Concept Inventories in Computer Science for the Topic Discrete Mathematics

Vicki L. Almstrum
The University of Texas at Austin
Department of Computer Sciences
Austin, TX 78712
+1 512 471 9737
almstrum@cs.utexas.edu

Peter B. Henderson
Butler University
Department of Computer Science
Indianapolis, IN 46208
+1 317 253 7198
phenders@butler.edu

Valerie Harvey
Robert Morris University
Department of Computer & IS
Moon Township, PA 15108
harvey@rmu.edu

Cinda Heeren
University of Illinois
Department of Computer Science
Urbana, IL 61801
heeren@cs.uiuc.edu

William Marion
Valparaiso University
Department of Math & CS
Valparaiso, IN 46383
bill.marion@valpo.edu

Charles Riedesel
University of Nebraska - Lincoln
Computer Science & Engineering
Lincoln, NE 68588
riedesel@cse.unl.edu

Leen-Kiat Soh
University of Nebraska - Lincoln
Computer Science & Engineering
Lincoln, NE 68588
lksoh@cse.unl.edu

Allison Elliott Tew
Georgia Institute of Technology
College of Computing
Atlanta, GA 30332
allison@cc.gatech.edu

ABSTRACT

This report describes concept inventories, specialized assessment instruments that enable educational researchers to investigate student (mis)understandings of concepts in a particular domain. While students experience a concept inventory as a set of multiple-choice items taken as a test, this belies its purpose, its careful development, and its validation. A concept inventory is not intended to be a comprehensive instrument, but rather a tool that probes student comprehension of a carefully selected subset of concepts that give rise to the most common and pervasive mismodelings. The report explains how concept inventories have been developed and used in other STEM fields, then outlines a project to explore the feasibility of concept inventories in the computing field. We use the domain of discrete mathematics to illustrate a suggested plan of action.

Categories and Subject Descriptors

K.3.2 [Computer and Information Science Education]; G.2 [Discrete Mathematics]; F.4.1 [Mathematical Logic]

ITiCSE '06, June 26-28, 2006, Bologna, Italy.

General Terms

Measurement, Experimentation, Human Factors

Keywords

Concept Inventory, Assessment, Discrete Mathematics, Assessment Tools, Distractors, Misconceptions.

1. INTRODUCTION

What is assessment? A simple definition is that

> *Assessment is the process of documenting, oftentimes in measurable terms, knowledge, skills, attitudes and beliefs. Assessment is often used in an educational context, but applies to many other areas as well [52].*

Assessment has the potential to help improve the education of students, whether in computing or another field. Assessment should provide sufficiently comprehensive feedback for improvement yet should not be invasive to the point of being impractical. Tools such as assignments, quizzes, and exams are used primarily to evaluate student performance for assigning grades. In general, however, such instruments are not effective at measuring student understanding because these tools are typically based upon a single person's perception of the field in question and are not validated. All too often, instructors observe that their students lack a basic understanding of concepts from prerequisite course(s) despite having earned good grades. Indeed, fundamental misconceptions can be deep-rooted and difficult to correct without the use of carefully crafted tools. In other words, a learner's fundamental misconceptions, whether grounded in their studies or their life experiences, are often difficult to identify, address, and correct in a timely manner. In this report, we explore concept inventories, which are assessment tools designed to identify such misconceptions precisely.

The idea for concept inventories arose in the 1980s, when the physics education community recognized that students' common-sense misconceptions (in this case, about the physical world) *were not dispelled in introductory undergraduate physics courses* [20, 21]. Educators observed that students seemed to understand the relevant concepts, learn basic formulas, and show the ability to solve individual problems, yet were unable to utilize the fundamental concepts correctly to analyze common real-world situations.

Concept inventories have been found to be effective instruments for supporting educational reform in various domains of physics (e.g., force [23] and electromagnetic phenomena [45]) and other areas of science (e.g., [55]). These successes have spurred interest in concept inventories for the computing disciplines (including, but not limited to, computer science, information systems, information technology, and software engineering).

Generally speaking, a concept inventory (CI) is a validated assessment tool with a focus on misconceptions students may have about fundamental concepts in a domain [53]. Concept inventories are commonly created using a student-driven process that elicits students' perceptions and understandings about a given concept.

A concept inventory has the following characteristics:

- It is a reliable, validated assessment instrument.
- It focuses on common student misconceptions.
- It covers a specific domain, but is not a comprehensive instrument (i.e., it is not a final exam).
- It is composed of multiple-choice items.
- It is designed to require at most 30 minutes to complete.
- Its scope may or may not match the scope of the corresponding course. Thus, it could be necessary to develop more than one CI to cover the topics of a course such as CS1.
- It can be administered as a post-test (for example, at the end of the course). Some concept inventories can also be given as pretests, allowing "before and after" comparisons.
- It can be used by instructors to identify aspects of instruction that would benefit from change, to assess the impact of modifications, and to compare pedagogical approaches.
- It should *not* be used by college administrators to assess teaching performance or by individual instructors for determining grades for individual students.

Researchers who wish to develop a CI must consider a range of issues, including the methodology to be used in creating the instrument and ways in which the data generated by the tool can be used for understanding and improving instruction. A key research question is whether there are characteristics unique to computing education that will have a bearing on how these issues are resolved.

This Working Group report proposes a multi-year process for designing one or more CIs in the well-defined area of discrete mathematics. The primary purpose of this effort would be to investigate the feasibility of using CIs in computing education research, with a concept inventory for discrete mathematics as the starting point.

This report has been written with two goals: (1) to inform the computing education community about the history, nature, and development of CIs and (2) to suggest a process for developing, testing, and using CIs in computing education. Section 2 motivates the use of CIs, while Section 3 provides context by describing how CIs have been developed and used in other fields. Section 4 explains the anatomy of a CI, including terminology and definitions. Section 5 gives a general outline of the process for developing CIs. Section 6 discusses the choice of discrete mathematics as the domain of focus in these early explorations. Section 7 explores several challenges associated with developing CIs. Section 8 suggests a process for developing a CI in the domain of discrete mathematics, and Section 9 discusses how computing educators might use a completed discrete mathematics concept inventory. The report is supported by a pair of appendices. The first appendix lists some fundamental concepts in the area of discrete mathematics and sample misconceptions. The second appendix illustrates criteria for meaningful items for a discrete mathematics concept inventory by developing several candidate items.

2. MOTIVATION

The literature is replete with articles in which educators describe their struggle with one of the central issues in teaching: the quest for pedagogical techniques that enhance student understanding of material in the curriculum. Educators often find it difficult to understand student misperceptions of concepts [4, 29] and could therefore benefit from instruments that provide insights into areas with which students are struggling. Results from studies of the use of CIs within the physics education and engineering education communities suggest that similar instruments could be valuable tools for the computing education community as well.

The SIGCSE community has a strong tradition of sharing their ideas for meeting such challenges, both informally and in more formal conference settings (e.g., [3, 24, 39, 48]). This report points the way for a project that will continue this tradition.

3. HISTORY OF CONCEPT INVENTORIES

The general idea of concept inventories (CI) as assessment tools emerged from the physics education community nearly two decades ago. David Hestenes [23] and his graduate students at Arizona State University sought to determine the extent to which students applied their classroom exposure to physics (in particular in the area of mechanics) to understanding the real world. Their work resulted in an instrument called the Force Concept Inventory (FCI). The FCI gained visibility when a prominent Harvard physicist, Eric Mazur [36], used the FCI and demonstrated the pervasiveness of his own students' misconceptions. Richard Hake, Emeritus Professor of Physics from Indiana University in Bloomington, led the effort to validate this instrument with data from over 6000 students [18]. The FCI is now considered to be a reliable instrument and has been used for a variety of purposes. The FCI has had a significant role in on-going reform in physics education, for example, in supporting a move from lecture-centered instruction to more active, hands-on approaches.

