom intervention by which the tools will be experimentally evaluated. So far I have only discussed the design space and constraints illuminated by the field activities. In chapters 4 and 5 I will discuss the designs of the two tools. It will be difficult as the designer and author to identify exactly what knowledge led to each decision (or even what decisions are worth remarking upon). I hope that a reader of this chapter will have developed a similar intuition. Regardless, I advocate that designers of education technology do conduct some amount of inquiry into the context and do so before they introduce their system. Further, that multiple courses be explored to develop an intuition for what varies and what remains constant between courses, and better design for operating at scale.

Finally, the rigor of this process and critical reflection at each stage can lead to a reframing of what the larger problem is that one is attempting to solve. Every design seeks to turn a current state into a preferred state. The lens by which we see now and conceive of what is preferred dictate the designs we can imagine. Domains for which the frame is especially prismatic are known as "wicked problems" (Rittel & Webber, 1973). That is, moving the frame or looking at it from different angles changes the scope and character of what is seen. Education writ large is such a wicked problem. Like urban crime, when you begin to operationalize the exact criteria of a preferred state, what is preferred becomes very slippery. For example, if we could make it so everyone had the same level of education would that be preferred? What if everyone could learn what they would in college without attending, is that something that institutions responsible for education would support? This isn't to say that improving education is intractably sociopolitical. It is to say that the frames of how we understand the problems in education can be moved and still lead us closer to someone's preferred state.

Through the active exploration of how learning in college could work, I came to see many aspects of the status quo as vestiges of a bygone era. The agrarian calendar when few are farmers, lectures in big halls when students can watch videos, assessments that can be fed into a Scantron, abstractions of student competencies into course grades and diplomas, the market bundling of instruction, assessment, certification and, for early undergraduate years, overnight camp. The business and technology of education are poised for a radical transformation. This fieldwork helped me imagine a preferred and achievable future that has different challenges than the present. One hallmark of the future of learning, where all knowledge is instantly available, will be decreased importance knowing a domain and the much greater importance of knowing how to learn efficiently from the overwhelming abundance of knowledge. That is the larger design frame that guides this work. How can we develop in students the best skills and dispositions for effective learning?