Developing an Instrument to Measure Engineering Student Misconceptions in Thermal and Transport Science

Ruth A. Streveler
School of Engineering Education
Purdue University

Ronald L. Miller
Department of Chemical Engineering
Colorado School of Mines

Mary A. Nelson
Department of Applied Mathematics
University of Colorado

Monica R. Geist
Department of Applied Statistics and Research Methods
University of Northern Colorado

Barbara M. Olds
Trefny Institute for Educational Innovation
Colorado School of Mines

ABSTRACT

This paper describes research designed to answer two questions: “What important concepts in thermal and transport science are difficult for engineering students to learn?” and “Can an instrument be developed to identify engineering student misconceptions of these difficult and important concepts?” The first research question was investigated by conducting a Delphi study with experienced engineering faculty while the second question was investigated by developing the Thermal and Transport Concept Inventory (TTCI). This instrument measures the conceptual understanding of key ideas in thermodynamics, fluid mechanics, and heat transfer by undergraduate engineering students. Items for the instrument were developed using student responses to open-ended questions followed by beta testing of multiple-choice items; student language was used to create plausible but wrong answers for each item. Beta test data from six engineering schools were used to calculate traditional measures of reliability. Future directions for development of the TTCI are discussed.

Keywords: concept inventory, Delphi study, difficult concepts, misconception

I. INTRODUCTION

Most tools and methods to assess engineering student learning focus on either procedural knowledge (e.g. solving specified classes of problems, designing a process or artifact, using appropriate engineering tools, oral and written communication) or development of affective and behavioral characteristics (e.g. teamwork, life-long learning, professional and ethical responsibility). Beginning in the 1970's, education researchers and educators began to identify conceptual shortcomings in students and the propensity for students to carry with them strongly-held misconceptions describing how the world around them worked (Duit 2007). One of the first systematic methods for assessing student's conceptual understanding was reported for undergraduate physics education by David Hestenes and his colleagues (Hestenes, Wells, and Swackhamer 1992). The instrument they developed, known as the Force Concept Inventory (FCI), consists of` 29 multiple choice items, each designed to probe students' understanding of Newtonian force concepts.

The breakthrough impact of the instrument occurred when Hake published FCI results for ~6000 undergraduate students which clearly showed the positive effect of active-learning and inquiry-based pedagogical techniques on understanding of the force concept as measured by FCI scores (Hake 1998). Later, Mazur at Harvard used the FCI with his students and found that, much to his surprise, student gains were no better than results reported in Hake's study (Mazur 1992). Along with other innovators, Mazur began the revolution in physics education in which a renewed focus on conceptual understanding replaced some of the emphasis on routine problem-solving.

As the positive effect of the FCI on physics education has become more widely known, concept inventories (CI's) have emerged in many science and engineering fields. New CI's are now available or under development in thermodynamics (Midkiff, Litzinger, and Evans 2001), fluid mechanics (Martin, Mitchell, and Newell 2003), heat transfer (Jacobi et al. 2003), materials engineering (Krause et al. 2003), strength of materials (Richardson et al. 2003), statics (Steif and Dantzler 2005), electromagnetic waves (Rhoads and Roedel 1999), signals and systems (Wage et al. 2005), electric circuits (Evans et al. 2003), and statistics (Stone et al. 2003) among other fields. These assessment tools have been subjected to varying degrees of validity, reliability, and bias testing.

The purpose of this paper is to report on a project that was inspired by the work of Hestenes, Hake and Mazur. Two research questions were investigated:

- What important concepts in thermal and transport science are difficult for engineering students to learn?
- Can an instrument be developed to identify engineering student misconceptions of these difficult and important concepts?

We will discuss the methodology used to answer the first research question and then describe the genesis of the Thermal and Transport Concept Inventory used to answer the second research question. To help understand the development process for the TTCI, we describe the development process of one example concept item from initial formulation through open-ended student interviews and written surveys to field testing of the draft multiple-choice item to the current version of the item. We will then present validity and reliability data for version 2.5 of the TTCI and discuss future work intended to continue improving the instrument.

II. WHAT IMPORTANT CONCEPTS IN THERMAL AND TRANSPORT SCIENCE ARE DIFFICULT FOR ENGINEERING STUDENTS TO LEARN?

The first task of the research was to develop a list of concepts in thermal and transport science that are important but difficult for undergraduate engineering students. A Delphi study was chosen as the method to develop this list. The Delphi method was deemed appropriate for this purpose because it is a structured process for collecting and distilling knowledge from a group of experts by means of a series of questionnaires interspersed with controlled opinion feedback (Adler and Ziglio 1996). Selection of appropriate participants is crucial in the Delphi methodology and therefore well-respected engineering faculty experts and prominent textbook authors were invited to participate. The participants were then asked to identify important concepts in thermal and transport science disciplines that are consistently difficult for students to understand and for which the students possess significant and robust misconceptions. The distinguishing features of the Delphi technique are its use of experts and its methodology. Delphi proponents recognize human judgment as a legitimate and useful input in generating predictions and therefore believe that the use of experts, carefully selected, can lead to reliable and valid results. In addition, the Delphi technique attempts to overcome weaknesses implicit in other methods such as relying on a single expert, a group average, or a round table discussion. Using a single expert puts too much weight on one person's opinion; the group average method fails because as Clayton notes, "the individuals consulted have neither the opportunity to provide their most thoughtful input nor the benefit of hearing other responses that might encourage a refinement of the contributions" (Clayton 1997); and the round-table approach is unreliable because some members of the group may unduly influence the decision. The Delphi method addresses the latter concern by soliciting input anonymously so that influences such as the professional reputation of a respondent or the forcefulness of a respondent's personality are neutralized. Thus all participants have equal stature in the process and their comments influence the other participants only through the logic of their argument, not their name recognition.

A. Our Application of the Delphi Process

According to Linstone and Turoff, "Usually Delphi, whether it be conventional or real-time, undergoes four distinct phases. The first phase is characterized by exploration of the subject under discussion, wherein each individual contributes additional information he [sic] feels is pertinent to the issue. The second phase involves the process of reaching an understanding of how the group views the issue (i.e., where the members agree or disagree and what they mean by relative terms such as importance, desirability, or feasibility). If there is significant disagreement, then that disagreement is explored in the third phase to bring out the underlying reasons for the differences and possibly to evaluate them. The last phase, a final evaluation, occurs when all previously gathered information has been initially analyzed and the evaluations have been fed back for consideration." (Linstone and Turoff 1996)

Panel Selection. Since the Delphi method relies on expert opinion, it was important to select the right experts. In some cases, Delphi participants are selected through a "nomination" process in which recognized experts are solicited but also asked to provide the names of other experts (Fish and Busby 1996). Selection criteria should be clearly articulated, e.g. number of years of experience, number of publications or other expert qualifications. For our panel, we started with a geographically distributed list of people with extensive expertise in the appropriate fields and considerable undergraduate teaching experience. As anyone who has taught engineering students will attest, classroom teaching experience and student interactions provide a rich source of anecdotal data about lack of student understanding and the presence of misconceptions. We asked these experts to help us identify others, including textbook authors in the relevant fields. Once we had identified approximately 35 experts, we sent them an email explaining the Delphi process and our project along with an invitation to join the group. Thirty-one experts agreed to participate, a number corresponding well with Clayton's rule-of-thumb that 15-30 people are an adequate panel size (Clayton 1997). The group included tenured and tenure-track engineering professors from research universities and undergraduate institutions. Five of the participants have authored well-known texts in thermodynamics, fluid mechanics, heat transfer, or thermal science. Although we purposely included new faculty members in the study, the average number of years taught was 23. We guaranteed confidentiality for all participants during the process, another important element of the Delphi procedure.

