Physics faculty beliefs and values about the teaching and learning of problem solving. II. Procedures for measurement and analysis

Charles Henderson[*]
Western Michigan University, Kalamazoo, Michigan 49008, USA

Edit Yerushalmi[*]
Weizmann Institute of Science, Rehovot, Israel, 76100

Vince H. Kuo,[†] Kenneth Heller, and Patricia Heller
University of Minnesota, Minneapolis, Minnesota 55455, USA
(Received 28 October 2005; published 12 December 2007)

To identify and describe the basis upon which instructors make curricular and pedagogical decisions, we have developed an artifact-based interview and an analysis technique based on multilayered concept maps. The policy capturing technique used in the interview asks instructors to make judgments about concrete instructional artifacts similar to those they likely encounter in their teaching environment. The analysis procedure alternatively employs both an *a priori* systems view analysis and an emergent categorization to construct a multilayered concept map, which is a hierarchically arranged set of concept maps where child maps include more details than parent maps. Although our goal was to develop a model of physics faculty beliefs about the teaching and learning of problem solving in the context of an introductory calculus-based physics course, the techniques described here are applicable to a variety of situations in which instructors make decisions that influence teaching and learning.

I. INTRODUCTION

Even the most effective curricula materials, tools, and pedagogies must be accepted by instructors if they are to have an impact on students. This is particularly true in higher education, where academic freedom gives faculty members a great deal of autonomy in choosing both a curriculum and its teaching. In spite of this freedom, however, there appears to be a uniform and stable set of educational practices in the teaching of introductory college physics that have been characterized as traditional instruction.[1] New ideas, especially those based on research in learning, must compete against these long-standing traditional methods. In terms of the marketplace, improved educational products must gain customer acceptance if they are to displace widespread traditional practices. In the commercial world, product designers know that to gain market share, they must construct their product to match the characteristics that influence customer choices. Following this model, the first step of effective dissemination of curriculum or pedagogy developed by physics education research must occur in the design process. Incorporating the beliefs and values as well as perceived needs and skills that underlie an instructor's choice of curriculum and pedagogy is a prerequisite to achieving a design that can be efficiently disseminated.

The few existing investigations into the beliefs[2] about teaching and learning held by college physics instructors have reported conflicting results. Some studies have found that instructors have beliefs that have been called "traditional"—they see teaching as transmitting information from themselves to students and learning as receiving this information.[3,4] Other studies have found that physics instructors see learning as students actively constructing understanding.[5–8] These instructors realize that teaching by telling is largely ineffective. The range of results suggests that there may be methodological difficulties in at least some of the measurements of faculty beliefs.

Modern consumer research uses contextual inquiry and ethnographic techniques to observe and interview people in natural settings doing normal activities.[9–15] This procedure has proved to be superior to older techniques that directly asked potential customers about their desires and needs, or simply observed them accomplishing some task. Directly asking, whether by interview or questionnaire, does not uncover important consumer characteristics, primarily because many issues are not a part of the interviewee's consciousness.[16] In addition, researchers have difficulty correctly interpreting either direct responses or observed actions because they cannot know the internal mental context in which they are given.[10]

In education, the type of artifact-based interview that is the basis of modern consumer research has long been used to create a natural context for exploring the ideas of students.[17–20] This type of interview is especially useful when the interviewee does not share a formal system of language with the interviewer. For example, an interview study found that some children would tell interviewers that the earth is like a ball. When asked to draw the ball, however, it became clear that these students conceived of the earth as having two hemispheres with the lower composed of rocks and dirt and the upper composed of air. People lived on the flat surface of the lower hemisphere.[20] Based simply on the language of the interview, one might erroneously conclude that a curriculum for these children could be based on the assumption that they have the accepted idea of a spherical earth.

Drawing on the tradition of artifact-based interviews in education and commercial product design we have developed a tool to probe physics faculty beliefs and values related to the teaching and learning of problem solving in the context of an introductory physics course. This paper describes the research procedure whose results are described for a study of six university physics faculty in the companion paper.[8] The goal of this paper is to make our measurement and analysis process as explicit and transparent as possible to begin the process of arriving at a better understanding of the physics instructor as a key step to any process of educational reform.

II. INSTRUCTIONAL SYSTEM

The goal of our study was to create a testable model of faculty beliefs about the people, things, activities, and environment related to the teaching and learning of problem solving in an introductory physics course for the purpose of documenting and ultimately influencing their instructional choices. The detailed motivation for this study and the research questions that framed it are given in the companion paper.[8] We chose the first semester, mechanics, of the calculus-based introductory physics course as the system of interest. The interview and its analysis were based on the three generic features of such an instructional system: (1) the initial condition of the students, (2) the desired final state of the students, and (3) the processes necessary to move students from 1 to 2.[19] This systems approach commonly underlies a research-based curriculum that emphasizes physics problem solving.[22–26] Instructional choices to accomplish this transition are affected by the instructor's beliefs about all three features as well as their values and their perception of any constraints external to this process. Since we are interested in the reasons that an instructor either uses, changes, or rejects curricular material, the investigation concentrates on the instructor's mental environment rather than actions.

III. DATA COLLECTION TOOL

A. Artifact-based interview

The data collection tool was designed to probe instructors' responses to a set of realistic teaching situations. This policy-capturing technique[27] asks instructors to make judgments about instructional artifacts similar to those in their teaching environment.[28–30] The artifacts are designed to create a context which would activate beliefs that could influence decisions when instructors select instructional material or pedagogical techniques. The technique allows us to standardize the data collection environment for multiple instructors.