The CI model has evolved and been applied by educators in many STEM (Science, Technology, Engineering, and Mathematics) communities to evaluate the effectiveness of teaching strategies and assess the quality of student learning. Researchers have created and are continuing to develop CIs for several fields, including electromagnetic waves [45], signals and systems [51], strength of materials [43], thermodynamics [38], materials science [30], statistics [49], heat transfer [26], fluid mechanics [35], chemistry [31], biology [7], electromagnetics [42], and circuits [22]. A coordinating force behind many of these efforts has been the Foundation Coalition [13], a group of educators dedicated to reform and innovation in engineering education. The area of concept inventories is one of seven that the Foundation Coalition is promoting to assist campuses that are engaged in efforts to improve their learning environments and curricula.

The efforts of researchers in these fields have helped in formalizing the cognitive focus of a CI and in fine-tuning the manner in which CIs are developed, validated, and deployed. By way of example, researchers at the Colorado School of Mines creating a CI for thermal and transport science went outside of the expected list of subject matter experts for their team by enlisting the services of Michelene Chi, a well-known cognitive psychologist [6].

Of particular interest to computing educators is the following work toward developing CIs in computing-related fields.

- The Computer Engineering Concept Inventory (CPECI) [37] project implemented a CI for use in a digital logic course. According to Jeff Jackson [25], one of the researchers on this project, CI consists of 26 items covering combinatorial and sequential circuits. It has been used for prerequisite assessment in subsequent courses, for the ABET assessment process, and for curriculum reform efforts.
- Longino, Loui, and Zilles [34, 56] have explored creating a CI for digital logic. This work identified a number of student misconceptions in digital logic and developed a draft assessment instrument as the basis for future work.
- This working group was proposed by the principal researchers on a newly funded National Science Foundation grant, which has as its main focus to develop a CI in the area of discrete mathematics [28]. The research team on this grant includes computing educators, an educational psychologist, subject matter experts (including four discrete mathematics textbook authors), and an advisory board of educators. Foundational work on this grant will build a community of computing educators interested in CI-based assessment. A panel at SIGCSE 2006 [2] and this ITiCSE 2006 working group are building blocks for this foundation, with the vision of a future coalition to share techniques and results among research teams developing concept inventories for computing-based disciplines.

The primary purpose for existing CIs has been to investigate the disparity between the concepts a student *should* be learning and what the student *actually* is learning. As such, instructors have used CIs as tools to evaluate student understanding relative to the goals of their courses. These assessment instruments have provided benchmarks for comparing the effect of innovative teaching techniques on student understanding of fundamental concepts (as distinct from achievement on examinations that test rote learning or the ability to plug values into formulas) within a single university environment. Indeed, CIs have provided direct feedback on student understanding of specific concepts for instructors who are using these tools as one component of their overall teaching assessment plan [11, 16].

4. ANATOMY OF A CONCEPT INVENTORY

This section explains the terms used to describe the structure of a concept inventory, and, in particular, of a CI item. A primary source for the wording in these definitions was [50]. The Working Group discussed alternatives to multiple-choice items, including multiple true-false (MTF) items [1], but eliminated them in favor of the multiple-choice format. Because the goal of a CI is to understand student misconceptions based on their responses, multiple-choice items provide a better basis for zeroing in on specific misconceptions.

The term *item* refers to one of the tasks (often referred to as *questions* or *problems*) on the instrument. Each item will include a stem (often referred to as a *statement* or *question*) and an n-tuple that includes the *key option* and (n-1) *distractors*. Note that this definition does not include explicit information about the ordering of the options.

A *multiple-choice item* is an objective test item with a statement (the *stem*) followed by a number of options as answers. The answer options are of two types:

- the *key option*, which is the correct choice
- some number of incorrect choice options, called *distractors*[1]

The *stem* can be viewed as the part of an objective test item that poses a question or sets a task for the student to complete.

An *instrument* is an examination paper or other form of assessment event constructed from one or more kinds of items and administered to a target group.

A *concept inventory* is an instrument that includes a small number of multiple-choice items designed to cover concepts from a particular domain. The goal for the project proposed in this report is to create an instrument that has the characteristics of a concept inventory and that is validated and reliable.

The development process for the concept inventory is likely to include the use of *open-ended questions* as a tool for identifying and determining the nature of students' misconceptions and the use of carefully designed *multiple-choice items* in the final instrument.

5. DEVELOPING A CONCEPT INVENTORY

As discussed in *Knowing What Students Know* [5], any tool designed to assess student learning must be built with an eye toward the following triad of concerns:

- *Cognitive model of student learning*: Any assessment must consider research on how, at what pace, and at what level of complexity students learn a concept.
- *Observations*: Items must be crafted to provide meaningful observations of student understanding (the data).
- *Interpretation*: Statistical analysis can be used to report whether or not the assessment measured what it was intended to measure.

---

[1] We have chosen to spell this term with the suffix "-or" rather than "-er". Both appear in the literature, but the Working Group found the spelling "distractor" better matched individual preferences as well as usage in some of our key sources, such as [19].

To develop a robust assessment instrument, researchers must consider these three concerns, perhaps iteratively, and consider how each influences the others.

The process of developing a concept inventory must be grounded in a cognitive model of student learning [5]. During development, the research team must observe student responses and interpret these responses with respect to the cognitive model. To optimize the results, the research team should include professionals with expertise in the domain area as well as experts in the areas of cognitive modeling and educational assessment. Because development is inevitably iterative (as discussed in sections 7 and 8) and carrying out the steps properly can be very time-consuming, the entire process of developing a validated CI is likely to be a multi-year endeavor.

In his overview of STEM CIs, Richardson [44] identified five broad activities that must be carried out in order to construct a concept inventory:

1. Determine the concepts.
2. Study the student learning process for those concepts.
3. Construct multiple-choice items.
4. Administer beta versions of the instrument to determine reliability and validity.
5. Revise the inventory to improve readability, reliability, validity, and fairness.

The remainder of this section explores each of these steps in turn.

To determine the concepts to be included in a concept inventory, researchers can survey domain experts to identify areas where students typically struggle while learning material. Because most CIs are designed to be completed in about 30 minutes, CI instruments are constrained to covering only a small number of concepts. For the Colorado School of Mines project [6], the cognitive psychologist on the research team suggested that while looking for student misperceptions to use in their CI developers should focus on emergent, rather than causal, principles. In other words, because students have a more difficult time understanding concepts for which the explanations do not follow from immediately observable events, a CI should focus on these less intuitive concepts.

Once the relevant set of concepts has been identified, the research team must devise situations that allow them to observe the processes students use to (mis)understand the concepts. These observations are synthesized over time into a cognitive model. Methods that researchers can use to probe student thought processes include individual interviews and focus group interviews.

In preparing to construct the multiple-choice items for a CI, researchers usually use a two-phase process to better understand the range of possible incorrect responses and the reasoning that leads students to these responses. In one approach, researchers first generate open-ended questions that focus on a single concept, then have students give written responses to the questions. Of primary interest are the incorrect responses, because these help reveal common misconceptions. In a complementary technique, the research team can interview students to probe why they believe these incorrect responses are correct. The misconceptions become the basis for designing distractors for the multiple-choice items, with the goal of having each incorrect response reveal clearly the mismodeling that led the student to give that response.