Generative round. We wanted our expert panel to generate a beginning list of difficult concepts in thermal and transport science rather than generate our own. Therefore, we included a pre-Delphi generative round for our project in which we asked panelists to describe concepts they found difficult for their students. Table 1 summarizes example comments received for the domain of heat transfer. Of the 31 experts who agreed to participate in the study, 23 provided approximately 60 concept ideas, which were coded and organized into a list of 28 concepts; this list included all concepts that had been submitted by at least two experts. These 28 concepts formed the basis for our subsequent rounds.

Round 1. We developed a questionnaire using the 28 concepts identified in the generative round and asked the experts to rate each concept based on two factors: 1) the proportion of his/her students that *understand* the concept and 2) how *important* it is for students to understand the concept. We used a scale of 0-10 for each question (0 = no one understands this concept to 10 = everyone understands this concept, and 0 = it is not at all important to understand this concept to 10 = it is extremely important to understand this concept). Thirty members of our expert panel ranked these 28 concepts. In all rounds, participants were told that "you will not have to rate any concept for which you don't feel you have sufficient expertise or classroom experience."

Table 1. Examples of Generative Round Comments by Delphi Participants Related to Heat Transfer Misconceptions

Confusion between temperature and heat transfer. Many students believe that if the path to a state involves a heat transfer input the temperature of the system will increase – even if the heat transfer is coupled with work leaving the system.
Students do not appear to have precise understanding of heat in the sense that it is used in thermodynamics.
Students have difficulty identifying heat and work interactions between the system and the surroundings.
[Students] have trouble with work and energy relationships.
There is always student confusion about heat vs. (internal energy).
There is always student confusion about internal energy and enthalpy.
Misconception: Temperature is a measure of energy. Example: Students often believe that if you add energy, heat for example, to any system, the temperature must go up. A corollary to this is that students often believe that if the temperature goes up the energy (internal energy or enthalpy) must have increased. A good example of a system that is very confusing is an evaporative cooling process in psychrometrics where the enthalpy of the moist air stays constant but the temperature decreases.
Heat, like energy, is a familiar term but its common use differs from thermodynamic definition.
Heat as transferred energy. No matter how often you make the point, some [students] insist on talking about the heat content of a system.
Confusion about the difference between heat and temperature. How can a process occur where heat is added but the temperature drops?

Round 2. For the second round of the Delphi study, we presented the panel of experts with the same 28 concepts and asked them to rank the concepts using the same scales as in round 1. In round 2 we also provided them with the median response to each question from round 1 and the interquartile range of responses (containing the middle 50 percent of responses). In the second round, if participants rated a concept outside of the interquartile established in round 1, they were asked to provide an explanation for their rating. In this way, the median ranking of each concept approached a stable value and the

interquartile range decreased in size, representing the consensus opinion of the participant group. Twenty-eight members of our expert panel ranked the 28 concepts in round 2.

Round 3. In the third round, we again asked the experts to rank all 28 concepts. We provided the median rating from round 2 and the anonymous comments that fellow panelists made justifying ratings outside of the specified range. Based on this final iteration, we identified 12 of the least understood but most important concepts in thermal and transport sciences which formed the content domain for developing TTCI test items.

Delphi Results. Based on responses from 30 participants for round 1, 28 participants in round 2 and 26 participants in round 3, we computed the ranking statistics for each of the 28 concepts in the study. These results are summarized in Table 2. The non-parametric median and interquartile range were used (rather than mean and standard deviation) because an ordinal scale was used to rate the concepts. These results indicate that the rankings for most concepts stabilized by round 2 (the median for 19 of the 28 concepts changed by a value of 0.5 or less) as suggested by other Delphi studies reported in the literature (Linstone and Turoff 1996).

The goal of this part of the study was to identify concepts which had high importance (those that were given a high ranking in the "importance" scale) and were also conceptually difficult (those that were given a low ranking on the "conceptual understanding" scale. A total of 12 concepts were identified as meeting the criteria of high importance but low conceptual understanding. These items included key topics in thermal science and transport disciplines such as the 2nd law of thermodynamics including reversible vs. irreversible processes, conservation of fluid momentum, viscous momentum transfer, several energy-related topics (heat, temperature, enthalpy, internal energy), and steady-state vs. equilibrium processes. Two of the concepts (differential vs. integral analysis and system vs. control volume analysis) were deemed mathematical rather than physical concepts and were not included in the TTCI instrument. At the request of several Delphi participants, we also included the ideal gas law and conservation of mass concepts in the TTCI since both are fundamental concepts in fluid mechanics and thermodynamics. Finally, we temporarily set aside the thermal radiation concept which will eventually be included in the instrument. This decision was made so that we could focus development of early versions of the instrument on what were deemed more fundamental heat transfer topics.

To independently validate the results of the Delphi process, we consulted the available misconception literature, especially a comprehensive bibliography of ~7700 studies reported by Duit (2007). The bibliography contains ~500 references to work on heat transfer misconceptions and slightly fewer numbers for thermodynamics and fluid mechanics.

Table 2. Results of Thermal and Transport Concepts Delphi Study[1]