In order to discriminate between a strongly held belief structure and fragmented or even conflicting beliefs, our interview had four different situations that emphasized different aspects of the same instructional goal. The artifacts were designed to encourage instructor introspection and elicit ideas that an instructor might not articulate spontaneously. Artifacts varied in their representation of aspects of expert and novice problem solving as well as the features of three

basic instructional paradigms: behavior, development, and apprenticeship.[31] These artifacts were also designed to allow one to test the consistency of ideas among the situations, a technique called triangulation, to facilitate the identification of strong conceptions. For example, one aspect of expert problem solving is planning. Each set of artifacts, given in the Auxiliary Appendix, was designed so that some explicitly represented planning (e.g., Student Solution D, Instructor Solution C) and some did not (e.g., Student Solution E, Instructor Solution A). Finally, the use of artifacts allows the instructors and the interviewer to use the instructors' natural language anchored by specifics in the artifacts rather than imposing the researchers' language with its inherent risk of biasing the data.

B. Specific artifacts

The specific types of artifacts evoke instructional situations familiar to any physics instructor. The three types of instructional artifacts were:

1. *Instructor solutions*: The instructor makes problem solutions available to students, in either verbal (e.g., during lecture) or written form. The artifacts are three possible instructor solutions for a single problem.
2. *Student solutions*: The instructor evaluates student solutions. The artifacts are five possible student solutions for a single problem.
3. *Problems*: The instructor selects or creates problems for students to solve. The artifacts are five possible statements of a single problem.

Each set of artifacts spanned the three instructional paradigms in ways that represented common practices. For the purpose of an interview, only a small number of short, concrete artifacts could be used. They are shown in the Appendix and briefly explained below.

1. Basing the interview tool on one physics problem

All of the artifacts were based on a single problem. This Core Problem (see auxiliary Appendix) was selected to be one that could reasonably be given in most calculus-based introductory physics courses while being rich enough to allow for interesting variations. Its solution required several important physics concepts (e.g., linear and circular kinematics, dynamics, and conservation of energy), and students could potentially work the problem in different ways. The problem was taken from a final exam for the introductory calculus-based physics course at the University of Minnesota. Since multiple lecture sections of the same course have a common final exam, this problem was designed and approved by a group of four physics instructors who taught the course. In addition, the student final exam solutions were available, providing a source of authentic student solutions.

The core problem was sent to physics faculty at several other institutions to verify that it could conceivably be given to their students. All reported that the problem was difficult for an introductory physics course, but one that they could reasonably assign to their students. It was important that the

TABLE I. Characteristics of student solution artifacts

	Student solution				
	A	B	C	D	E
Qualitative reasoning	No evidence	Evidence	Evidence	Evidence	No evidence
Organization	Disorganized	Somewhat logical	Logical	Logical	Unclear organization
Final answer	Incorrect	Incorrect	Incorrect	Correct	Correct
Diagram	No	Yes	Yes	Yes	No
Reasoning	Unclear	Unclear	Clearly incorrect	Clearly incorrect	Unclear
Evaluation of final answer	No	No	Yes	No	No
Correct math	Yes	No	Yes	Unclear	Yes
Evidence of hierarchical knowledge organization	None	Unclear	Solution could be based on fundamental principles	Solution based on fundamental principles	None
Evidence of metacognition	None	Categorizes problem	Planning via subproblems, reflection on progress	Planning via subproblems	None

problem be considered difficult enough by an instructor to require an average student to use an exploratory decision making process as opposed to an algorithmic procedure.[21,22] Although each artifact reflected certain instructional and learning styles, as discussed below, none were designed to be flawless in order to facilitate the interview discussion.

2. Instructor solutions

These artifacts consisted of three different Instructor Solutions to the Core Problem. A review of posted instructor solutions found that most were of two basic types. The first type is a brief, "bare-bones" solution that offers little description or commentary. This type of solution leaves many of the minor steps to be filled in by the reader and is found in many textbook solution manuals. Instructor Solution 1 was modeled after this type of solution. The other common type of solution was more descriptive than the bare-bones type of solution. In this type of solution all of the details of the solution were explicitly written out. Instructor Solution 2 was modeled after this type of solution. Although providing a good representation of the range of actual instructor solutions, these first two instructor solutions were missing two aspects, characteristic of expert decision making,[22] that are recommended by some curriculum developers.[24–26] First, both of the previously described solutions proceed from the given information to the problem goal. Research has shown that expert problem solvers often begin with the problem goal and attempt to relate it to the known information.[22] Another missing aspect was an indication of the reasoning that motivated the particular steps of the problem solution by

describing an approach to the solution. Instructor Solution 3 contains these aspects.

3. Student solutions

The artifacts consisted of five different student solutions to the Core Problem. These were designed to reflect both common components found in an analysis of several hundred student final exam solutions and of the differences between expert and novice problem solving from the research literature.[21,22,32,33] The solutions reflected the dichotomies shown in Table I. Most features were triangulated in at least two artifacts. These artifacts were designed to trigger thoughts about different instructional approaches and their impact on the instructor and student. They have also been used as the basis of a previous paper on grading practices.[34]

Problems. These artifacts were four different restatements of the Core Problem based on an analysis of problem types used by physics instructors on tests. They reflected the research literature about problems[21,22] and were designed to trigger thoughts about the distribution of roles between instructor and student and the desired and perceived reasoning of students. Problem A includes a visualization of the problem, explicitly states the target of the solution, and is broken into three explicit subproblems that provide a guide to the solution. Problem B does not include a visualization of the problem and requires students to determine their subproblems. It has multiple-choice numerical answers that result from common conceptual difficulties. Problem C does not include a visualization of the problem, requires students to determine their subproblems, simulates a "realistic" context,

does not pose an explicit question, and is overly verbose. It also does not use SI units. Problem D includes a visualization of the problem, explicitly states the target of the solution, and adds three explicit subproblems that guide a qualitative analysis of the problem.