In pilot studies, the research team would administer beta versions of the CI to large numbers of students. Carried out correctly, this process allows researchers to establish the reliability and validity of the CI in identifying specific mismodelings. The research team can use results of the pilot studies to revise and improve items in the CI in order to improve its reliability and validity to the desired level. Reliability is concerned with ensuring consistency, that is, that students will answer items similarly if they take the CI more than once, while validity is concerned with whether the CI addresses the content accurately, that is, that the items truly are exposing the misconceptions they are designed to reveal [17].

Reliability can be established through statistical analysis of the results, while validity must be addressed throughout the development of the instrument. To ensure content validity, subject matter experts review both the process by which the CI has been created and the instrument itself to verify that the items adequately represent the domain [19]. Construct validity, which is concerned with ensuring that each item measures what it was intended to measure, is perhaps the most critical component and typically involves a number of activities [14, 17]. These activities can include the following: experts carefully reviewing the focus and nature of each item; statistical analyses to establish that the answers to items about similar concepts are similar; and analyses to correlate item scores with other measures. CI scores may also be compared with other established measures of student ability to establish criterion-related validity [14].

6. THE CASE FOR DISCRETE MATHEMATICS

While the long-term vision is to have well designed, validated concept inventories for many computing topics, the current project is limited to a single CI for the subject domain of discrete mathematics. This Working Group chose discrete mathematics, rather than topics from a course such as CS1, for several reasons.

First, topics from discrete mathematics are emerging as a stable component of computing curricula. The ACM/IEEE Task Force Report *Computing Curricula 2001* [27] includes topics from discrete mathematics among the core knowledge areas that every computer science student should learn. There is reason to believe that there is more consensus for typical discrete mathematics topics than for CS1-related or other course-related topics, thus offering wider applicability for a CI in the discrete mathematics domain.

Second, it may be possible to match the scope of a typical discrete mathematics course with the scope of a single CI. A key question for the research team is the level of granularity at which discrete mathematics concepts should be covered on the CI.

Third, the development of CIs in other domains, such as undergraduate physics, has demonstrated that it is very difficult to create multiple-choice items that fulfill the criteria of a concept inventory. Because the area of discrete mathematics has received significant attention from various groups within the computing education community, there is a useful body of work pertaining to the areas where students struggle. One example is the Math Thinking group [54], a loosely organized network of computing educators, whose website and discussion list focus on the role of mathematics and mathematical reasoning in computing curricula. Another example is the authors of discrete mathematics textbooks for students in computing-related programs, particularly highly experienced authors whose books have gone through several significant editions. Such resources promise to be useful in developing the initial questions used to probe for student misunderstandings and misconceptions.

Finally, many fundamental concepts from discrete mathematics are recognized as themes that permeate computing, sometimes embedded in other computing courses and sometimes offered as separate courses ([27], Section 7.4). Example topics include logical thinking, inductive thinking, recursive thinking, developing valid arguments, and creating abstractions. This gives hope that the experience gained in developing a CI for discrete mathematics can transfer naturally into the development of CIs for other computing topics.

7. ISSUES AND CHALLENGES

This section highlights a variety of issues that a research team must consider while developing a concept inventory. The goal of creating a CI is to have a validated assessment instrument that educators can use to probe how and why students answer items correctly and incorrectly. However, many factors can confuse the picture and prevent researchers from identifying the how and why, leading to incorrect identification of mismodelings.

7.1 Breadth of Concepts

One issue is the level of detail needed for a CI to cover the concepts of the domain adequately. If the goal were to develop a comprehensive assessment instrument, then a detailed taxonomy of relevant topics would be useful. However, CIs are, by nature, not comprehensive. The underlying philosophy is that using a limited instrument that has been validated is better than using none. A CI should be concise, yet accurate, and should probe into potential student mismodelings of fundamental concepts. At the same time, student performance on the CI should be consistent with performance on a more comprehensive instrument for the same domain. If the number of relevant fundamental misconceptions is large, researchers must consider the trade-offs between coverage and the length of time a student requires to complete the CI. In some circumstances, the best recourse may be to have multiple CIs, each focusing on a small subset of concepts.

7.2 Role of Vocabulary and Notation

Sometimes student difficulties arise from the definitions used for a concept. For example, in the area of mathematical logic students may think of the standard English definitions for the words 'or' and 'implies' rather than the formal logical meanings. Problems can also arise when mathematical notation is complex or relatively unfamiliar to the student. Other difficulties come from more fundamental misconceptions, even when students have been presented with good definitions. For example, students often have difficulties understanding the concepts of vacuousness and (uncountable) infinity.

The computing field is notorious for its dependence on notations and conventions (e.g., little-endian vs. big-endian). CI items can be written with different styles, for example notation-heavy, notation-light, or with concrete objects. In fields such as mathematics, the notation used to express a concept may be as important as understanding the concept itself (as with set operators). The critical point is that the manner in which an item is stated may influence how students interpret that item. A notation-heavy item may end up testing students both on their knowledge of the notation and their understanding of the concept. Appendix B, Part B.3, provides an example of this in the area of set theory.

Definitions and notation are generally important components of expected student learning in a course, so it is desirable to detect whether students have misconceptions about definitions, notation, or both. However, using a specialized vocabulary or notation in writing CI items may mask misconceptions of a more fundamental nature, particularly if the CI is used as a pretest (when students have not yet taken the course and may not yet know the vocabulary). This situation could be ameliorated if multiple items probe the same concept, yet use different styles.

7.3 Isolating Concepts of Concern

It can be difficult to write CI items that clearly focus on a single concept or topic. It might be that a candidate CI item is too broad and thus spans too many concepts. This would make it difficult to isolate specific misunderstanding(s). Some concepts can be approached from multiple directions, for example, in discrete mathematics some problems can be modeled either from the point of view of set theory or propositional logic. These characteristics make it more difficult to design CI items that clearly diagnose a student's mismodeling because it is more difficult to isolate the particular misconceptions.

7.4 Size Constraint of the CI

Because a CI must be brief enough to complete in about 30 minutes without time pressure, it can contain only a limited number of items. It is a challenge when designing a CI to ensure that it includes the smallest number of items necessary to reveal the most information about student misconceptions. Current practice in developing CIs appears to favor including from three to five items addressing each potential misconception.

7.5 Elusiveness of Good Distractors

Two separate issues determine whether a distractor is "good" or effective. First, each distractor should represent a single misconception. When chosen, there should be little or no ambiguity as to why a student chose that response, thus pinpointing the corresponding misconception. Second, each distractor should have drawing power, that is, students should feel enticed to select it if they have the associated misconception [17, 19]. Devising distractors for an item so that each distractor is comparably effective may prove very difficult.

7.6 Target Audience Differences

A significant challenge in creating any instrument for a diverse audience is the diversity itself. Relevant issues can be related to a range of student characteristics, such as whether they are native speakers of English or have English as a second language; whether they have a scientific or non-scientific background; and a variety of differences due to cultural, ethnic, or regional background.

While a key target audience for a topic such as discrete mathematics is students studying computing, a CI for discrete mathematics has a much broader potential audience. Subpopulations can be selected based on factors such as who teaches the material (e.g., mathematics instructors or computing instructors) and how the material is incorporated into the curriculum (e.g., as an independent course or integrated into computing-based courses). In addition, many other disciplines, such as mathematics, biology when focused on genetics, and economics for discrete modeling, embrace the material of discrete mathematics. While core discrete mathematics concepts are fundamental for numerous disciplines, developing and validating multiple instruments for various student audiences would require an excessive amount of effort. The Working Group suggests developing a single discrete math concept inventory as a means for exploring the usefulness of such instruments throughout computing education.