Concept	"Understanding" Data Median (interquartile range)			"Importance" Data Median (interquartile range)		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
1.Adiabatic vs. Isothermal Processes	7.5 (6-8)	8 (6-8)	8(6.75-8.25)	9 (8-10)	9 (9-10)	9 (9-10)
2.Bernoulli Equation	7 (4-8)	6 (5-7)	6 (5-7)	9 (7-10)	9 (8-9)	9 (8-9)
3. Compressible vs. Incompressible Flow	5 (3-7)	6 (4-6.5)	6 (5-7)	7.5 (6-8)	7 (7-8)	7.5 (7-8)
4. Conservation of Linear Momentum	5 (3-6)	5 (4-6)	5.5 (5-6)	9 (8-10)	9 (8-10)	9(8-9.25)
5. Differential vs. Integral Analysis	4.5 (3-6)	4 (3-5.25)	4 (4-5)	7 (6-9)	8 (6-8)	8 (7-9)
6. Dimensional Analysis	6 (4-7)	5.5 (4.25-7)	6 (5-6.25)	7 (5-7)	6 (5-8)	7 (5-8)
7. Entropy & 2nd Law of Thermodynamics	4 (2-6)	4 (3-5)	5 (3-5.25)	8 (7-9)	9 (8-9)	9 (8-10)
8. Extensive and Intensive Properties	8 (6-9)	8 (7-8)	8 (7-9)	7 (6-9)	8 (7-9)	8 (7-9)
9. First Law of Thermodynamics	8 (7-9)	8 (7-9)	8 (8-9)	10 (10-10)	10 (10-10)	10 (10-10)
10. Fluid vs. Flow Properties	7 (5-8)	6 (5-7)	6 (5-6)	7 (5-9)	7 (5-8)	7 (5-8)
11. Heat Transfer Modes	8 (6-9)	8 (6.25-8)	8 (7-9)	9 (8-10)	9 (9-10)	9 (9-10)
12. Heat vs. Energy	6 (5-8)	6 (5-7)	6.5 (5-7)	9 (8-10)	9 (8-10)	9 (8-10)
13. Heat vs. Temperature	6 (4-8)	6.5 (5-8)	7 (6-8)	9 (8-10)	10 (9-10)	10 (9-10)
14. Ideal Gas Law	8 (7-9)	8 (8-9)	8 (8-9)	9 (8-10)	9 (9-10)	9 (9-10)
15. Internal Energy vs. Enthalpy	6 (3-7)	5 (4-6)	6 (5-6.25)	8 (7-9)	9 (8-9)	9 (8-9)
16. No-slip Boundary Conditions	8 (6-9)	8 (7-9)	8 (8-9)	8 (7-9)	9 (8-9)	9 (8-9)
17. Nozzles and Diffusers	6 (5-8)	6 (6-7.5)	7 (6-7)	7 (5-9)	7 (6-8)	7 (6-8)
18. Pressure	8 (6-9)	8 (7-8)	8 (7.75-9)	9 (8-10)	10 (9-10)	10(9.75-10)
19. Reversible vs. Irreversible Processes	5 (4-7)	5 (4-6)	5 (5-6)	8 (8-9)	9 (8-9)	9 (8-9)
20. Spatial Gradient of a Function	4 (3-7)	5 (4-6)	5 (4-5)	7 (3-9)	7 (6-8)	7 (6-8)
21. Specific Heat Capacity	7 (6-8)	7 (6-7)	7 (6-8)	8 (7-10)	9 (8-9)	9 (8-9)
22. Steady-state vs. Equilibrium Process	5 (3-8)	5 (3-6)	5 (4-5.25)	8 (5-10)	9 (7-9)	9 (8-9)
23. Steady-state vs. Unsteady-state Process	8 (7-8)	8 (7-8)	8 (7-8)	9 (8-10)	9.5 (9-10)	9.5 (9-10)
24. System vs. Control Volume	7 (4-8)	6 (5-7)	6 (6-7)	8 (6-10)	9 (8-10)	9 (8.5-10)
25. Temperature Scales	7 (5-9)	8 (8-9)	9 (8-9)	8 (6-10)	9 (8-10)	9 (9-10)
26. Thermal Radiation	6 (4-8)	5 (5-6)	5 (5-6)	7 (5-9)	8 (6.75-8)	8 (7-8.25)
27. Thermodynamic Cycles	7 (5-8)	7 (6-7)	7 (7-8)	8 (8-10)	9 (8-10)	9 (8-9.25)
28. Viscous Momentum Flux	5 (3-7)	4 (3.75-5)	4 (3-4)	7.5 (6-9)	8 (7-8)	7 (7-8)

"Understanding" Scale	**"Importance" Scale**
0 = no one understands the concept	0 = no at all important to understand the concept
10 = everyone understands the concept	10 = extremely important to understand the concept

[1]*Italicized* concepts are those the Delphi study identified as poorly understood but highly important (i.e. low scores on the "understanding" scale but high scores on the "importance" scale)

Confusion about thermal processes and heat transfer have been identified in students of all ages and focus on the following five conceptual themes (Carlton 2000; Thomas and Schwenz 1998; Thomaz et al. 1995):

- heat and temperature are equivalent (related to concept 13 from Table 2)
- temperature determines how "cool" or "warm" a body feels (related to concept 12 from Table 2)
- heat is a substance transferred between bodies (related to concept 12 from Table 2)
- addition of energy as heat always increases the temperature in a body (related to concept 15 from Table 2)
- temperature should change in a phase transition (e.g. boiling) since energy is being added or removed (related to concept 15 from Table 2)
- thermal equilibrium is equivalent to the steady flow of heat (related to concept 27 from Table 2)

As indicated in the list, we found that each of the "important but poorly understood" heat transfer misconceptions identified in the Delphi study is prominently mentioned in the misconception literature. Each of these themes has been discussed in more detail elsewhere (Streveler et al. in press).

The misconception literature for thermodynamics is less exhaustive but does indicate that the following misconceptions persist in significant numbers of students (Arnold and

Millar 1996; Clark and Jorde 2004; Johstone and McDonald 1977; Kesidou and Duit 1993):

- energy quality does not degrade and all processes are reversible (related to concepts 7 and 19 in Table 2)
- heat is a property of a system (related to concepts 12 and 15 in Table 2)
- objects at the same temperature are not necessarily in thermal equilibrium (related to concept 22 in Table 2)

Again, all the misconceptions identified in the Delphi study are prominently mentioned in the thermodynamics misconception literature. In addition, the literature describes a multitude of misconceptions associated with energy conservation and conversion, a concept which was rated extremely important by Delphi experts but understood well enough not to be included in the TTCI instrument.

Finally, the misconception literature for fluids and fluid flow mentions the following common misconceptions (Besson 2004; Kariotogloy, Koumaras, and Psillos 1993; Psillos and Kariotogloy 1999):

- pressure (rather than pressure gradient) causes flow (related to concept 2 in Table 2)

- pressure is a function of fluid volume rather than depth (i.e. pressure is related to horizontal as well as vertical dimensions) (related to concept 2 in Table 2)
- buoyancy is not related to fluid depth or density (related to concept 2 in Table 2)
- fluid pressure increases if flow cross-sectional area is reduced since flow is "pushed" through the smaller area (related to concept 2 in Table 2)
- viscosity is not a fluid property or viscosity does not affect behavior of fluid motion (related to concept 28 in Table 2)
- linear momentum is not conserved or is not related to forces acting on a fluid (related to concept 4 in Table 2)

We conclude that all of the fluid mechanics misconceptions identified in the Delphi study have been a focus of previous misconception work reported in the literature. Overall, results from this analysis lend an additional degree of content validity to results from the Delphi study and formed the basis for developing relevant thermal and transport questions for our concept inventory.

III. ANSWERING RESEARCH QUESTION 2: CAN AN INSTRUMENT BE DEVELOPED TO MEASURE ENGINEERING STUDENT MISCONCEPTIONS OF DIFFICULT BUT IMPORTANT CONCEPTS IDENTIFIED IN THE DELPHI STUDY?

Once we had identified concepts from the Delphi study that would be included in a multiple-choice misconception instrument patterned after the Force Concept Inventory, we began developing candidate items for each concept. Each item was developed using a seven-step process which included: 1) drafting open-ended questions about the concept, 2) collecting student response data orally (think-aloud problem solving sessions) and in written form, 3) using the responses to convert the open-ended questions to multiple choice items with distractors describing plausible but incorrect answers, 4) collecting expert reviews on each item, 5) beta testing the drafted items on groups of engineering students, 6) revising the items based on statistical performance and expert feedback, and 7) collecting additional beta test data. To illustrate this process, we will describe the genesis and development of one TTCI item known as ***Hotplate***.