C. The interview protocol

To reflect on the artifacts, it was necessary for the instructors to solve the problem. During pilot testing, we found that, under interview conditions, many physics faculty made mistakes while solving the problem. For this reason, the problem was given to the instructors several days before the interview and they were asked to solve it. This was done by all of the interviewees. During the interview, a tripod-mounted video camera made an audio record of the conversation and visually recorded any pointing at features of the artifacts.

The interview protocol was designed to probe instructors' beliefs across the three instructional situations using both general and specific questions related to the elements that compose the instructional system. The auxiliary Appendix gives the interview script. The interview protocol avoided, as much as possible, imposing any language on the interview by using the artifacts to elicit the instructor's natural language. Once elicited and clarified, this language was then used by the interviewer. The use of a structured protocol reduces the likelihood that the conceptual frameworks of individual interviewers will bias interviewee responses. Data from different levels of generality across different contexts permitted a significant degree of triangulation in the analysis process. The interview protocol allowed us to complete the data collection for a single instructor in 1.5 h—the maximum amount of time that we thought most physics faculty would agree to be interviewed.

The interview was broken into four distinct parts. The first three parts each dealt with one of the three types of instructional artifacts and focused primarily on how the artifacts might be used during the instructor's course. The intent of these parts of the interview was to probe both the perceived and desired interactions between the students and instructor within the learning system. The fourth part of the interview was intended to focus on the instructor's perceptions of the students' initial and final state of knowledge about problem solving, the instructor's beliefs about the nature of problem solving, and the instructor's satisfaction with their instruction. This part of the interview used a fourth artifact generated by the interview process itself and described below. All of the interview questions fit into one of three basic categories: (1) general, open-ended questions designed to gather information about an instructor's ideas;[35,36] (2) specific questions, often related to a specific artifact, to gather information about a simulated instructional decision—their ideas-in-use;[35,36] and (3) questions about the problem-solving process to gather information about the instructor's beliefs regarding student problem solving. The following paragraphs briefly summarize each part of the interview.

1. Parts 1–3: Using the standard artifacts

The first part of the interview dealt with the Instructor Solution artifacts. It began with general questions about how and why the instructor used example problem solutions. The interviewer then gave the three Instructor Solutions to the instructor, who was asked to compare them to the materials they used in classes and to explain their reasons for making those choices. This part concluded by asking the instructor to reflect on the important parts of the problem-solving process either as represented in or missing from the artifacts.

The second part of the interview dealt with the Student Solution artifacts and began with general questions about how and why the instructor grades student problem solutions. The interviewer then gave the instructor the student solutions and asked the instructor to assign a grade to each solution and to explain their reasoning. This part of the interview concluded by asking the instructor to reflect on their student's problem-solving process as represented or not represented in student solution artifacts.

The third part of the interview began by giving the instructor the problem artifacts and asking them to compare the problem statements to those they typically used. This part concluded by asking the instructor to reflect on what problem-solving processes were either required or not required by the different ways of posing a problem.

2. Part 4: Using an artifact created in the interview process

During each of the first three parts of the interview, the interviewer recorded each feature of the problem-solving process that the instructor mentioned, using the instructor's words. Each feature was written on an individual index card. The collection of these cards, which ranged from 13 to 25 for the different interviews, constituted the final artifact. The instructor was given the cards and asked to categorize them in any way they wished. The instructor was then asked to name each category and explain why the statements were in that category. Questions were asked regarding these categories to elicit the instructor's belief about their students' proficiency in each of the categories at the beginning and at the end of the introductory physics sequence. Finally, the interviewees were asked questions to elicit their satisfaction with the outcome of their current course. This type of card sorting task is used to capture the instructor's language and knowledge organization.[37]

From instructor feedback at the completion of the interview process, we believe that the interview helped the instructors clarify their own thinking about instruction. In addition, even though the instructors were reluctant to give up 1.5 h of their time, most commented that the interview went by very quickly.

IV. DATA ANALYSIS: CREATING MULTILAYERED CONCEPT MAPS

The goal of the analysis technique is to represent prevalent beliefs and values of the population of physics faculty engaged in making pedagogical decisions about teaching using problem solving in an introductory physics course. It is not designed to characterize an individual instructor. Analyzing the rich data derived from these interviews posed several challenges: (1) Efficiently representing the complex interrelationships among the instructor beliefs underlying the inter-

view narrative. (2) Characterizing instructors' beliefs, to the extent possible, within an instructional system while allowing an authentic depiction of those beliefs as represented by their discourse. (3) Describing both the agreement and diversity among instructors' ideas. (4) Providing a tool to trace the interpretation of data that allows testing the analysis process, auditing the study, and evaluating the results.[38]

Our methodology merges two qualitative research perspectives that, in their extreme interpretations, conflict. Creating a hybrid analysis technique has, however, been found to be fruitful in other studies.[39] Some qualitative methods emphasize allowing categories and connections to emerge from the interview data.[40–42] We were interested in these instructor generated categories and connections. Other methods analyze such data in terms of *a priori* constructs,[43] which in our case was an instructional systems approach to pedagogy. Our hybrid process alternatively employs an *a priori* systems view analysis and an emergent categorization analysis to construct multilayered concept maps. The *a priori* analysis essentially allows the construction of an initial framework of categories. This is followed by an emerging category analysis that is used to elaborate or even change the initial framework. This process is described in detail below.