7.7 Capturing Student Mental Models

Donald Norman [41] distinguishes between mental models and conceptual models: "Conceptual models are devised as tools for the understanding or teaching of physical systems. Mental models are what people really have in their heads and what guides their use of things." It is vital for the research team to incorporate direct student input as they build the student mental model. This model then supports the creation of a well-grounded CI. While years of teaching experience in a topic such as discrete mathematics may tempt researchers to make expert guesses about the thought processes behind students' incorrect responses, the issues that lead students to answer an item correctly or incorrectly can differ from those an instructor might hypothesize as part of the conceptual model. Building the understanding of misconceptions upon genuine student input can help researchers avoid mis-identifying the reasons behind incorrect responses.

7.8 Need for an Iterative Process

The development of a CI will almost certainly be an iterative process with embedded feedback loops. There are likely to be iterations at every step, including developing a list of concepts, a set of viable open-ended questions, and a collection of high quality multiple-choice items, as well as in the process of beta testing, pilot testing, and field testing. At any point, the results may show the need for revision. This iterative process of designing, deploying, analyzing, and revising the CI serves not only to improve the instrument, but also to contribute to the knowledge base about common misconceptions and how best to probe student understanding of concepts.

8. PROPOSED PROCESS FOR CREATING A DMCI

In this section, the Working Group presents the development process for a Discrete Math Concept Inventory (DMCI) as it was conceived during the ITiCSE 2006 conference in Bologna, Italy. This description is meant to give a general sense of possible development scenarios. This report does not address issues relating to acquiring the monetary or human resources required to carry out such an effort. Considering the challenges discussed in the previous section, we anticipate that creating a validated CI instrument will take two to three years of effort, assuming at least five or six iterations through various versions of the instrument.

The group of people who will carry out the project is referred to as the *research team*. In parallel with the steps described below, the research team has several additional tasks:

- Develop the instructions that will be distributed to volunteers who assist during the early stages development.
- Establish a web presence to ensure that information about the instrument is readily available to interested individuals.
- Assist all participating faculty members in obtaining approval from their Institutional Review Boards.
- Create guidelines, which should include a discussion of uses and potential misuses; this document will accompany the DMCI after it is finalized.

Figure 1 illustrates the suggested DMCI development process. The following sections discuss the various steps in more detail, including major issues that arise and our initial response to each.

8.1 Identifying Fundamental Discrete Math Concepts

Because a focus of the DMCI is to investigate common student mismodelings when working with discrete mathematics concepts, the research team must identify a set of relevant concepts and the associated misconceptions. Appendix A lists fundamental discrete mathematics concepts with associated difficulties as a sample starting point.

The research team can choose any of a number of approaches to identifying student misconceptions [40]. Existing taxonomies, curricula, and syllabi can contribute to this understanding. Another approach is to get input from subject matter experts, that is, educators who have taught discrete mathematics for a long time. Authors who have written discrete mathematics textbooks released in several significant editions are yet another good source [9, 15, 47]. These authors have put serious thought into the fundamental concepts of discrete mathematics and the misconceptions students have about those concepts. For instance, Susanna Epp has published an article entitled "The role of logic in teaching proof" [8] and Kenneth Rosen has developed a list of common mistakes students make as a complement to his textbook [46]. Volunteer subject matter experts can create lists of concepts with associated misconceptions. Subject matter experts can also share test items, either multiple-choice or open-ended, that may yield useful insights into concepts and misconceptions.

As the master list of concepts develops, the concepts can be prioritized according to perceived importance, and misconceptions can be categorized based upon perceived difficulty (for example, on a 5-point scale or with respect to cognitive level). The resulting information can be used to generate an initial list of fundamental concepts and perceived relevant misconceptions.

8.2 Developing the Pilot Items

Using the list of concepts and misconceptions as a resource, the research team can create items designed to elicit student responses and thinking. These items will comprise a pilot instrument to be given to students in the process of identifying their actual misconceptions. At one end of the spectrum researchers could develop open-ended pilot items, while at the other end they could develop multiple-choice pilot items. An advantage of using open-ended items is that students can write responses in their own words, allowing researchers to discover misconceptions not considered earlier. A disadvantage of using open-ended items is that the analysis process can be fairly involved: researchers must begin with a content analysis on the responses in order to determine the set of relevant misconceptions for each item. Using multiple-choice items has the advantage of a much more straightforward analysis process: researchers need only conduct a frequency analysis on the number of times students chose each option. However, a key disadvantage of using multiple-choice items is that researchers must find a means to create a set of distractor options that includes the most likely misconceptions associated with each stem.

One approach to developing multiple-choice pilot items could be to enlist the assistance of subject matter experts, individuals who regularly teach the discrete math course. These experts could propose candidate CI items, perhaps providing the stem, the correct response, and one or more potential distractors along with the rationale behind the options. Another approach would be to ask the subject matter experts to propose open-ended items. If each expert focuses on a limited number of concepts, this could

