AC 2009-1795: DEVELOPMENT OF A CONCEPT INVENTORY IN HEAT TRANSFER

Michael Prince, Bucknell University

Margot vigeant, bucknell

Katharyn Nottis, Bucknell University

Development of a Concept Inventory in Heat Transfer

Keywords: heat, misconceptions, concept inventory

Abstract

Initial research with chemical engineering students suggests several areas where students appear to have robust misconceptions. In heat transfer, those areas include (1) temperature vs. energy, (2) temperature vs. perceptions of hot and cold, (3) factors which affect the rate of transfer vs. those which affect the amount of energy transferred and (4) the effect of surface properties on radiation. This study reports on the development of a concept inventory to assess these concept areas. Data was collected from approximately 400 chemical engineering students enrolled in about a dozen undergraduate programs over a two-year period of instrument development. Content validity was assessed by panels of engineering faculty who teach in these areas. Internal reliability was assessed through calculation of split-half reliabilities and KR20 (Kuder-Richardson Formula) values. Reliability was assessed for the instruments as a whole and for each specifically targeted misconception area.

Introduction

There is a growing recognition that students enter classrooms with preconceptions which act as filters for new learning (Smith, diSessa, & Roschelle, 1993). This prior knowledge can interfere with concept mastery. There is also a broad realization that meaningful learning of science content requires conceptual understanding rather than memorization of facts and formulas (Bransford, Brown, & Cocking, 2000; Lightman & Sadler, 1993), along with a growing appreciation that traditional instructional methods can be ineffective at altering students' preconceptions (Suping, 2003).

Engineering education has started to examine students' conceptual understanding and the instructional methods used in undergraduate courses. Guidance for addressing these issues in engineering education can be found in physics education (Hake, 1998; Laws et al., 1999). However, what has prevented engineering education from capitalizing extensively on the success in physics education has been the lack of knowledge of the relevant literature, the lack of concept inventories to assess conceptual understanding in engineering, and the lack of inquiry-based activities in engineering similar to those shown to be effective in physics. This study contributes by developing an assessment instrument for heat transfer, which is a required topic for chemical, mechanical and other engineering fields.

Confusion among concepts such as heat, energy and temperature is widely recognized in the literature (e.g., Carlton, 2000; Jasien & Oberem, 2002; Thomaz, Malaquiz, Valente, & Antunes, 1995). A Delphi study identified several concepts in thermal and transport

science that were both important and difficult for students to master (Streveler, Olds, Miller, & Nelson, 2003). While the Delphi study cited identified general areas of misconceptions, further research provided more specific information about possible misconceptions among undergraduate engineering students (Miller, Streveler, Olds, Chi, Nelson, & Geist, 2006; Prince & Vigeant, 2006). For example, it was found that engineering students had difficulty distinguishing between factors that affect the rate of heat transfer and those that affect the total amount of energy transferred in a given physical situation. Confusion in these areas was also found to persist, even when students successfully completed relevant coursework (Miller, Streveler, Olds, Chi, Nelson, & Geist, 2006).

The purpose of the current study was to develop a Heat Transfer Concept Inventory that could both document conceptual change and detect the presence of previously identified misconceptions. The current instrument was patterned after concept inventories developed in other disciplines such as the Force Concept Inventory in physics (Hestenes, Wells, & Swackhamer, 1992).

Identification of the Targeted Concepts and Associated Misconceptions

The concept inventory was designed to cover the following content areas identified as problematic through previous research (Miller, Streveler, Olds, Chi, Nelson, & Geist, 2006; Prince & Vigeant, 2006): (1) factors which affect the rate versus amount of heat transferred, (2) distinctions between temperature and heat, (3) distinctions between energy and temperature, (4) the relationship between heat transfer and temperature change, and (5) the effects of surface properties on heat transfer by radiation. Table 1 lists commonly found misconceptions in each of those areas.