Concept maps[44] compactly represent a large number of interconnections. A multilayered concept map is a hierarchically arranged set of maps where child maps include more details than parent maps. Concept maps are a useful analysis tool because the visual and explicit nature of the connections between concepts forces the research team to discuss, confront, and clarify their diverse interpretations. In contrast to a situation where concepts are addressed sequentially, concepts on the map are connected to other concepts and derive their meaning from these interconnections. This construction process often reveals different researcher interpretations of an idea or term that requires clarification for the purposes of the analysis. The arrangement of the multiple layers provides a technique of organizing data from all of the interviewees that preserves the connection to original statements so that the interpretations of those statements can be tested. The visual nature of the concept maps provided a relatively straightforward comparison of individual and composite maps that allowed the extraction of any shared ideas.

In the following sections we will describe the analysis process in detail. First, we will explain the construction of statements from interview transcripts. Then, we describe the construction of a tentative main map that reflects the principal categories regarding the learning and teaching of problem solving and the connections between these categories that we found prominent in the transcripts. Next, we constructed child maps that included details of each instructor's beliefs about every principal category on the main map. Finally, composite maps were constructed representing commonalities and differences among instructors. Each stage includes a validation process that resulted in the iterative revision of the previous analysis stage.

A. Transcripts→statements

The audio portion of the videotape from each interview was transcribed into approximately 30 pages of text by a professional transcriber unrelated to this research. The researchers then checked the transcripts by viewing the videotapes and inserted references to visual information when needed to clarify the text. A qualitative analysis requires that this interview data be broken into smaller units for further analysis.[43] We chose the unit of analysis to be a statement,[45] each representing a single idea related to one of the research questions described in the companion paper.[8]

Each statement was designed to stand alone and ranged in size from short, three word sentences to more complex sets of three or four sentences. The words used by the participant were always retained. Occasionally extraneous words were removed for clarity, with this removal noted in the statement. More frequently, related discourse necessary to understand the context of the statement was added to the statement in parentheses. Each of the interviews yielded approximately 400 statements that were numbered sequentially for reference. Table II shows a portion of one instructor's transcript and the resulting statements.

Because making statements involves researcher interpretation, it was important that the reliability of this process be monitored. All of the statements used in this study were created by two of the authors. Initially both extracted statements from the same transcript passages and compared their work. Upon comparison, differences were discussed with the entire research team. This process allowed the identification of individual biases and the clarification of the statement-making procedure. The criterion for agreement was that the extracted statements convey the same information. Initially, the statements agreed at about the 70% level. By the end of the first transcript, the statements for the entire transcript agreed at the 86% level before discussion and at the 100% level after discussion. By the end of the second transcript, the statements for the entire transcript agreed at the 93% level before discussion and at the 100% level after discussion. This prediscussion level of agreement was considered acceptable and the remaining transcripts were broken into statements by only one of the researchers.

B. Transcript→tentative main map

The main map consists of the principal categories of instructor beliefs and their interconnections. We began the analysis process in a top-down manner with a tentative main map that represented prominent commonalities observed by the researchers in a holistic reading of the transcripts from the *a priori* perspective of an instructional system. Principal categories were represented by boxes and connections by links between the boxes. Each link was labeled so that the boxes and their links formed meaningful sentences. To ensure the credibility[46] of the map, the entire research team discussed all of the initial candidate main maps to falsify categories or links based on specific evidence from the transcripts. This process was continued until consensus was reached on a tentative main map that could be used in the subsequent analysis.

At this stage of the analysis, it was possible that discourse

TABLE II. An example of statements made from a portion of the interview transcript.

Interviewer: "Ok. I want to talk about two different kinds of students. And looking at your chart [created earlier in the interview] there's students that come in knowing stuff, which is great. But there are also a lot of students who don't come in being able to handle these areas. And of those students that come in without being able to handle them, some of the students get better and some of the students don't. So I'm wondering what the difference between those two types of students is—the students who improve during the class and the students that don't."

Transcript	Statement No.	Statement
Instructor 6: "Well, I mean, there's certainly a lot of categories. First of all, there's the ones that just don't care, that aren't gonna get any better. And of course, there's the other extreme, the people that really have the intelligence and motivation to look into these things. I think problem solving in general is something that some people find fun, and some others don't. I mean, some people like going through, and I think probably most physicists are in it because they like doing it. And so I think the people that enjoy a challenge, that enjoy the idea of working these things out, and coming up with knowledge that they didn't have before. I mean, I think that's the sort of sense of wonder sort of thing. I think on the negative end of things there's a lot of people that just think all this stuff is just totally beyond them, they'll never be able to do it. And therefore they're not going to try. I think some people have a sort of feeling that if they're not going to be good at it, why worry about it. It's not going to be important for them. Here are these things about…there was a newspaper article that [name] used to have on his office a long, long time ago, which was some columnist saying, "why do we have to know algebra anyway? I never see any want ads for an algebra-doer!" or things like that. So some people, they have a tendency to disparage what they can't do. And so they won't care about it. I think that's the biggest problem with teaching these big general courses, is you get that bigger fraction that just don't care."	294	"Well, I mean, there's certainly a lot of categories."
	295	"First of all, there's the students that just don't care, that aren't gonna get any better."
	296	"And of course, there's the other extreme (as opposed to students who just don't care and aren't gonna get any better), the people that really have the intelligence and motivation to look into these things.
	297	"I think problem solving in general is something that some people find fun, and some others don't. (This is a difference between students who improve and students who don't.)."
	298	"(I think problem solving in general is something that some people find fun, and some others don't.) I mean, some people like going through, and I think probably most physicists are in it because they like doing it."
	299	"(I think problem solving in general is something that some people find fun, and some others don't.) And so I think the people that enjoy a challenge, that enjoy the idea of working these things out, and coming up with knowledge that they didn't have before (will improve in the class). I mean, I think that's the sort of sense of wonder sort of thing."
	300	"I think on the negative end of things there's a lot of students that just think all this stuff is just totally beyond them, they'll never be able to do it."
	301	"(I think on the negative end of things there's a lot of people that just think all this stuff is just totally beyond them, they'll never be able to do it.) And therefore they're not going to try. I think some people have a sort of feeling that if they're not going to be good at it, why worry about it. It's not going to be important for them."
	302	"Here are these things about…"
	303	"There was a newspaper article that [name] used to have on his office a long, long time ago, which was some columnist saying, "why do we have to know algebra anyway? I never see any want ads for an algebra-doer!" or things like that."
	304	"So some people, they have a tendency to disparage what they can't do. And so they won't care about it. I think that's the biggest problem with teaching these big general courses, is you get that bigger fraction that just don't care."