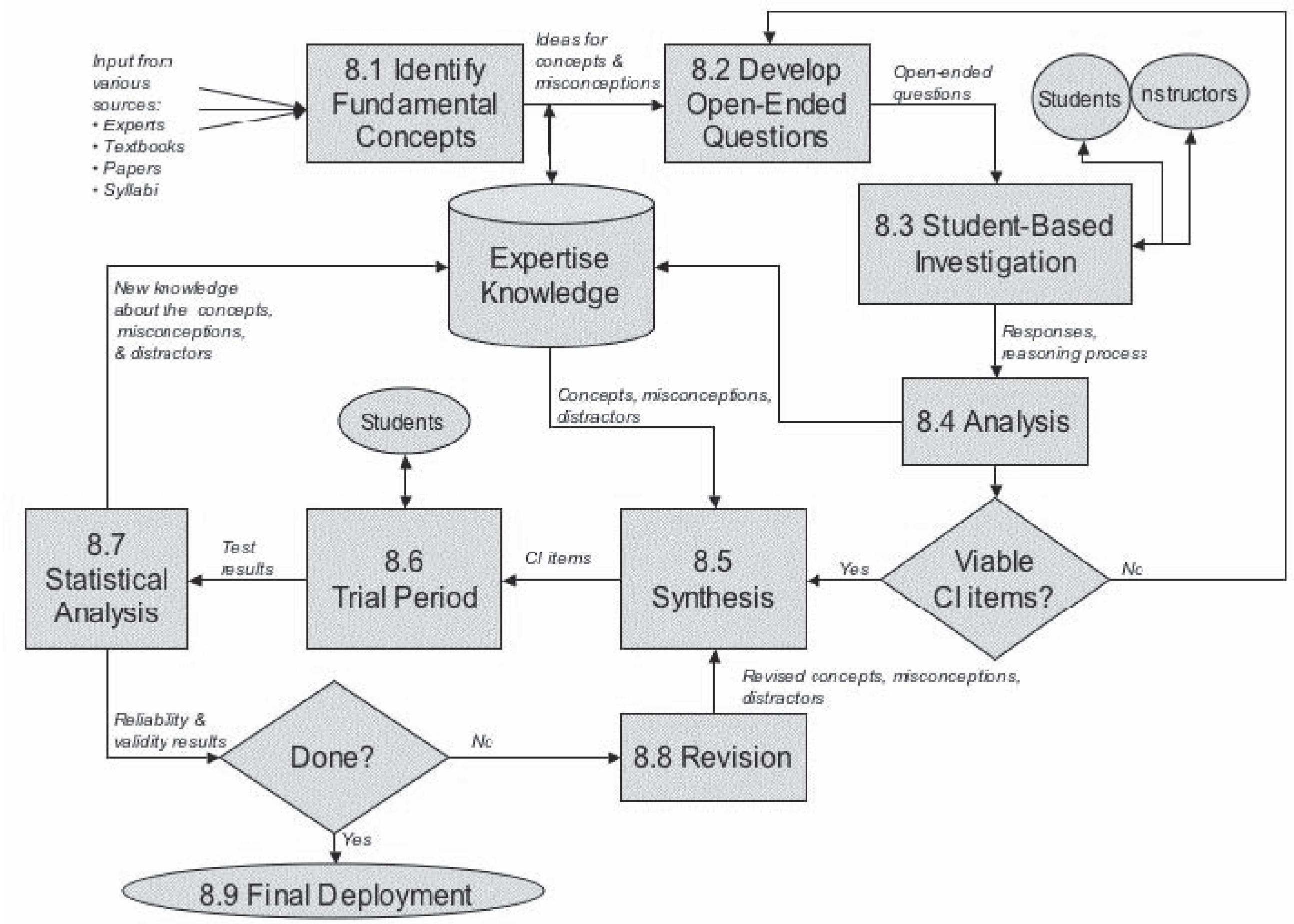

Figure 1: Proposed development process for discrete mathematics CI

result in a good spread of candidate items. Candidate items can then be reviewed by the researchers and perhaps by a panel of the subject matter experts. Appendix B includes example open-ended questions of the sort that might be generated during this phase.

The ultimate responsibility for creating the pilot items rests with the research team.

8.3 Preliminary Student-Based Investigation

In order to collect hard evidence about students' misconceptions, the research team can recruit instructors willing to administer the pilot instrument to a sampling of students. These instructors can be members of the research team, the subject matter experts mentioned in the previous section, and other volunteers willing to participate in the piloting process. Because students will be involved, the research team must ensure that all participating faculty members have obtained any necessary Institutional Review Board approvals.

An important part of this piloting process is to observe students' reasoning processes. One approach can be to ask students to justify their answers in writing as they complete the pilot instrument, although students might have difficulties expressing these explanations in a manner that is usable. A more effective approach can be to have direct interaction between members of the research team and individual students, although this can be time-consuming and may require specialized training to do well. The research team can apply techniques such as think-aloud [10], individual interviews [32], and focus group interviews [33]. A practical approach can be to use graduate students to conduct these interviews as part of their thesis or dissertation projects.

This stage may have to be repeated several times with different students at different sites before the research team converges on a set of items that can be used to generate a useful instrument.

8.4 Analysis of Investigation

For each concept, the research team must review the responses and thought traces coming from the student-based investigations. The goal should be to delineate the logic behind students' incorrect responses through a process called *deconstruction*. Each sample CI item in Appendix B includes a brief deconstruction discussion.

If the pilot items were open-ended, the analysis must include a content analysis of the prose students used to respond to each item. This content analysis will categorize the responses and build a set of misconceptions. Whether the pilot items were open-ended or multiple choice, the analysis must include a frequency analysis to reveal the set of misconceptions most likely to occur for each pilot item.

8.5 Synthesis

As the research team studies the results from the pilot items, one task will be to create stems for the candidate CI items. The data from the analysis of student-based investigations, along with suggestions from subject matter experts, will provide the basis for defining the pool of candidate distractors for each stem. The research team must narrow the set of distractors for each stem to a manageable level, retaining an optimal combination of options in terms of variety, clarity, lack of ambiguity, and conciseness (which should help prevent an item from being too time-consuming to complete). Each option must be expressed clearly and unambiguously, so that students can easily and quickly read and understand the item. At the end of this stage, the research team must consider whether the steps to this point have led to a pool of viable candidate CI items. A realistic goal is to have a minimum of three or four good distractor items per stem. The team must also consider the ordering of the options within each item.

Besides composing individual CI items, the research team must consider the overall composition of the full instrument. One consideration is how to order the items (e.g., randomly, topic-by-topic, easy-to-difficult, or according to a specific distribution algorithm) to provide a good flow as students complete the instrument. The ordering is significant, since the validation established in the next stage holds only for the version of the instrument subjected to pilot testing. Any variations must be validated separately.

8.6 Trial Period

By this stage, the research team will have created a candidate DMCI. Now the research team must methodically investigate the validity and reliability of the instrument. In order to make it feasible to engage a sufficient number of volunteer instructors for the investigation, the DMCI must be easy to administer and require little if any training to administer properly. The research team must establish a protocol for collecting the results in a timely, efficient fashion.

During the trial period, the research team will distribute the instrument to participating faculty, who in turn will administer the instrument to a diverse set of students. Because students will be involved, the research team must once again ensure that the participating faculty members have obtained the necessary Institutional Review Board approval.

8.7 Statistical Analysis

Once there is sufficient data collected from administering the candidate DMCI, the research team must conduct statistical analysis to determine reliability. It is beyond the scope of this Working Group report to specify the details of this analysis. The research team must consider validity throughout the process (see discussion in Section 5). Once the results meet pre-established criteria regarding validity and reliability, as well as issues such as readability and fairness, the instrument will be ready for final deployment (Section 8.9). Otherwise, the next step will be revision (Section 8.8), perhaps in several rounds.

8.8 Revision

Depending upon the results of the statistical analysis, some of the preceding steps may require repetition. The research team may decide to rephrase, replace, or eliminate some of the distractors if reliability cannot be assured or if the distribution of responses differs from desired patterns. If there is excessive redundancy or if the outcomes do not profile the concepts appropriately, ineffective CI items can be adjusted or replaced. This would lead to a new round of pilot testing in order to determine the reliability of the modified instrument.

8.9 Final Deployment

Once the DMCI has been deemed ready for broad use, the research team can publicize it and make it available to interested faculty and administrators. The instrument should be packaged with guidelines related to its uses and potential misuses. The delivery mode for the DMCI can be on paper or online as long as adequate controls are in place to ensure the security of the instrument.

The final version of the DMCI represents a fixed instrument that must be kept secure in order to ensure it is used only in authorized settings. Validation concerns the instrument as a whole, with exactly the provided set of items in the given order, with the exact wording and the exact ordering of options. Any variations, whether to rearrange the items or even reword slightly, require a new round of careful testing to establish validity and reliability. This is one of the reasons developing a CI is a demanding and time-consuming process. Thus, even though evolution of the DMCI is inevitable, the fact that a new validation round is required for every set of changes emphasizes the importance of thorough and careful initial development.

9. FUTURE WORK WITH THE DMCI

At this point, one project to create a DMCI has been funded by the National Science Foundation [28] and the Working Group knows of two other relevant proposals that are still under review as of this writing. This report, rather than being prescriptive for such research groups, is intended to be descriptive in order to allow the Computing Education community to better understand the issues and possibilities inherent in concept inventories.

When completed, the Discrete Math Concept Inventory (DMCI) will provide the computing education community with a new standardized tool for broad use in computing education research. Building on the model established by the physics education community with the Force Concept Inventory (FCI), a long-term vision is to establish a normed national standard for performance on the DMCI based on thousands of scores. Then, regardless of the specific use, the DMCI performance of a select group of students could be measured against this norm, allowing objective conclusions regarding relative performance.