Table 1: Initial Targeted Concept Areas (Phase 1)

Content Area	Misconception
1. Rate vs. Amount	Many students seem to believe that factors which increase the rate of heat transfer always increase the amount of heat transferred as well. These misconceptions carry over to related fields such as mass transfer.
2. Temperature vs. Heat	Many students think that temperature is a measure of how hot or cold things feel. Many students do not understand that other factors, such as the rate of heat transfer, frequently affect how hot or cold something feels.
3. Energy vs. Temperature	Students commonly believe that temperature is a direct measure of the energy in an object, so something at a higher temperature always has more energy.
4. Heat Transfer vs. Temperature Change	Students frequently believe that a change in temperature automatically tells you something about the rate or amount of energy transferred.

5. Radiation	Students are often confused about the effect of surface properties on the rate of radiative heat transfer.

In the initial phase of this research project, a concept inventory was developed which sought to measure conceptual understanding in each of these targeted areas.

Methodology

<u>Phase 1</u>

The concept inventory was composed of 28 multiple choice questions, several with open-ended segments which asked students to explain their thinking in more detail. These open-ended questions were intended to provide both richer assessment of students' conceptual understanding and to provide possible distractors on multiple choice questions in future development of the instrument. The 28 questions were broken down into 7 questions for concept area one, 4 questions for concept area two, 4 questions for concept area three, 7 questions for concept area four and 6 questions for concept area 5. Questions came from three main sources. Some questions were drawn (with permission) from those being developed and tested in drafts of the Thermal and Transport Science Concept Inventory (Miller, 2008) and the Heat Transfer Concept Inventory (Mitchell, 2007). New questions were also developed by the authors themselves.

The instrument was piloted with a sample of convenience of 119 undergraduate engineering students from four different institutions; 88 were chemical engineering majors, 31 were mechanical engineering majors. Approximately 56% were juniors, about 41% were seniors, and the remainder in other years of schooling. Seventy-nine (approximately 66%) were taking a course on heat transfer at the time they completed the concept inventory. The remainder had previously taken a course in heat transfer. Professors of classes where students were given the inventory were given detailed directions for administering the instrument in order to provide similar test-taking conditions among the different schools.

Faculty administering the concept inventory were also asked to complete a feedback form and to note whether each question assessed a particular concept; a ratio of agreement was calculated for the purposes of content validity. Questions with a low content validity ratio were targeted for significant revisions.

Results were also examined for reliability, with an eye towards improvements in subsequent rounds of testing. Classical Test Theory guided the item analysis, with a focus on item difficulty and item discrimination. It is recognized that a major limitation of using this theory is that both item difficulty and discrimination are dependent upon the participants (Fan, 1998), which is why care was taken to obtain a sample that would be similar to the group of students who might use the inventory in the future. "Difficulty" measures what fraction of students answers a given question correctly, while "Discrimination" correlates a students' score on a particular question with their score overall. Questions with high difficulty (close to 1.0) are answered incorrectly by the

majority of students, while questions with high discrimination index (close to 1.0) tend to be answered correctly only by those individual with high test scores overall. For the phase 1 instrument, item difficulties ranged from 19% of the entire sample getting a question correct to approximately 95% of the entire sample getting a question correct. Four questions had a negative discrimination index, meaning students who had high overall scores tended to answer those questions incorrectly. Eight other questions had a discrimination index of less than .20. These questions were targeted for revision or elimination.

Phase 2

Based on the results of Phase 1, several revisions were made to the instrument with the hope of improving the reliability numbers. The instrument itself was increased to 36 multiple choice questions, with no open-ended responses in order to reduce the time required for completion.

One of the general changes to the instrument was to revise and regroup the targeted concept areas, as shown in Table 2. As can be seen by comparison with Table 1, two previous concept areas were merged. This was done for two reasons. First and most importantly, it was thought that there was significant overlap between these two, since misconceptions regarding a change in temperature and when energy is transferred depends on the more fundamental and direct relationship between the concepts temperature and energy. Secondly, reducing the number of misconception area allowed the test to increase the number of questions per misconception area without making the instrument too time consuming for students and instructors. Since reliability generally improves as the number of questions increases, this provided an opportunity to increase the reliability of specific sections of the instrument targeted for individual content areas.