A. Phase One: From Concept to Open-Ended Item

Hotplate was designed to assess students' conceptual understanding of the relationship among energy (specifically internal energy), temperature and heat. The genesis for the item was a similar question included in the Chemistry Concepts Inventory authored by Doug Melford as a M.S. student at Purdue in 1996 (Mulford 1996). The original open-ended version of ***Hotplate*** is shown in Figure 1.

Two identical beakers contain equal masses of liquid at a temperature of 20 ^{O}C. One beaker is filled with water and the other beaker is filled with ethanol (ethyl alcohol). The temperature of each liquid is increased from 20 ^{O}C to 40 ^{O}C using identical hot plates.

It takes 2 minutes for the ethanol temperature to reach 40 ^{O}C and 3 minutes for the water to reach 40 ^{O}C. Once a liquid has reached 40 ^{O}C, its hot plate is turned off.

To which liquid was more energy transferred during the heating process?

Why did you answer the way you did (i.e. explain your reasoning)?

Figure 1. Open-ended Version of ***Hotplate*** Item

In this item, each fluid is heated in an identical beaker between the same starting and ending temperature using the same rate of heating with identical hot plates. Based on student think-aloud data, we expected that students who were confused about the relationship between energy, temperature, and heat capacity would not generally be able to answer this item correctly.

Six students (all juniors or seniors majoring in chemical or mechanical engineering) individually participated in think aloud sessions to consider the hotplate item. The role of the interviewer in this session was to elicit more detailed student answers to ***Hotplate*** and to elicit explanations about why students answered as they did. Student responses were audiotaped, transcribed, and analyzed for evidence of conceptual understanding and prevalent misconceptions about energy and temperature. As predicted by the Delphi results, the majority of students participating in think-alouds demonstrated limited understanding of the concepts being addressed, though some provided reasonably correct answers. Examples included:

<u>Incorrect answers based on misconceptions</u>:

"They both received the same energy because the temperature change was the same."

"We can't tell because we don't know the [fluid] heat capacities."

"It has nothing to do with heat transfer, only temperature."

"The amount of time it takes to heat is based only on the heat transfer coefficient in the beakers."

<u>Correct answers</u>:

"Just because ethanol gets hotter faster does not mean it gains more heat. Just that the ethanol has a lower heat capacity."

"Water because it was heated longer at the same rate of heating."

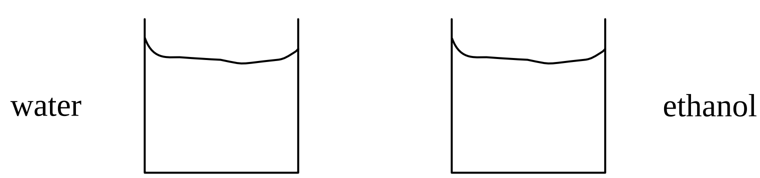
B. Phase Two: From Open-Ended Question to Multiple Choice Test Item

During coding of the ***Hotplate*** think-aloud data, we identified several misconception patterns in student responses including confusion about temperature vs. energy or heat, what heat capacity of a substance means, and equating the rate of heat transfer with the amount of energy transferred. These results confirmed the predictions of the Delphi experts (Table 1). As mentioned earlier, these misconceptions have been prominently reported in the thermal science misconception literature and gave us confidence that our think-aloud strategy was eliciting significant student thinking which was worthy of inclusion as plausible distractors for the multiple-choice version of ***Hotplate***. Using student comments like those listed above combined with input from the literature and Delphi participants, we drafted four ***Hotplate*** distractors (answers b-e) along with the correct answer (answer a). The original multiple choice version of ***Hotplate*** is shown in Figure 2.

Hotplate TTCI Test Item

Two identical beakers contain equal masses of liquid at a temperature of 20 ^{O}C as shown below. One beaker is filled with water and the other beaker is filled with ethanol (ethyl alcohol). The temperature of each liquid is increased from 20 ^{O}C to 40 ^{O}C using identical hot plates.

It takes 2 minutes for the ethanol temperature to reach 40 ^{O}C and 3 minutes for the water temperature to reach 40 ^{O}C. Once a liquid has reached 40 ^{O}C, its hot plate is turned off.



<u>To which liquid was more energy transferred during the heating process?</u>

a. Water because more energy is transferred to the liquid that is heated longer.

b. Alcohol because more energy is transferred to the liquid that heats up faster (temperature rises faster).

c. Both liquids received the same amount of energy because they started at the same initial temperature and ended at the same final temperature.

d. Can't determine from the information given because heat transfer coefficients for water and ethanol are needed.

e. Can't determine from the information given because heat capacities of water and ethanol are needed.

Figure 2. Version 1 of ***Hotplate*** Multiple Choice Test Item

C. Phase Three: Initial Beta Testing

The version of ***Hotplate*** shown in Figure 2 was beta tested with 87 students at four engineering institutions participating in the TTCI beta study. A summary of students' responses is shown in Figure 3 and indicates that about 50 percent of the students answered the item correctly. Distractors c (~15 percent), d (~13 percent), and e (~19 percent) were all selected by a significant number of students suggesting that the item may have been eliciting the types of students' misconceptions and incorrect thinking that we expected to see. Distractor "b" was only selected by two students which was a bit surprising given interview data which indicated student confusion between rate of heating and amount of energy transferred. Good test construction practice recommends removing or replacing distractor "b" and that may happen in later versions of the instrument, but for now it remains part of the item until additional test data are collected to make a more reasoned decision.

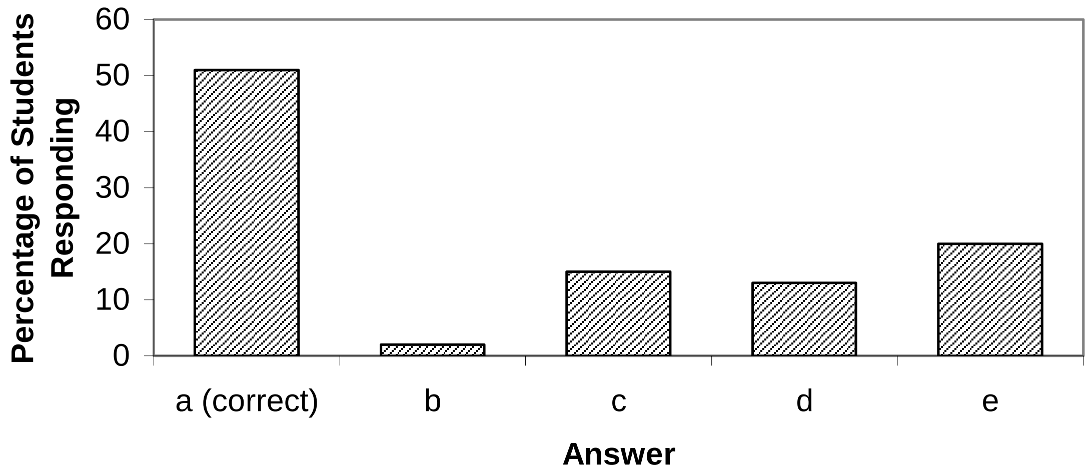


Figure 3. Summary of Student Responses to Hotplate Item

Other statistical results for ***Hotplate*** were also computed including the item difficulty index (=0.51) and item discrimination index (=0.58). These statistical markers will be discussed in the next section as part of the TTCI psychometric analysis discussion, but both values were judged to be well within acceptable range for items in the TTCI. See section V.A for more discussion of difficulty and discrimination indicators.