We list here three general examples:

- A key purpose of the DMCI is to give instructors insight into their students' understanding of fundamental discrete mathematics concepts. The experiences of physics educators who have employed the FCI indicate that instructors are likely to be surprised by their students' performance on the assessment. The DMCI is intended to stimulate thoughtful reconsideration of instruction for topics for which students have persistent misconceptions. Instructors can use this assessment tool to evaluate the impact of innovative instructional techniques by measuring student performance relative to a national norm, both with and without the teaching innovation. As an example, Fagen, et al [12] used the FCI to evaluate the use of Peer Instruction in introductory physics courses.
- The computing education community is concerned about the classroom learning experiences of students of varied histories and demographics. The DMCI can help assess, for example, whether different approaches to teaching and

learning discrete mathematics are more or less effective for students with various characteristics. The lessons gleaned from such research can be used to establish appropriate course prerequisites or to plan appropriate interventions for particular groups of students. In an interesting example of this application, researchers have concluded that a student's prior math experience is more strongly correlated to success on the Statistics CI than is their statistics experience [49].

- The DMCI can be used for studies of long-term student outcomes. Given sufficient longitudinal data, researchers would be able to determine whether the DMCI as a whole is predictive of student success in a computing major, or even if particular topics from discrete mathematics are predictive. Such results can be used to make programmatic changes and to increase opportunities for students to master specific topics throughout their undergraduate course of study.

As described in this report, the development process of the DMCI, including rigorous testing for validity and reliability, builds upon best practice in educational research. This would make the DMCI a new and welcome addition to the set of assessment tools researchers in computing education have available for evaluating pedagogy and student outcomes.

10. ACKNOWLEDGMENTS

The Working Group members gratefully acknowledge the participation of David Klappholz and Steven Condly, who collaborated with Vicki Almstrum and Peter Henderson to propose this working group initially but were unable to participate during the conference itself.

The Working Group acknowledges the support of the National Science Foundation through two different grants:

- NSF-DUE #0536270, with PI David Klappholz (Stevens Institute of Technology) and co-PI Vicki Almstrum (The University of Texas at Austin), which provided partial support for Peter Henderson's travel to *ITiCSE 2006*.
- NSF-DUE # 0512064, with PI Barbara Moskal (Colorado School of Mines), and co-PIs Mark Guzdial (Georgia Institute of Technology), Stephen Cooper (Saint Joseph's University, and Wanda Dann (Ithaca College)), which supported Allison Elliott Tew's travel to *ITiCSE 2006*.

Several of our institutions provided travel support, in particular the Dean's Office for the College of Arts and Sciences at Valparaiso University (for William Marion) and Butler University (for Peter Henderson).

The discussions held with *ITiCSE'2006* conference delegates, both during the Working Group session and in private conversations, were invaluable, and in particular we are grateful to the Working Group coordinator, Myles McNally, for his support before, during, and after the conference.

During the warm days of July and editing this report, we received valuable comments from David Klappholz, Judith Gersting, and Michael Loui. Thank you as well to John Clement for a late night phone call that helped us chase down an elusive reference.

Finally, we want to express our appreciation for the contributions of the anonymous referees. Their comments were very helpful in improving this report.

11. REFERENCES

[1] Albanese, M.A., Kent, T.H., & Whitney, D.R. (1977). A comparison of the difficulty, reliability and validity of complex multiple choice, multiple response and multiple true-false items," *Annual Conference on Research in Medical Education.* 16:105-110.

[2] Almstrum, V. L., Klappholz, D., Condly, S., and Clement, J. (2006). Are they learning what (we think) we're teaching? In *Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education* (Houston, Texas, USA, March 3–5, 2006). ACM Press, New York, NY, p. 195-196.

[3] Begel, A., Garcia, D.D., & Wolfman, S.A. (2004). Kinesthetic learning in the classroom, *SIGCSE '04: Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*, Norfolk, VA, ACM Press, pp. 183–184.

[4] Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How People Learn: Brain, Mind, Experience, and School* (expanded edition). Washington, D.C.: National Academy Press.

[5] Center for Education (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. The National Academies Press. Available: www.nap.edu/books/0309072727/html/

[6] Colorado School of Mines (2006). Developing an Outcomes Assessment Instrument for Identifying Engineering Student Misconceptions in Thermal and Transport Sciences. (NSF DUE - 0127806) [July 19, 2006]. Available: www.mines.edu/research/cee/Misconceptions.html

[7] Doxas, I., Garvin-Doxas, K., & Klymkowsky, M.W. (2006) *Bioliteracy Project Home Page*. July 2, 2006 [July 23, 2006]. Available: bioliteracy.net/.

[8] Epp, S. (2003). The role of logic in teaching proof, *The American Mathematical Monthly*, 110(10), pp 886-99.

[9] Epp, S. (2004). *Discrete Mathematics with Applications*, 3rd Edition, Brooks/Cole Publishing.

[10] Ericsson, K.A. & Simon, H.A. (1984). *Protocol Analysis: Verbal Reports as Data*, Cambridge, MA: MIT Press.

[11] Evans, D. L., moderator (2003). Panel session: Progress on concept inventory assessment tools, *FIE'2003: Proceedings of the 33rd Annual Frontiers in Education*, pp. 1–8.

[12] Fagen, A.P., Crouch, C.H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher*, *40*(4), pp. 206-209.

[13] The Foundation Coalition. (2006). *The Foundation Coalition Home Page*. [July 27, 2006]. Available: www.foundationcoalition.org/

[14] Gay, L.R. & Airasian, P. (2000). *Educational Research: Competencies for Analysis and Application*, 6th Edition, Upper Saddle River, NJ: Prentice-Hall, Inc.

[15] Gersting, J. (2007). *Mathematical Structures for Computer Science*, 6th Edition, W. H. Freeman Publishing.

[16] Gray, G.L., Evans, D., Cornwell, P., Costanzo, F. & Self, B. (2003). Towards a nationwide dynamics concept inventory assessment test, *Proceedings of the 2003 American Society*

for Engineering Education Annual Conference and Exposition, Nashville, TN.

[17] Gronlund, N.E. & Linn, R.L. (1990). *Measurement and Evaluation in Teaching*, Sixth Edition, New York, NY: Macmillan Publishing Company.

[18] Hake, R.R. (1998) Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, pp 64-74. Available: www.physics.indiana.edu/~sdi/ajpv3i.pdf

[19] Haladyna, T.M. (2004). *Developing and Validating Multiple-Choice Test Items*, 3rd Edition, Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[20] Halloun, I.A., & Hestenes, D. (1985a). Common sense concepts about motion, *American Journal of Physics*, *53*, pp.1056–1065.

[21] Halloun, I.A., & Hestenes, D. (1985b) The initial knowledge state of college physics students, *American Journal of Physics*, 53, 1043–1055.

[22] Helgeland, B. & Rancour, D. (2003). *Circuits Concept Inventory*, The Foundation Coalition, [July 19, 2006]. Available: www.foundationcoalition.org/home/keycomponents/concept/circuits.html

[23] Hestenes, D., Wells, M., & Swackhamer, G. (1992) Force Concept Inventory, *Physics Teacher*, *30*, pp. 141-158. Available modeling.asu.edu/R&E/Research.html

[24] Hundhausen, C.D. & Brown, J.L. (2005). Personalizing and discussing algorithms within CS1 studio experiences: an observational study. *ICER '05: Proceedings of the 2005 International Workshop on Computing Education Research*, Seattle, WA, ACM Press, pp. 45–56.

[25] Jackson, J. (2006). Private Communication. September 6.

[26] Jacobi, A., Martin, J., Mitchell, J., & Newell, T. (2003). A concept inventory for heat transfer. *33rd ASEE/IEEE Frontiers in Education Conference*, Boulder, CO.

[27] Joint Task Force on Computing Curricula, Computing Curricula 2001, *Journal on Educational Resources in Computing*, 1(3es), ACM Press, New York, NY.