Table 2: Targeted Conceptual Areas for Phase 2

Content Area	Misconception
1. Rate vs. Amount	Many students seem to believe that factors which increase the rate of heat transfer always increase the amount of heat transferred as well. These misconceptions carry over to related fields such as mass transfer.
2. Temperature vs. Heat	Many students think that temperature is a measure of how hot or cold things feel. Many students do not understand that other factors, such as the rate of heat transfer, frequently affect how hot or cold something feels.
3. Energy vs. Temperature	Students commonly believe that temperature is a direct measure of the energy in an object, so something at a higher temperature always has more energy.
4. Radiation	Students are often confused about the effect of surface properties on the rate of radiative heat transfer.

The new concept inventory was re-tested with a sample of convenience of 228 undergraduate engineering students from six different institutions. The questions were

grouped into 8 questions in concept area one, 9 questions in concept area two, 9 questions in concept area three and 12 questions in concept area four. Some questions were considered to assess two concept areas.

Sample Conceptual Questions and Instrument Revisions

Here we provide an example of questions used in both Phase 1 and Phase 2 to illustrate the structure of the problems and the types of revisions that were made during the development of the instrument. Example 1 below (Mitchell, 2007) is taken from the series of questions designed to assess students' understanding of the effect of surface properties on radiation.

Example 1:

Radiation Question: Phase 1

A person walks toward two diffuse, gray surfaces that are maintained at 1000K (see figure below). Surface 1 has an emittance of 0.95. Surface 2 has a reflectance of 0.95.

Surface 1
T=1000K
Emittance = 0.95

Surface 2
T=1000K
Reflectance = 0.95
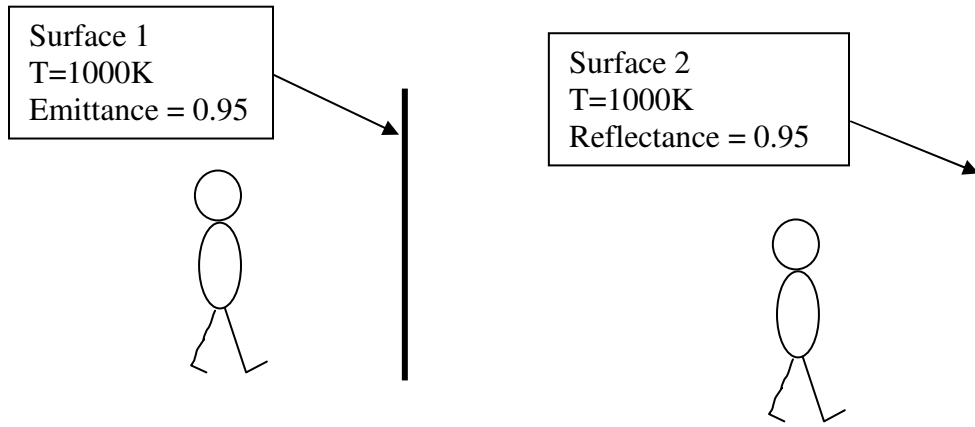
Question 7. Which statement is true?
     a) The person will feel warmer as they approach surface 1
     b) The person will feel warmer as they approach surface 2
     c) The person will feel the same warmth in both cases.
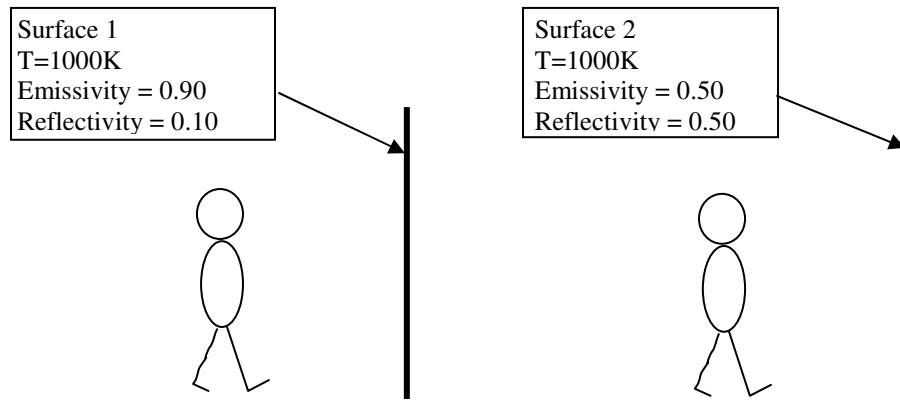   d) Not enough information given