D. Phase Four: Expert Review

After we completed initial beta testing, each item in the TTCI was reviewed by two technical experts in the disciplines of fluid mechanics, heat transfer or thermodynamics. Expert feedback and comments about ***Hotplate*** focused on three issues:

- Effect of evaporation losses in the open beakers
- Effect of using hotplates where the effect of fluid properties might affect actual heat transfer rates for water and alcohol

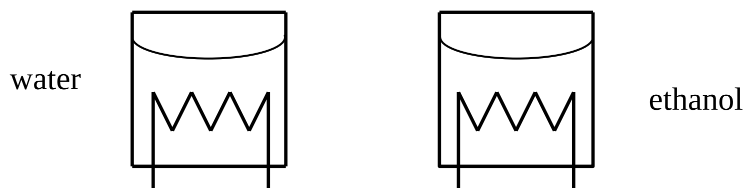
- Wordiness in the item set-up and some of the distractors

As a result of this valuable feedback, we replaced the use of hotplates and open beakers with immersion heaters in closed beakers to minimize the effects of evaporation and varying fluid properties on the heating rates. To improve clarity, we also rewrote the item and shortened distractors a-d. Version 2 of ***Hotplate*** is shown in Figure 4.

Hotplate TTCI Test Item

Two identical closed beakers contain equal masses of liquid at a temperature of 20 ^{O}C as shown below. One beaker is filled with water and the other beaker is filled with ethanol (ethyl alcohol). The temperature of each liquid is increased from 20 ^{O}C to 40 ^{O}C using identical heaters immersed in the liquids. Each heater is set to the same power setting.

It takes 2 minutes for the ethanol temperature to reach 40 ^{O}C and 3 minutes for the water temperature to reach 40 ^{O}C.



Ignoring evaporation losses, to which liquid was more energy transferred during the heating process?

a. Water because it is heated longer

b. Alcohol because it heats up faster (temperature rises faster)

c. Both liquids received the same amount of energy because they started at the same initial temperature and ended at the same final temperature

d. Can't determine from the information given because heat transfer coefficients from the water and alcohol beaker surfaces are needed

e. Can't determine from the information given because heat capacities of water and ethanol are needed

Figure 4. Version 2 of ***Hotplate*** Multiple Choice Test Item

E. Phase Five: Additional beta testing

To begin assessing the performance of the revised version of ***Hotplate***, we administered the item to 29 seniors in chemical engineering. Table 4 compares chemical engineering student responses for the new version of ***Hotplate*** (Figure 4) with results collected from mechanical and chemical engineering students at four beta test sites using the old version (Figure 2). Although based on limited data and with different demographics for each beta test, the results show that significant misconceptions persist with both groups of students and with each version of the ***Hotplate*** test item. The distribution of distractor selection varied, particularly involving distractors “c” and “d.” Comparing results from the two data sets showed that far fewer chemical engineering seniors believed that heat transfer coefficients were required to answer the question (a good result), but many more believed that both liquids received the same amount of energy for a given temperature change or that heat capacity data was required to answer the item (a troubling and incorrect result). Thus, we concluded that the item still identified key misconceptions held by engineering students, even those who have completed numerous courses in thermal and transport science topics.

Table 4. Comparison of Student Response Distribution for Versions 1 and 2 of the Hotplate Item

Response	Version 1 of ***Hotplate*** (n=87)	Version 2 of ***Hotplate*** (n=29)
a (correct)	51 percent	41 percent
b	2 percent	2 percent
c	15 percent	31 percent
d	13 percent	2 percent
e	19 percent	24 percent

V. Psychometric Analysis of the TTCI Instrument

Each question in the TTCI was developed using the process described in section IV in which open-ended draft questions were prepared and student responses collected in oral (structured interviews) and written (open-ended responses collected in engineering classes) forms. These initial responses were used to: 1) refine each question, 2) draft language for the correct answer and distractors and 3) identify problems in terminology, labeling or graphics. Draft multiple-choice questions were then used with small groups of students (~10) to collect alpha test data which was used to further refine question language.

After alpha testing, each revised question was included as part of the TTCI instrument and was used in beta testing at six engineering institutions in the United States (a total of 102 students answered the fluids items, 68 the heat transfer items, and 178 the thermodynamics items). Version 2.21 of the instrument was used for beta testing and consisted of 12 items in fluid mechanics, 7 in heat transfer, and 13 in thermodynamics. Some items posed one question to the students while others posed two so that a total of 16 questions were asked in fluid mechanics, 9 in heat transfer, and 20 in thermodynamics. Every item was written to elicit information about both what the students thought the right answer was and <u>why</u> they answered as they did. Each beta test institution chose which portions of the instrument to use with their students so the number of student responses varied from item to item.

A. Initial screening of TTCI items

Two statistics based on classical test theory (Lord and Novick 1968; Pedhazur and Schmelkin 1991) were used to screen the beta test data and determine which items were performing satisfactorily. The item difficulty index is defined as the proportion of examinees who answered each posed question correctly. Since the TTCI focuses on <u>difficult</u> concepts, lower difficulty index values (i.e. fewer students answering correctly) are acceptable whereas with traditional achievement tests a range of low to high values would be desirable. (However, more difficult questions often result in instruments with lower reliability. This issue will be addressed later in this section.) To screen TTCI items, we retained items which fell in the difficulty index range of approximately 0.25-0.75 (between 25 percent and 75 percent of the students answered the question correctly) based on the premise that the TTCI should focus primarily on the middle two quartiles of student performance and should exclude very easy or very difficult test items (Kline 2005).

The item discrimination index was used as the second measure of item performance. This index is a measure of how well a question discriminates or differentiates between higher performing and lower performing examinees and is computed for each TTCI question as follows: 1) student scores in the beta dataset were rank ordered by the total number of questions answered correctly, 2) student rankings were divided into top one-third, middle-third, and bottom-third groups, and 3) item discrimination index was computed by finding the difference between the proportion of students in the top third and the proportion in the bottom third who answered a question correctly. Thus, the closer an item discrimination index is to 1, the more discriminating the question. For purposes of evaluating TTCI items, we followed the advice of Thorndike [23] and only accepted questions with item discrimination indices greater than 0.20. Low or negatively discriminating items are not desirable because they are likely measuring something other than what the rest of the instrument is measuring. Negatively discriminating items may be measuring nothing, they may be the result of poorly worded or confusing questions, they may be so easy that nearly all students got them correct, or they may be so difficult that most students guessed and thus little difference was detected between the high and low scoring students.

Table 5 summarizes the performance of questions in each subtest using the item difficulty and item discrimination criteria. In all cases, a majority of the items performed well and were retained.