[28] Klappholz, A.D. and Almstrum, V.L. (2006). *Collaborative Research: Discrete Mathematics in Computing Education*. National Science Foundation. Available: www.nsf.gov/awardsearch/showAward.do?AwardNumber=0536270

[29] Kolodner, J. (1997). Educational implications of analogy: A view from case-based reasoning, *American Psychologist*, 52(1), pp. 57-66.

[30] Krause, S., Decker, J. C., & Griffin, R. (2003). Using a materials concept inventory to assess conceptual gain in introductory materials engineering courses. *33rd ASEE/IEEE Frontiers in Education Conference*, Boulder, CO.

[31] Krause, S., Birk, J., Bauer, R., Jenkins, B., & Pavelich, M. J. (2004). Development, testing, and application of a chemistry concept inventory. *34th ASEE/IEEE Frontiers in Education Conference*, Savannah, GA.

[32] Kvale, S. (1996). *Inter Views: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage.

[33] Lewis, M. (1999). *Focus Group Interviews in Qualitative Research: A Review of the Literature*. August 24, 1999 [July 19, 2006]. Available: www.scu.edu.au/schools/gcm/ar/arr/arow/rlewis.html

[34] Longino, J.T. (2006). *Boolean blunders: identification and assessment of student misconceptions in a digital logic course*. Tech. Rep. UILU-ENG-06-2215 CRHC-06-11, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign.

[35] Martin, J. K., Mitchell, J., & Newell, T. (2004). Work in progress: Analysis of reliability of the Fluid Mechanics Concept Inventory. *34th ASEE/IEEE Frontiers in Education Conference*, Savannah, GA.

[36] Mazur, E. (1997). *Peer Instruction: A User's Manual*, Prentice Hall.

[37] Michel, H., Jackson, J., Fortier, P., & Liu, H. (2006). *Computer Engineering Concept Inventory*. The Foundation Coalition. July 23, 2006. Available: www.foundationcoalition.org/home/keycomponents/concept/computer.html

[38] Midkiff, K.C., Litzinger, T.A., & Evans, D.L. (2001). Development of engineering thermodynamics concept inventory instruments. *31st ASEE/IEEE Frontiers in Education Conference*, Reno, NV.

[39] Naps, T.L., Eagan, J.R., & Norton, L.L. (2000). JHAVÉ — an environment to actively engage students in Web-based algorithm visualizations. *SIGCSE'00: Proceedings of the 31st SIGCSE Technical Symposium on Computer Science Education*, Austin, TX, ACM Press, pp. 109–113.

[40] National Research Council, Committee on Undergraduate Science Education (1997). *Science Teaching Reconsidered: A Handbook*, Washington, D.C.: National Academy Press.

[41] Norman, D.A. (1983). Some Observations on Mental Models. In Gentner, D.R., Stevens, A.L. (Eds.) *Mental Models*. Lawrence Erlbaum Associates.

[42] Notaros, B.M. (2002). Concept inventory assessment instruments for electromagnetic education, *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, San Antonio, TX.

[43] Richardson, J., Steif, P., Morgan, J., & Dantzler, J. (2003). Development of a concept inventory for strength of materials. *33rd ASEE/IEEE Frontiers in Education Conference*, Boulder, CO.

[44] Richardson, J. (2004) Concept Inventories: Tools for uncovering STEM students' misconceptions. *Invention and Impact: Building Excellence in Undergraduate Science, Technology, Engineering and Mathematics (STEM) Education*. Available: www.aaas.org/publications/books_reports/CCLI/PDFs/02_AER_Richardson.pdf

[45] Roedel, R.J., El-Ghazaly, S., Rhoads, T.R., & El-Sharawy, E. (1998). The Wave Concepts Inventory: An assessment

tool for courses in electromagnetic engineering. *ASEE/IEEE Frontiers in Education Conference*.

[46] Rosen, K. (2003). Common Mistakes in Discrete Math. In *Discrete Mathematics and Its Applications Student Solutions Guide*, McGraw Hill Publishing. Available: www.mhhe.com/math/advmath/rosen/r5/student/mistakes .html

[47] Rosen, K. (2007). *Discrete Mathematics and It*s *Applications*, 6th Edition, McGraw Hill Publishing.

[48] Soh, L.K. (2006). Implementing the jigsaw model in CS1 closed labs. *ITiCSE '06: Proceedings of the 11th annual SIGCSE Conference on Innovation and Technology in Computer Science Education,* Bologna, Italy, ACM Press, pp.163–167.

[49] Stone, A., Allen, K., Rhoads, R.R., Murphy, T.J., Shehab, R.L., & Saha, C. (2003). The Statistics Concept Inventory: A pilot study. *33rd ASEE/IEEE Frontiers in Education Conference*, Boulder, CO.

[50] The Teaching Resources Compendium for Higher Education (2002). Glossary of Assessment Terms. [July 27, 2006]. Available: www.lmuaut.demon.co.uk/trc/edissues/assgloss.htm

[51] Wage, K.E. & Buck, J.R. (2001). Development of the Signals and Systems Concept Inventory (SSCI) assessment instrument. *31st ASEE/IEEE Frontiers in Education Conference*, Reno, NV.

[52] Wikipedia, “Assessment”, July 10, 2006 [July 19, 2006]. Available: en.wikipedia.org/wiki/Assessment

[53] Wikipedia, “Concept Inventory”, July 1, 2006 [July 19, 2006]. Available: en.wikipedia.org/wiki/Concept_inventory

[54] Working Group on Integrating Mathematical Reasoning into Computer Science Curricula. [July 26, 2006]. Available: www.math-in-cs.org/

[55] Wyckoff, S. (2001). Changing the culture of undergraduate science teaching. *Journal of College Science Teaching*. *30*(5), 306-312.

[56] Zilles, C., Longino, J., and Loui, M. (2006). Student misconceptions in an introductory digital logic design course. *Proceedings of the 2006 ASEE Annual Conference and Exposition*, Chicago, IL., June 18–21, 2006, CD-ROM.

APPENDIX A
Potential Fundamental Discrete Mathematics Concepts

The Working Group developed this list of fundamental discrete mathematics concepts to provide sample areas where students commonly experience difficulties. This list is by no means complete, as it lacks many notions that should probably appear on the completed Discrete Math Concept Inventory. The concepts are ordered alphabetically.

Discrete Mathematics Concept	**Discussion of Typical Student Difficulties**
Argument by contradiction	Difficulty understanding what it means to argue that the implication p ◊ q is false.
Counting and combinatorial arguments	Difficulty finding an equivalent simplifying model to which a counting argument can be applied.
Empty set and empty string	Empty, or nothing, is often confusing for students because it tends to be a special case.
Functions	Several areas can lead to misunderstandings in the area of functions: • A valid function must satisfy uniqueness and completeness properties. • Co-domain vs. range (students often confuse these two concepts) • Generality of f: A ◊ B (most students think of functions as real ◊ real) • Properties (these concepts are often confused or not understood) o One-to-one o Onto • Inverse (when can a function have an inverse? Use of an inverse.) • Difficulty understanding and applying the fundamental concept composition
Implication and conditionals	Difficulty understanding the rows in the truth table when the hypothesis/antecedent is false can lead to logical reasoning misconceptions.
Inductive reasoning	Difficulty understanding that this method involves building up cases one step at a time until a pattern emerges.
Negation	Applicable in both propositional logic and predicate logic. For example in propositional logic, the negation of a logical “or” is a logical “and,” and in predicate logic the negation of a universally quantified predicate is an existential quantified predicate.
Pigeonhole principle	Difficulty recognizing that the pigeonhole principle is an application of a function whose domain and co-domain are finite and the size of the domain is greater than the size of the co-domain.
Recursive thinking	Difficulty recognizing that the smaller case to which a problem is reduced must have the same properties (“structure”) as the original problem.
Vacuous truth	Difficulty recognizing that universally quantified predicates are true when the set over which they are defined is empty.