data existed for each instructor that related to all of the system features on the main map. It was also possible that certain system features might not be found in some instructors' discourse data, or that the data revealed categories that could not be incorporated into the system view. Another possibility was that most of the instructors' discourse data would not match the system view. Our initial main map is shown in Fig. 1. For example, in reading what instructors said about students' initial state, we realized that they all emphasized that they direct their instruction towards a partial group of the students. This feature can be seen prominently as "Some College Students" in our initial draft map (Fig. 1).
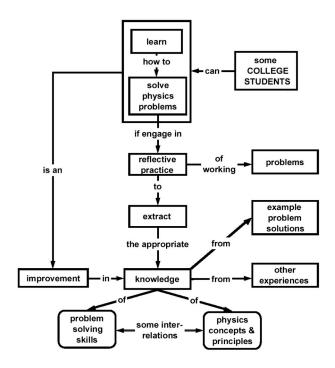
FIG. 1. Tentative main map developed at the beginning of the analysis process.

C. Tentative main map→individual child map for one instructor

The next step was to create child maps for a single instructor. A child map uses the interview statements to elaborate an instructor's beliefs about a principal category on the main map. This initial set of child maps was then used as the starting point for the analysis of additional instructors. The previous stage of the analysis that created the tentative main map is top down motivated by the system view. In this stage

data is analyzed from the bottom up to allow as many categories as possible to emerge from the statements in a manner guided by the general structure of the tentative main map. This confined approach to open coding[41] focuses the analysis on the research goals while allowing instructor generated categories and connections to emerge.

This analysis stage began with the sorting of all the statements related to a principal category on the tentative main map from the transcript of a single instructor. Using these statements, as many subcategories as possible were defined through open coding. These subcategories were grouped into higher order categories by identifying commonalities. For most maps, the final product was a list or tree structure such as the "Students Who Can Improve" map, Fig. 2. When the principal category was a process, the child map had a process structure such as the "Solve Physics Problems" map in the companion paper.[8]

The child maps also provide a direct link to the interview data. The numbered statement from the interview is recorded in each box representing a category built from that statement. Since the statements were numbered sequentially, it is easy to determine how many interview statements support a particular category and if those statements came from a single interview context. For example, Fig. 2 shows an individual's child map for the important feature of "Students Who Can Improve." The idea that "not caring is a characteristic that is detrimental to learning" is expressed in different contexts, once early in the interview (statement No. 87) and again later (statements No. 295 and No. 304).

Even this simple child map has implications for curriculum and professional developers. It shows that this instructor believes that only the middle group of students is affected by their instruction. Their efforts are not for the best students, contrary to a common assumption that physics faculty focus on "good" students who are most like themselves. In addition, they believe that students who have the intelligence to
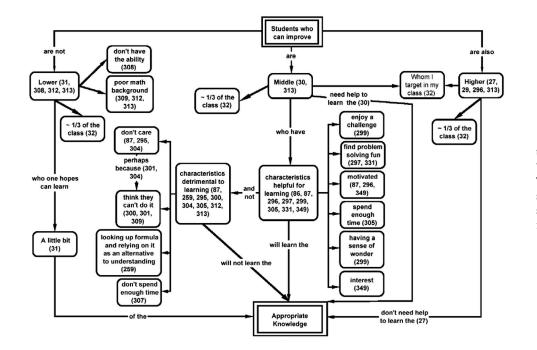

FIG. 2. Child map of instructor 6 for the principal category of "Students Who Can Improve." The numbers in parentheses are statement numbers from the transcript of this instructor that justifies the existence of this category.
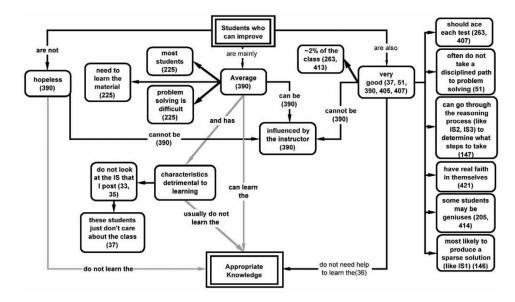
FIG. 3. Child map of instructor 3 for the principal category of "Students Who Can Improve." The numbers in parentheses are statement numbers from the transcript of this instructor that justifies the existence of this category. Gray links indicate that the connection was inferred by the researchers, from several statements.

learn but do not have the motivation, self-efficacy, or that use a plug-and-chug approach to problem solving are not the objective of their instruction. This belief is clearly at odds with many skill and attitude building, research-based curricula in physics.[1,21,47]

The open coding process has the potential to introduce systematic error due to researcher bias. For this reason, each initial child map was carefully validated by at least one additional researcher by returning to and reinterpreting the original statements used to build that map. Disagreements were discussed and most often resolved by these two researchers. When the two researchers could not resolve a disagreement, the entire research team discussed the issue. This discussion and reexamination of the transcript sometimes resulted in a revision of the principal categories on the main map and the consequent resorting of statements for the child

maps. These iterations of the main map, based on the bottom-up analysis used to create child maps, allows for categories to emerge on the main map that were not part of the initial *a priori* organization.