Student results indicated that this question was quite difficult, with only 25% of students identifying the correct response. In addition, the question was a poor discriminator between students who did well on the instrument and those who did not, with a discrimination index that was actually negative (-0.08). Because of this, the question was significantly revised in Phase 2, as shown below.

Radiation Question: Phase 2

A person walks toward two diffuse grey surfaces that are maintained at 1000°K (see figure below).

Surface 1 has an emissivity of 0.90 and a reflectivity of 0.10
Surface 2 has an emissivity of 0.50 and a reflectivity of 0.50


```
Surface 1
T=1000K
Emissivity = 0.90
Reflectivity = 0.10
```

```
Surface 2
T=1000K
Emissivity = 0.50
Reflectivity = 0.50
```

Question: Which statement is true?
   a)   The person will feel warmer as they approach surface 1
   b)   The person will feel warmer as they approach surface 2
   c)   The person will feel the same warmth in both cases.
   d)   Not enough information given

The changes were made to make the question a more direct assessment of the impact of emissivity on the amount of radiation heat transfer. The first version of the question required students to *infer* that the emissivity of surface 2 was less. The second version of the question provides this information explicitly, and so is a better measure of students' understanding of the effect of emissivity on radiation heat flux. Results indicate that this change was effective. On the revised question, 57% of students were able to answer it correctly and the discrimination index went from -0.08 to 0.32, a significant improvement.

Example 2:

Temperature vs. Energy Question: Phase 1

Question: Assuming either stream below leaves a turbine at the same conditions, which stream has the potential to produce more total electricity in such a turbine?

   Stream 1: Steam flow rate = 10 kg/s.   Temp. = 200C   Pressure = 2 atm
   Stream 2: Steam flow rate = 100 kg/s   Temp = 190 C   Pressure = 2 atm.

a. Stream 1 has the potential to produce more electricity
b. Stream 2 has the potential to produce more electricity
c. Either stream has the potential to produce the same amount of electricity
d. Not enough information given

Explain your reasoning.

This question was somewhat easy, with 71% of students being able to answer it correctly. However, it was a poor discriminator with a discrimination index of only 0.02. Because of this, it was significantly revised for phase 2 testing. Note that the question included an open-ended response in addition to the multiple choice section. Student responses to this question were used to craft distractors to the question for phase 2 testing.

Temperature vs. Energy Question: Phase 2

Question: High pressure steam is commonly used to produce electricity by expansion through a turbine. Assuming that either stream below leaves the turbine as a liquid at 100ºC and atmospheric pressure, which stream has the potential to produce more total electricity in the turbine?

Stream 1: Steam flow rate = 10 kg/s.   Inlet T = 200ºC   Inlet P = 2 atm
Stream 2: Steam flow rate = 100 kg/s   Inlet T = 190ºC   Inlet P = 2 atm.

a. Stream 1 because a higher temperature yields a higher engine efficiency
b. Stream 1 because the potential electricity produced depends primarily on the temperature difference between the inlet and outlet temperatures
c. Stream 2 because the higher mass flow rate more than compensates for the slightly lower temperature
d. Either stream has the potential to produce the same amount of electricity because this depends primarily on the inlet and outlet pressures

Results indicated that the revision successfully achieved its goal. While the difficulty of the question remained fairly low, with 78% of students answering it correctly, the discrimination index did improve from 0.02 to 0.33.