APPENDIX B
Development of Potential CI Items

In this Appendix, the Working Group presents four items as examples that could appear on a Discrete Math Concept Inventory (DMCI). Included with each candidate item is the following: an open-ended question that could be used during the pilot phase to investigate students' misconception, the stem for a multiple-choice item, the key option (correct answer) and possible distractors, a discussion of each option, and a brief paragraph that deconstructs the item.

Examples B.2 (on inductive reasoning) and B.4 (on probability) were somewhat contentious during discussions when the Working Group met at the *ITiCSE'2006* conference in Bologna, Italy. These two items may be poor candidates for inclusion on a concept inventory. We invite the reader to look through these examples and contact us (via one of the two Working Group leaders) with reactions or advice.

Example B.3, the sets example, is given in two versions, one "notation-heavy" and the other "notation-light".

Example B.4, the probability example, includes two different open-ended questions and two corresponding multiple-choice items.

B.1 Concept: Propositional Logic

Potential misconception: The student does not understand logical implication.

Potential open-ended question:

> A teacher said to a student, "If you receive an A on the final exam, then you will pass the course." The student did not pass the course. Give one valid conclusion and one invalid conclusion.

Candidate CI item:

> A teacher said to a student, "If you receive an A on the final exam, then you will pass the course." The student did not pass the course. Which of the following conclusions are valid?

a. The student received an A on the final exam. [*Invalid* – For this option to be valid, either the outcome or the teacher's statement would have to be invalid.]

b. The student did not receive an A on the final exam. [*Valid* – Matches the condition given in the teacher's statement.]

c. The student flunked the final exam. [*Invalid*: The teacher's statement only establishes a relationship between an A on the final exam and passing the course; no conclusions can be drawn regarding final exam grade except that it is not an A.]

d. If the student passed the course, then he or she received an A on the final exam. [*Invalid* – It is possible that a student could pass the course by earning a B or C on the final exam or even under other circumstances.]

e. None of these conclusions is valid. [*Invalid* – B is valid.]

Brief deconstruction: This item addresses fundamental misconceptions regarding logical implication in the context of the English language. Students frequently make invalid inferences.

B.2 Concept: Inductive Reasoning

Potential misconception: Students often struggle with inductive reasoning, that is, with identifying patterns in repetitive information or processes.

Potential open-ended question:

> A man and a woman are walking side-by-side and their right feet touch the ground at the same time. The woman takes three steps for each two steps of the man, and they continue walking side-by-side. Use either mathematical notation or an English statement to explain how many steps the man takes before their left feet touch the ground at the same time.

Candidate CI item (written to imply a definitive answer)

> A man and a woman are walking side-by-side and their right feet touch the ground at the same time. The woman takes three steps for each two steps of the man, and they continue walking side-by-side. Which of the following statements most accurately represents the number of steps the man takes until their left feet touch the ground at the same time?

a. 6 [*Invalid* – Plausible since 6 = 2 * 3]

b. 6 • n, for all n = 1, 2, 3, … [*Invalid* – Plausible since 6 is a factor of 6 •n]

c. Some k = 1, 2, 3, …, 10,000 [*Invalid* – Plausible if the student is guessing rather than reasoning.]

d. 12 • n, for all n = 1, 2, 3, … [*Invalid* – Similar to b, but unclear why 12 is used – maybe the student finds the sum of the steps incorrectly.]

e. None of the above [*Valid* – their left feet never touch the ground at the same time only their right feet hit the ground together. We explicitly did not include "never" among the options.]

Brief deconstruction: This question can reveal student understanding about inductive reasoning. The item includes extraneous information, which is potentially confusing. For example, students might confuse right and left steps.

B.3 Concept: Sets

Potential misconception: Student does not understand the empty set, basic set operations, and notation.

"Notation-heavy" Version

Potential open-ended question:

Let S be any well-defined set and { } be the set containing no elements. Using S and { } for each of the set operations $\subseteq$ and $\in$, write one true statement.

Candidate CI item:

Let S be any well-defined set and { } be the set containing no elements. Which of the following statements is always true?

a. { } $\in$ S [*Invalid* – Would be true if { } is a member of S, but the empty set is not automatically an element of any given set.]

b. { } $\subseteq$ S [*Valid* – The empty set is a subset of every set.]

c. { } $\subset$ S [*Invalid* – False when S contains no elements.]

d. { } = { { } } [*Invalid* – The empty set can never be equivalent to the set that has the empty set as an element.]

Brief deconstruction: If the student understands the basics of sets and set notation, this item is relatively easy to understand, although the phrase "always true" may cause confusion. Students confused by the notation and by the meaning of the empty set will likely select one of the distractors.

"Notation-light" Version

Potential open-ended question:

Let S be any well-defined set. Write one true statement about the relationship between the set Sand the empty set using each of the following set operators.

1) is a member of

2) is a subset of

3) is a proper subset of

4) is the same set as

Candidate item:

Let S be any well-defined set. Which of the following statements is always true?

a. The empty set is a member of S

b. The empty set is a subset of S

c. The empty set is a proper subset of S

d. The empty set is the same as a set containing the empty set

Brief deconstruction: This item is conceptually identical to the "notation-heavy" one given above. This "notation-light" version removes the additional hurdle of notation, perhaps making it easier for some students to answer correctly.

B.4 Concept: Probability

Potential misconception: The student is confused about the probability of independent events.

Two potential open-ended questions:

1. Consider a sequence of independent flips of a fair coin. Suppose the first four flips are all heads. What is the probability that the fifth flip will be tails?
2. A coin is biased so that the chance of getting heads is 75% and of getting tails is 25%. Explain what combinations of flips of this coin could be used to yield a probability of 1/2.

Two candidate CI items:

1. Consider a sequence of independent flips of a fair coin. Suppose the first four flips are all heads. The probability that the fifth flip will be tails is

   a. 0 [*Invalid* – Not a fair coin, similar to (e)]

   b. 1/5 [*Invalid* – Not an independent event]

   c. 1/2 [*Valid* – events are independent]

   d. 4/5 [*Invalid* – Not an independent event]

   e. 1 [*Invalid* – Not a fair coin, similar to (a)]

2. A coin is biased so that the chance of getting heads is 75% and of getting tails is 25%. Which of the following combinations of flips of this coin makes a 50/50 bet?

   a. Flip it once [*Invalid* – Clearly fails, since this results directly in a 25/75 chance.]

   b. Flip it twice [*Invalid* – Of the 4 possible outcomes, only T-H and H-T yield the required 50/50 chance.]

   c. Flip it three times [*Invalid* – Of the eight possible outcomes, none yield the required 50/50 chance.]

   d. Flip it an even number of times [*Valid* – By flipping twice and continuing until getting either T-H or H-T, this results in the required 50/50 chance]

   e. Flip it an odd number of times [*Invalid* – While plausible, this does not guarantee 50/50]

Brief deconstruction: The stems of both items are relatively clear, although the first has less information than the second and may therefore be easier for a student to understand. While the five responses for the first item are very easy to understand, several of the responses reveal similar, but not identical, misconceptions (e.g., a & e, and b & d). For the second item, a student might require a long time to analyze each option to determine the correct response. This could be construed as a tricky item because determining the correct answer requires fairly deep, indirect reasoning.