D. Individual child map for one instructor→individual child maps for all instructors

Child maps for additional instructors were created using the existing maps as a starting point. We began this process with a second instructor who appeared as different from the first instructor as possible. Using the first instructor's child maps, we attempted to place each of the second instructor's statements into the categories of the first instructor's map. If statements fit, we recorded their numbers in the appropriate box or along the appropriate link. Statements that did not fit
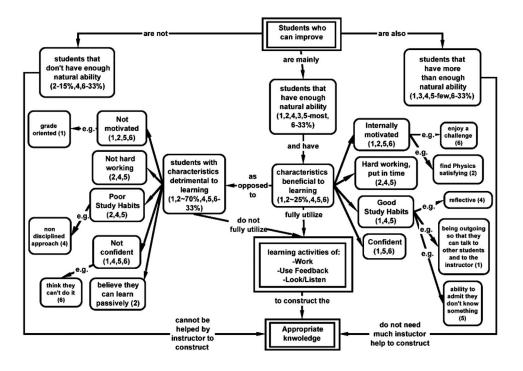

FIG. 4. Composite map for the principal category of "Students Who Can Improve." The numbers in parentheses represent the identification of the instructor who expressed that category.
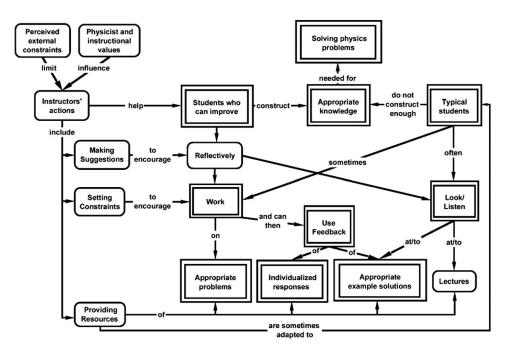
FIG. 5. Final main map resulting from the analysis process. Categories in double boxes have composite maps.

were used to create new categories. Statements from the first instructor were then reclassified into these new categories if possible. If such statements were found, the first instructor's child map was modified to include the new category or have the new links. If at the end of the process, boxes or links on the first instructor's map were not populated by any statements, then these aspects were deleted. In a similar way, child maps were constructed for each instructor and previous maps were modified. The validation process used in constructing these child maps was the same as that described above for the first child map.

The participant child map for Instructor 3 is shown in Fig. 3. Notice that the general structure of subcategories for Instructor 3 is very similar to those of Instructor 6, as described earlier. They are similar in the way that they define the average students, who do not improve, as being unmotivated to work hard, and not serious and systematic in their study habits. They differ in the amount of detailed statements they provide and in where those details fall.

This process facilitates the ultimate consolidation of each of the six different sets of child maps into single composite maps representing the similarities and differences of all of the instructors. It also strengthens all of the child maps since it requires the revision of existing child maps based on an emerging understanding of the data by the research team.

E. Individual child maps for all instructors→composite child maps and final main map

The final step was to construct a final main map and composite maps to represent the interview data of all the six instructors.

1. Composite maps

A composite map shows both the similarities and differences among the six instructors with respect to a single prin-

cipal category on the main map. The composite map, "Students Who Can Improve," is shown in Fig. 4; other composite maps from this study are shown in the companion paper.[8] The numbers in each box indicate which of the instructors expressed that idea. This makes it possible to see how many participants have beliefs consistent with a particular box or link. Links are only numbered when necessary for clarity. The wording used on the composite concept maps is the hypothesis of the research team designed to express the commonality and range of instructor beliefs.

These maps show the common instructor beliefs as well the range of differences. For example, the composite map "Students Who Can Improve" (Fig. 4) shows that all six instructors have the instructional belief that only the middle group of students in their class can be affected by their instruction. They differ, to some extent, on the definition of that group. Other composite child maps, however, show that instructors have qualitatively different beliefs. For example, the composite child map "Solve Physics Problems"[8] shows three distinct beliefs about the process of students solving physics problems.

Differences on the composite child maps can arise for several different reasons. The first reason is that the instructors actually have different beliefs. However, it is possible that a common belief justified by the statements of one instructor did not occur in the statements of another instructor because the interview was not adequate to trigger them. It is also possible that the process of qualitative analysis produced this difference artificially due to either underinterpretation or overinterpretation of a set of statements made by an interviewee. Making this belief structure explicit allows other investigators using other techniques to test both the specific similarities and differences among instructors.

The validation procedures used for the individual maps were also applied to the composite maps and resulted in iterations of the categories on both the child maps and the main map. For example, when developing the combined

TABLE III. Summary of methodological goals and solutions described in this paper.

Data collection	
Goal	Solution
Elicitation of beliefs that shape instructors' teaching decisions	Artifact-based interview tool probes instructors' responses to a set of realistic or semirealistic teaching situations. The artifacts are based on the research literature and on observations of common practice.
Discrimination between strongly held and fragmented or even conflicting beliefs	Artifact based interview tool has four different situations that emphasize different aspects of instruction.

Data analysis	
Goal	Solution
Comparison between beliefs underlying research-based curriculum and faculty beliefs	Confined open coding characterizes instructors' beliefs, to the extent possible, within the learning-teaching system. This system commonly underlies research-based curriculum. This type of coding also allows an authentic depiction of instructors' beliefs as represented by categories that emerge from their discourse.
Description of both the mutual core of instructors' ideas, as well as the diversity between instructors	Composite concept maps represent both the similarities and the differences between individual child maps.
Explicit communication tool required for verifying meaning in the analysis process, as well as allowing an external researcher to audit the study and evaluate the results	Multilayered concept maps provide a direct and transparent data trail between interviewee categories as represented on the concept maps and the raw data.
	Concept mapping forces an explicit categorization of interrelations in the data and provides a concrete anchor to clarify researcher interpretations.
Efficient presentation of results of the complex interrelationships among instructor beliefs	Concept maps have multiple layers that make it possible to use higher-order maps to navigate and quickly "zoom in" on an area of interest

child concept maps, the idea of "feedback" frequently occurred. This caused the research team to elevate the category to the main map, while differentiating it to two forms: "example problem solutions" and "individualized response."