Table 5. Summary of Item Difficulty and Item Discrimination Results from TTCI Beta Test Data

Question category	# of questions beta-tested	# of questions meeting item difficulty criterion [1]	# of questions meeting item discrimination criterion [2]	# of questions retained in TTCI
heat transfer	9	7	7	7
thermodynamics	20	17	14	13
fluid mechanics	16	15	12	11

[1] item difficulty range of 0.25-0.75
[2] item discrimination > 0.20

B. Reliability and Validity of the Instrument

After eliminating poorly performing items, we used the remaining 21 items (encompassing 31 questions since some items include 2 questions) which met our established difficulty and discrimination criteria to create version 2.5 of the TTCI. Re-analysis of the beta test data for these items was used to study instrument reliability and validity. Tables 6-8 provide a brief description of each item in version 2.5 and the percentage of students in the beta sample who answered each question correctly.

Table 6. Description of Fluid Mechanics Items in the TTCI version 2.5

Item	Description	1 or 2 question item?	Concept	percent correct responses to each question
1.A	pressure changes during steady flow through a venturi meter	2	Bernoulli principle	46, 65

2.A	direction of velocity and pressure force vectors in steady pipe flow	2	linear momentum conservation	52, 45
2.B	forces acting on a vertical jet of water impinging on a wall	2	linear momentum conservation	25, 36
8.A	direction of momentum transfer for laminar shear flow between two parallel plates	1	viscous momentum flux	45
8.B	comparison of laminar shear flow between parallel plates for fluids of different viscosity	1	viscous momentum flux	47
10.A	steady gas flow in a pipe with varying density	1	conservation of mass in flowing process	65
10.B	steady gas flow in a pipe with varying temperature in the flow direction	1	conservation of mass in flow process	29
10.C	steady flow of water in conduit of varying cross-sectional area	1	conservation of mass in flow process	64

Table 7. Description of Heat Transfer Items in the TTCI version 2.5

Item	Description	1 or 2 question item?	Concept	percent correct responses to each question
4.A	apparent temperature of tile and carpet floors when walking with bare feet	1	distinguish between temperatureand rate of energy transfer	64
5.A	heating of closed containers of water and ethanol with immersion heaters ("***Hotplate***")	1	distinguish between temperature and rate of energy transfer	51
5.B	apparent temperature difference when submerged in room temperature water or air	1	distinguish between temperature and rate of energy transfer	69
5.G	ice melting using metal blocks of different temperatures	2	distinguish between temperature and rate of energy transfer	32, 26
7.F	heating of water flowing through a pipe	2	distinguish between steady-state and equilibrium processes	44, 37

Table 8. Description of Thermodynamics Items in the TTCI version 2.5

Item	Description	1 or 2 question item?	Concept	percent correct responses to each question
3.B	ability of steam at different temperatures to do mechanical work	2	distinguish between energy quality and quantity (2^{nd} law)	65, 65
3.D	analyze proposed cycle for converting thermal energy to mechanical work	1	conversion of heat to work (2^{nd} law)	45
6.B	comparison of adiabiatic, frictionless turbine and adiabatic, frictionless piston	1	distinguish between internal energy and enthalpy	25
6.C	comparison of water heated in a pipe with water flowing through a pump	2	distinguish between internal energy and enthalpy	43, 35
7.B	dissolution of dye pellets in a beaker of water	2	distinguish between steady-state and equilibrium processes	75, 80
7.D	dissolution of table salt in a beaker of water	2	distinguish between steady-state and equilibrium processes	53, 78
9.A	compare density of gases at same temperature and pressure but different molecular weights	1	ideal gas law	33
9.B	change in gas density when number of moles change during a chemical reaction	2	ideal gas law	50, 60

A total of 68 heat transfer student responses, 38 thermodynamics student responses, and 102 fluid mechanics student responses were used in this analysis. An additional 140 student responses to the thermodynamics questions were set aside when we discovered that none of these students had taken a course in thermodynamics and answered on average only two questions correctly. Thus, the thermodynamics data set for beta testing was smaller than the other two sets. As the instrument continues to be refined and used by more students, we intend to continue monitoring reliability and validity statistics.

As indicated in Tables 6-8, the number of correct responses for each item ranges from 25 to 80 percent with average scores of 52 percent correct for the heat transfer items, 41 percent correct for the fluid mechanics items and 55 percent correct for the thermodynamics items. These data combined with student interviews and open-ended written reponses indicated that engineering students still possess a wide variety of strongly held misconceptions in thermal and transport science and that these misconceptions can be identified using the TTCI.

Instrument reliability. Reliability for each part of the instrument was computed using Kuder-Richardson KR-20 reliability coefficients. A KR-20 coefficient is a special case of Cronbach's alpha used when data are dichotomous (i.e. right or wrong) and is computed using correct/incorrect scoring data for each student respondent (that is, the computations are made without regard to which distractor a student has chosen if he or she answered an item incorrectly). These results are shown in Table 9 and suggest that, based on limited data, each part of the instrument is showing reliability values among domain-related items which are slightly lower than we would like (and lower than guidelines for achievement tests of 0.7 for groups of students and 0.8 for individual student assessment) (Doran et al. 2002).

We speculate that there are at least two issues working against higher KR-20 values for the current version of the TTCI. First, each domain covered by the instrument includes several related but not equivalent concepts (i.e the heat transfer questions span several concepts including energy, heat, temperature, internal energy, enthalpy, and thermal equilibrium) for which students may possess varying degrees of understanding or misconception. For the TTCI, several new pairs of items focusing on specific heat transfer concepts have been written and tested. In each case, initial student respond data gives correlation coefficients ranging between 0.78 and 0.86, indicating that misconceptions in narrowly focused conceptual subdomains can be reliabily identified. More focused concept inventories such as the statics concept inventory concentrate on a narrower conceptual domain and report reliability coefficients approaching 0.9 (Steif and Dantzler 2005).

Second, by intentionally including very difficult questions (as opposed to achievement or diagnostic instruments in which most of the questions are easier), the overall reliability of the instrument can be expected to be lower since students will more likely guess on some questions or at least choose an answer in which they are not completely confident (Kline 2005).

Table 9. KR-20 Values for each Part of TTCI Version 2.5

Question category	KR-20 value
heat transfer (5 items, 7 total questions)	0.66
thermodynamics (8 items, 13 total questions)	0.63
fluid mechanics (8 items, 11 total questions)	0.56

Instrument Validity. Although several types of validity have been proposed to establish instrument performance, we focused on the two most relevant to a concept inventory: content and construct. Other important validity measures such as predictive validity have not been computed because they apply more directly to achievement tests. Predictive validity measures the ability of an instrument to predict differences in performance for use is assigning course grades or ranking students. Our concept inventory is most appropriately used as an assessment for identifying the existence of important misconceptions and for pre-post studies of misconception repair for a specific student population.

Content validity focuses on the question of whether items included in the instrument span the appropriate and desired technical domain. During TTCI development, we have worked in several ways to ensure content validity. The Delphi process discussed earlier represented an intensive exercise to identify important concepts which are poorly understood by students in thermal and transport sciences as defined by a panel of experts in these domains. This exercise achieved consensus about the key concepts and related misconceptions to be included in the instrument and informed us while individual test items were developed. As discussed earlier, an extensive literature search also confirmed that many of the difficult concepts identified in the Delphi study were also mentioned by previous researchers working in the heat transfer, thermo, and fluids domains. Thus, we claim that the conceptual domain covered by the TTCI is the domain we intended to cover. It is important to note, however, that we do not claim that the TTCI covers all important concepts in thermal and transport sciences but rather focuses on the key, but often misunderstood, concepts in these subject areas.

Construct validity attempts to answer the question of whether instrument items measure the concepts we think they are measuring. To address this issue, all items were drafted by domain experts and then reviewed by at least two additional experts who teach, conduct research, and in some cases write textbooks in the TTCI domains. Expert feedback was used to revise item wording and accompanying graphics as we strived for question correctness and clarity. In the case of each item, experts agreed that the intended concept and misconceptions were correctly targeted.

We also used think-aloud sessions with engineering juniors and seniors to confirm that students could identify the concept associated with each draft item, that the item text and graphics were understandable, and that each item raised a conceptual difficulty with most of the students interviewed. We recorded and coded all think-alouds to ensure a complete and accurate picture of student responses. Student responses gave us important feedback which, in some cases, uncovered items that were misunderstood because of poor question construction or graphics, or because we inadvertently used unfamiliar vocabulary, symbols, or notation.

As the database of TTCI student responses grows, we also plan to use item response theory to test specific items for bias. We will also conduct confirmatory factor analyses to confirm if our factor structure (i.e. number of concepts we believe we are testing) matches the data.

VI. FUTURE WORK

Based on the performance results of TTCI, version 2.5, and the need for additional questions and student response data, we have developed version 3.0 of the instrument which is now undergoing beta testing. A total of 24 new fluid mechanics, thermodynamics, and heat transfer items have been added focusing on key concepts including the “heat transfer rate vs. amount of energy transferred” concept discussed earlier as well as energy quality, fluid momentum conservation, equilibrium vs. steady-state processes, and the Bernoulli equation. Results from version 2.5 suggest that student misconceptions about these concepts are particularly robust.

Further work will also be done to determine if *levels* of student conceptual understanding can be recognized. Both Minstrel (1992) and Meyer and Land (2006) argue that conceptual understanding is not an all or nothing phenomenon, but might develop gradually. If this is the case, dichotomous measures like the KR-20 coefficient, which only take into account right or wrong responses, may not be the most appropriate measure for instruments like the TTCI. Therefore we plan to include use of pattern recognition methods such as Q-matrices (DiBello, Roussos, and Stout 2007) in future analyzes.

VII. SUMMARY AND CONCLUSIONS

This paper has reported on a project which investigates two research questions: 'What important concepts in thermal and transport science are difficult for engineering students to learn?' and 'Can an instrument be developed to measure misconceptions of these difficult and important concepts?' A list of important and difficult concepts in thermal and transport science was developed using Delphi study with expert engineering faculty and textbook authors as participants. Concepts identified align with the literature on scientific misconceptions reported in over 7000 studies complied by Duit (2007). The second question was investigated through the development and psychometric testing of the Thermal and Transport Concept Inventory. This instrument is designed to identify undergraduate engineering student misconceptions about important concepts in fluid mechanics, heat transfer, and thermodynamics. Instrument items were developed by drafting open-ended questions addressing each concept, collecting student response data orally (think-aloud problems solving sessions) and using written essays, using the responses to convert the open-ended questions to multiple choice items with plausible distractors as incorrect answers, collecting expert reviews on each item, beta testing the drafted items on groups of engineering students, revising the items based on statistical performance and expert comments and collecting additional beta test data. Student response data indicate that each item does identify student misconceptions but results of psychometric analysis of beta test data indicate that overall reliability within a conceptual domain (i.e. thermodynamics, fluid mechanics, and heat transfer) is not yet high enough to use the instrument for individual or group diagnostic or achievement assessment.

As a result, we envision the present version of the TTCI being used by engineering faculty in thermal or transport science courses as a formative assessment tool to detect the existence of important, robust misconceptions that their students hold at the beginning of a course. The instrument can also be used for pre- and post-testing to measure the degree to which misconceptions have been repaired in a course or degree program. Education researchers can also use the instrument for identifying the presence of robust misconceptions in thermal and transport sciences in students and as a tool for measuring the effect of interventions on students' conceptual change.

Acknowledgments

We wish to thank the National Science Foundation for supporting this work through grant number DUE-0127806. We also gratefully acknowledge discussions with Dr. Michelene Chi at the University of Pittsburgh and the participation of Delphi study participants.

References

Adler, M., and E. Ziglio, eds. 1996. *Gazing into the oracle: The delphi method and its application to social policy and public health*. London: Jessica Kingsley Publishers.

Arnold, M., and R. Millar. 1996. Learning the scientific "story": A case study in the teaching and learning of elementary thermodynamics. *Science Education* 80 (3):249-281.

Besson, U. 2004. Students' conceptions of fluids. *International Journal of Science Education* 26 (14):1683-1714.

Carlton, K. 2000. Teaching about heat and temperature. *Physics Education* 35 (2):101-105.

Clark, D., and D. Jorde. 2004. Helping students revise disruptive experientially supported ideas about thermodynamics: Computer visualizations and tactile models. *Journal of Research in Science Teaching* 41 (1):1-23.

Clayton, M.J. 1997. Delphi: A technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology* 17:373-386.

DiBello, L., L Roussos, and W. Stout. 2007. Review of cognitively diagnostic assessment and a summary of psychometric models. Invited chapter in *Handbook of Statistics, Volume on Psychometrics*, edited by C. V. Rao and S. Sinharay.

Doran, Rodney, Fred Chan, Pinchas Tamir, and Carol Lenhardt. 2002. *Science educator's guide to laboratory assessment* Arlington, VA: NSTA Press.

Duit, Reiner. *Teachers' and students' conceptions and science education*. http://www.ipn.uni-kiel.de/aktuell/stcse/. (cited May 18, 2007).

Evans, D. L., G. L. Gray, S. Krause, J. Martin, C. Midkiff, B. M. Notaros, M. Pavelich, D. Rancour, T. Reed-Rhoads, P. Steif, R. Streveler, and K. Wage. 2003. Progress on concept inventory assessment tools. *Proceedings, 33rd Frontiers in Education Conference*.

Fish, L.S., and D.M Busby. 1996. The delphi method. In *Research methods in family therapy*, edited by D. H. Sprenkle and S. M. Moon. New York: Guilford Press.

Hake, R. R. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics* 66 (1):64-74.

Hestenes, D., M. Wells, and G. Swackhamer. 1992. Force concept inventory. *The Physics Teacher* 30 (3):141-151.

Jacobi, A., J. Martin, J. Mitchell, and T. Newell. 2003. A concept inventory for heat transfer, *Proceedings, IEEE Conference*, Westminster, CO.

Johstone, A.H., and J.J McDonald. 1977. Misconceptions in school thermodynamics: An interpretive study. *Physics Education* 12:248-251.

Kariotogloy, P, P. Koumaras, and D. Psillos. 1993. A constructivist approach for teaching fluid phenomena. *Physics Education* 28:164-169.