2. Final main map

The final main map (Fig. 5) emerged from the analysis stages previously described. For example, the tentative main map shown in Fig. 1, which was based on an instructional systems approach, has only one category describing initial student characteristics. However, the analysis process uncovered that the interviewees talked about initial student characteristics in two distinct and disjointed ways. Student characteristics necessary for success were located in the "students that improve" category and represented the majority of the class to whom instruction was directed. At the same time, instructors believed that the typical students in the class did not have these characteristics. A second category, "the typical student," was created to describe the instructors' conflicting belief about the initial state of the student population, and was elevated to the main map. There was no evidence from the interview transcripts that the instructors recognized this dichotomy. An example of the evolution of the concept maps

as the analysis process progressed can be seen in the preliminary versions reported at conferences.[48–52]

Another example is the addition of the "reflectivity" box on the main map. Each of the learning activity child maps showed the instructor view that students can engage in each of the learning activities either reflectively or not reflectively and that this difference between student approaches to learning significantly altered the outcome. Thus reflectivity was promoted to the main map, yet the details of reflectivity are not found on a single "reflectivity" child map, but rather on each of the individual learning activity child maps.

V. SUMMARY

There are significant measurement and analysis challenges facing researchers investigating faculty beliefs that influence their choices of curriculum material and pedagogy. These instructors have a complex net of beliefs that can be context-specific and inconsistent. Because of the complex nature of this belief structure and how it is activated, it is not possible to deduce it from observing instructors' actions or from an open interview. Instead, we have developed a tool using an artifact based interview which uses multilayered concept maps as its primary analysis mechanism. It is impor-

tant to keep in mind that, although the concept map boxes and links come directly from the interview data, the maps themselves are a construct of the analysis process. They are a model of the beliefs of instructors in the context of making decisions about instructional material that should be tested by other means. Because the data represented on the concept maps result from instructor discourse in interview situations that represent common instructional contexts, we expect that these maps will be a useful representation of instructor beliefs and values that must be respected in any attempt at curricular change within these contexts.

The goal of this study is to construct a model, on an empirical basis, of the beliefs of physics faculty about the teaching and learning of problem solving in introductory physics that is definite enough to be tested by further research. To that end, this paper makes explicit the data collection and analysis methods used. Table III summarizes the design goals and data collection and data analysis tools. After the hypotheses represented by our concept maps are sufficiently tested and modified, we expect that the results of this line of research will be useful in assisting curriculum developers and those engaging in professional development to target their efforts more effectively to disseminate improved curriculum and pedagogy. Furthermore, we believe that this technique will prove to be useful for investigating the basis of instructor decisions in other domains of teaching and learning.

ACKNOWLEDGMENTS

APPENDIX: INTERVIEW ARTIFACTS AND PROTOCOL

See separate auxiliary material for interview artifacts and protocol.

*Previously at School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA.

†Present address: Department of Physics: Colorado School of Mines, Golden, CO, 80401, USA.

[1] E. F. Redish, *Teaching Physics with the Physics Suite* (Wiley, Hoboken, NJ, 2003).

[2] A. G. Thompson, in *Handbook of Research on Mathematics Teaching and Learning*, edited by D. A. Grouws (MacMillan, New York, 1992), pp. 127–146.

[3] M. Prosser and K. Trigwell, *Understanding Learning and Teaching: The Experience in Higher Education* (St. Edmundsbury Press, Great Britain, 1999).

[4] D. Kember and K. P. Kwan, in *Teacher Thinking, Beliefs and Knowledge in Higher Education*, edited by N. Hative and P. Goodyear (Kluwer, Dordrecht, The Netherlands, 2002), p. 219.

[5] K. Samuelowicz and J. D. Bain, Conceptions of teaching held by academic teachers, Res. Higher Educ. **24**, 93 (1992).

[6] J. G. Donald, Professors' and students' conceptualizations of the learning task in introductory physics courses, J. Res. Sci. Teach. **30**, 905 (1993).

[7] C. Henderson and M. Dancy, in *Proceedings (peer reviewed) of the 2004 AAPT Physics Education Research Conference*, edited by S. Franklin, J. Marx, and P. Heron (American Institute of Physics, Melville, NY, 2005).

[8] E. Yerushalmi, C. Henderson, K. Heller, P. Heller, and V. Kuo, Phys. Rev. ST Phys. Educ. Res. **3**, 020109 (2007).

[9] L. Anschuetz and S. Rosenbaum, in *Proceedings of 2003 Conference for Human-Computer Interaction* (Association for Computing Machinery, Ft. Lauderdale, FL, 2003).

[10] J. Simonsen and F. Kensing, Commun. ACM **40**, 82 (1997).

[11] D. Gilmore, Business: Understanding and overcoming resistance to enthographic design research, Interactions **9**, 29 (2002).

[12] M. E. Raven and A. Flanders, Using contextual inquiry to learn about your audiences, SIGDOC Asterisk J. Comput. Doc. **20**, 1 (1996).

[13] L. E. Wood, Semi-structured interviewing for user-centered design, Interactions **4**, 48 (1997).

[14] M. Ozer, A survey of new product evaluation models – Concepts. Methods and Strategy, Prod. Invent. Manage. J. **16**, 77 (1999).

[15] H. Beyer and K. Holtzblatt, *Contextual Design: Defining Customer-Centered Systems* (Morgan Kaufmann, San Francisco, 1998).