Kesidou, S., and R. Duit. 1993. Students' conceptions of the second law of thermodynamics: an interpretive study. *Journal of Research in Science Teaching* 30 (1):249-281.

Kline, T.J.B. 2005. *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks: Sage.

Krause, Stephen, J. Chris Decker, Justin Niska, Terry Alford, and Richard Griffin. 2003. Identifying student misconceptions in introductory materials engineering classes. *Proceedings, 2003 American Society for Engineering Education Annual Conference*.

Linstone, H.A., and M. Turoff, eds. 1996. *The delphi method: Techniques and applications*. Reading:MA: Addison-Wesley Publishing Company.

Lord, F.M., and M.R. Novick. 1968. *Statistical theories of mental test scores*. Menlo Park: Addison-Wesley.

Martin, J., J. Mitchell, and T. Newell. 2003. Development of a concept inventory for fluid mechanics, *Proceedings IEEE Conference,* Westminster, CO.

Mazur, E. 1992. Qualitative vs quantitative thinking: Are we teaching the right thing? *Optics and Photonics News* 3:38.

Meyer, J.H.F. , and R. Land. 2006. Threshold concepts and troublesome knowledge: Issues of liminality. In *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge*, edited by J. H. F. Meyer and R. Land. New York: Routledge.

Midkiff, K. C., T. A. Litzinger, and D. L. Evans. 2001. Development of engineering thermodynamics concept inventory instruments, at Reno, NV.

Minstrell, J., R. Duit, F. Goldberg, and H. Niedderer Eds. 1992. Facets of students' knowledge and relevant instruction. *Proceedings, International Workshop – Research in Physics Learning: Theoretical Issues and Empirical Studies, The Institute for Science Education*.

Mulford, D. 1996. An inventory for measuring college students' level of misconceptions in first semester chemistry. Unpublished Master's Thesis, Purdue University.

Pedhazur, E.J., and L.P. Schmelkin. 1991. *Measurement, design and analysis: An integrated approach*. Upper Saddle, NJ: Prentice-Hall.

Psillos, D., and P Kariotogloy. 1999. Teaching fluids: Intended knowledge and students' actual conceptual evolution. *International Journal of Science Education* 21 (1):17-38.

Rhoads, Teri Reed, and Ron J. Roedel. 1999. Wave concept inventory - a cognitive instrument based on bloom's taxonomy. *Proceedings - Frontiers in Education Conference* 3:13-1.

Richardson, J., P. Steif, J. Morgan, and J. Dantzler. 2003. Development of a concept inventory for strength of materials, at Westminster, CO, USA.

Steif, Paul S., and John A. Dantzler. 2005. A statics concept inventory: Development and psychometric analysis. *Journal of Engineering Education* 94 (4):363-371.

Stone, A., K. Allen, T. R. Rhoads, T. J. Murphy, R. L. Shehab, and C. Saha. 2003. The statistics concept inventory: a pilot study, *Proceedings, IEEE Conference*, Westminster, CO, USA.
Streveler, R., T. A. Litzinger, R. Millar, and P. Steif. in press. Conceptual knowledge in engineering science: What engineering students find difficult and how we can help them learn on a conceptual level. *Journal of Engineering Education*.
Thomas, P. L., and R. W. Schwenz. 1998. College physical chemistry students' conceptions of equilibrium and fundamental thermodynamics. *Journal of Research in Science Teaching* 35 (10):1151-1160.
Thomaz, M.F., I.M. Malaquias, M.C. Vanente, and M.J. Antunes. 1995. An attempt to overcome alternative conceptions related to heat and temperature. *Physics Education* 30:19-26.
Wage, K. E., J. R. Buck, C. H. G. Wright, and T. B. Welch. 2005. The signals and systems concept inventory. *IEEE Transactions on Education* 48 (3):448-61.

Authors' Biographies

Ruth A. Streveler is an Assistant Professor in the School of Engineering Education at Purdue University. She holds a PhD in Educational Psychology from the University of Hawaii at Manoa, a MS in Zoology from The Ohio State University, and a BA in Biology from Indiana University. Prior to coming to Purdue, she was the founding Director of the Center for Engineering Education at the Colorado School of Mines. She has been a Co-PI on several projects funded by the National Science Foundation and served a stint as Acting Director of the Center for the Advancement of Engineering Education. She earned the 2006 and 2007 Helen Plants Awards for best non-traditional session at the Frontiers in Education Conference.

Address: School of Engineering Education, Purdue University, Neil Armstrong Hall of Engineering, 701 W. Stadium Drive, West Lafayette, IN 47904; telephone: (765) 496-9031; email: streveler@purdue.edu.

Ronald L. Miller is Professor of Chemical Engineering and Director for the Center for Engineering Education at Colorado School of Mines. He earned degrees in chemical engineering from the University of Wyoming and Colorado School of Mines. Dr. Miller has received three university-wide teaching awards and has held a Jenni teaching fellowship at CSM. He has received grant awards for educational research from the National Science Foundation, the U.S. Department of Education FIPSE program, the National Endowment for the Humanities, and the Colorado Commission on Higher Education and has published widely in the areas of engineering education assessment, pedagogy, and curricular design.

Address: Chemical Engineering Department, Colorado School of Mines, Golden, CO 80401; telephone: (303) 273-3892; email: rlmiller@mines.edu.

Mary A. Nelson is director of assessment for the Department of Applied Mathematics at the University of Colorado - Boulder. She holds a masters degree in Mathematics from George Mason University and a PhD in Research and Evaluation Methods from University of Colorado -

Boulder. Her current research interest is in how students develop conceptual understanding of basic concepts, and she is currently introducing oral assessments into large lecture calculus classes. Dr. Nelson was awarded the 2006 Peebles Award for excellence in teaching.

Address: Department of Applied Mathematics, UCB 526, University of Colorado, Boulder, 80309; telephone (303) 492-4152; email: mary.nelson@colorado.edu.

Monica R. Geist is a doctoral candidate in Applied Statistics and Research Methods with an emphasis in Research Methods at the University of Northern Colorado. She also teaches mathematics at Front Range Community College. She received the Apple Award for Excellence in Teaching from Phi Theta Kappa in 1998. She received the Faculty of the Year in 2005 and the Jerome Wartgow Teaching with Technology Award in 2006.

Address: Front Range Community College, 3645 W. 112th Avenue, Westminster, CO 80020; telephone: 303-404-5280; email: monicageist@comcast.net

Barbara Olds is Associate Vice-President for Educational Innovation and Professor of Liberal Arts and International Studies at the Colorado School of Mines. She recently returned to Colorado after spending three years at the U.S. National Science Foundation where she served as a Division Director in the Education and Human Resources Directorate. She is also currently a Visiting Professor in Purdue University's Engineering Education Department. Her research interests lie primarily in understanding and assessing engineering student learning. Her current research focuses on development of concept inventories for engineering topics. She has participated in a number of curriculum innovation projects and has been active in the engineering education research and assessment communities. She is a Fellow of the American Society for Engineering Education and was a Fulbright lecturer/researcher in Sweden in 1999.

Address: Trefny Institute for Educational Innovation, Colorado School of Mines, Golden, CO 80401; telephone: (303) 384-2410; email: bolds@mines.edu.