[16] S. Wilcox and W. Reese, Medical Device Diagnostic Ind.http://www.devicelink.com/mddi/01/09.html 68 (2001).

[17] J. Piaget, *The Child's Conception of the World* (Harcourt, Brace, New York, 1929).

[18] J. Piaget, *The Child's Conception of Movement and Speed* (Routledge, London, 1970).

[19] R. Driver and J. Easley, Stud. Sci. Educ. **5**, 61 (1978).

[20] R. Driver, E. Guense, and A. Tiberghien, *Children's Ideas in Science* (Open University Press, Buckingham, UK, 1985).

[21] F. Reif, Millikan Lecture 1994: Understanding and teaching important scientific thought processes, Am. J. Phys. **63**, 17 (1995).

[22] D. Maloney, in *Handbook of Research on Science Teaching and Learning*, edited by D. Gabel (MacMillan, New York, 1994).

[23] J. P. Mestre, R. J. Dufrense, W. J. Gerace *et al.*, Promoting skilled problem-solving behavior among beginning physics students, J. Res. Sci. Teach. **30**, 303 (1993).

[24] A. Van Heuvelen, Overview, Case Study Physics, Am. J. Phys. **59**, 898 (1991).

[25] P. Heller and M. Hollabaugh, Teaching problem solving through cooperative group. Part 2: Designing problems and structuring groups, Am. J. Phys. **60**, 637 (1992).

[26] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, Am. J. Phys. **60**, 627 (1992).

[27] R. J. Shavelson and P. Stern, Research on Teachers' Pedagogical Thoughts, Judgments, Decisions, and Behavior, Rev. Educ. Res.

51, 455 (1981).

[28] H. L. Swanson, J. E. O'Connor, and J. B. Cooney, An Information Processing Analysis of Expert and Novice Teachers' Problem Solving, Am. Educ. Res. J. **27**, 533 (1990).

[29] K. S. Cushing, D. S. Sabers, and D. C. Berliner, Olympic Gold: Investigations of Expertise in Teaching, Educ. Horiz. **70**, 108 (1992).

[30] C. LeMaistre, What is an expert instructional designer? Evidence of expert performance during formative evaluationEduc. Technol. Res. Dev. **46**, 21 (1998).

[31] Farnham-Diggory, Paradigms of Knowledge and Instruction, Rev. Educ. Res. **64**, 463 (1994).

[32] R. Zajchowski and J. Martin, Differences in the problem solving of stronger and weaker novices in physics: Knowledge, strategies, or knowledge structure?, J. Res. Sci. Teach. **30**, 459 (1993).

[33] A. Van Heuvelen, Learning to think like a physicist: A review of research-based instructional strategies, Am. J. Phys. **59**, 891 (1991).

[34] C. Henderson, E. Yerushalmi, V. Kuo *et al.*, Grading Student Problem Solutions: The Challenge of Sending a Consistent Message, Am. J. Phys. **72**, 164 (2004).

[35] R. J. Menges and W. C. Rando, What are your assumptions? Improving instruction by examining theories. College Teaching **37**, 54 (1989).

[36] C. Argyris and D. A. Schon, *Theory in Practice: Increasing Professional Effectiveness* (Jossey-Bass, San Francisco, 1974).

[37] M. T. Chi, P. J. Feltovich, and R. Glaser, Categorization and Representation of Physics Problems by Experts and Novices, Cogn. Sci. **5**, 121 (1981).

[38] J. W. Creswell, *Qualitative Inquiry and Research Design: Choosing Among Five Traditions* (Sage, Thousand Oaks, CA, 1998).

[39] S. B. Merriam, *Qualitative Research and Case Study Applications in Education* (Jossey-Bass, San Francisco, 1998).

[40] C. Moustakas, *Phenomenological Research Methods* (Sage, Thousand Oaks, CA, 1994).

[41] A. L. Strauss and J. Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques* (Sage, Newbury Park, CA, 1990).

[42] A. L. Strauss, *Qualitative Analysis for Social Scientists* (Cambridge University Press, Cambridge, 1987).

[43] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook* (Sage, Thousand Oaks, CA, 1994).

[44] J. D. Novak and D. B. Gowin, *Learning How to Learn* (Cambridge University Press, New York, 1984).

[45] R. H. Hycner, Some guidelines for the phenomenological analysis of interview data, Hum. Stud. **8**, 279 (1985).

[46] Y. S. Lincoln and E. G. Guba, *Naturalistic Inquiry* (Sage, Beverly Hills, CA, 1985).

[47] L. McDermott, Guest Comment: How we teach and how students learn—A mismatch?, Am. J. Phys. **61**, 295 (1993).

[48] C. Henderson, K. Heller, P. Heller, V. Kuo, and E. Yerushalmi, in *Proceedings of the 2001 Physics Education Research Conference*, edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, Loudonville, NY, 2001), pp. 79–82.

[49] C. Henderson, K. Heller, P. Heller, V. Kuo, and E. Yerushalmi, in *Proceedings of the 2002 Physics Education Research Conference*, edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, Loudonville, NY, 2002), pp. 31–34.

[50] V. Kuo, K. Heller, P. Heller, C. Henderson, and E. Yerushalmi, in *Proceedings of the 2001 Physics Education Research Conference*, edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, Loudonville, NY, 2001), pp. 87–90.

[51] V. Kuo, K. Heller, P. Heller, C. Henderson, and E. Yerushalmi, in *Proceedings of the 2002 Physics Education Research Conference*, edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, Loudonville, NY, 2002), pp. 35–38.

[52] P. Heller, K. Heller, C. Henderson, V. Kuo, and E. Yerushalmi, in *Proceedings of the 2001 Physics Education Research Conference*, edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, Loudonville, NY, 2001), pp. 71–74.