Results

The main focus of our initial work on this project was to develop an instrument with acceptable internal reliability. The bulk of results presented relate to that objective. In addition, the instrument's main purpose is to assess the conceptual understanding held by undergraduate engineering students and preliminary results related to student performance on the instrument are also presented.

Reliability Results

Phase 1

Internal reliability was determined for the entire inventory and for sub-test questions, grouped by content areas, since it was anticipated that future users of the instrument might not always use the entire inventory. Two measures of internal reliability were used, split-half reliability and KR20 values. Split-half is a measure indicating to what extent the results on a randomly selected half of the questions mirror the results in the remaining half (1.0 = perfect symmetry). KR20 is a measure of test reliability and internal consistency, with 1.0 indicating perfect consistency. For this instrument, minimum acceptable overall value for the instrument was set at 0.60 on both measures and the target reliability was set by the authors to be 0.70 on both measures.

Split-half reliability for the entire instrument was 0.61 and KR20 was 0.64, which may be considered acceptable for research purposes. Removing questions with negative discrimination indices in an attempt to increase the instruments' reliability did not significantly alter the analysis (KR20 = 0.67, split half reliability = 0.53). The instrument was found to have poor reliability with the mechanical engineering sample, indicating there may be differences in content emphasis in heat transfer courses taught in engineering, although additional testing should be done before drawing a firm conclusion. When reliability was calculated using only the results of the chemical engineering majors the reliability did increase marginally. The split-half reliability was 0.63 and the KR20 was 0.70.

In addition to considering the instrument as a whole, reliabilities were also calculated for each individual content area of the instrument. These reliabilities, shown in Table 3, had much more variability and were generally less acceptable. The exception was concept area 1 (rate vs. amount) which had a split half reliability of 0.79 and a KR20 of 0.74. This was a positive finding indicating that this section of the exam had good reliability and did not require significant revision. However, reliability numbers for the other sections of the instrument were notably poorer. The results for radiation were particularly discouraging with negative reliability values, indicating the need for significant revisions.

Table 3. Reliability Results for Phase 1 Instrument by Content Area

Content Area	KR20	Split Half
1. Rate vs. Amount	0.74	0.79
2. Temperature vs. Heat	0.44	0.25
3. Energy vs. Temperature	0.22	0.13
4. Heat Transfer vs. Temperature Change	0.16	0.16
5. Radiation	-0.08	-0.10
Overall	**0.64**	**0.61**

Phase 2

Reliability results for the instrument in Phase 2 improved, as shown in Table 4. Overall reliability of the instrument increased significantly, with a KR20 of 0.83 and a split-half reliability of 0.80. In addition to the improvement in the reliability of the overall instrument, the reliability measures for each subcategory also increased significantly, with the exception of category 1 which was not revised significantly because of its good performance in phase 1. The improvement in the radiation subsection of the instrument was particularly noteworthy. Improvements in reliability may be attributed both to the increased number of questions per content area in some cases and the revision of questions with poor difficulty levels or discrimination values from Phase 1 testing.

Table 4. Reliability Results for Phase 2 Instrument by Content Area

Content Area	KR20	Split Half
1. Rate vs. Amount	0.79	0.75
2. Temperature vs. Heat	0.55	0.51
3. Energy vs. Temperature	0.54	0.46
4. Radiation	0.61	0.57
Overall	**0.83**	**0.80**

Student Performance Results

In addition to presenting psychometric analysis of the instrument itself, it is of interest to look at what the instrument tells us about the degree of conceptual understanding of engineering students. Since the instrument developed in Phase 2 showed superior reliability results both for the instrument overall and for individual concept areas, only students results obtained with this instrument in Phase are presented. Student performance in Phase 2 for the instrument as a whole and in individual concept areas is presented in Table 5.

Table 5. Student Performance Data from Phase 2

Content Area	Student Performance (% correct)
1. Rate vs. Amount	28%
2. Temperature vs. Heat	60%
3. Energy vs. Temperature	54%
4. Radiation	45%
Overall	**50%**

The results suggest that students have significant misconceptions in these areas, given that they were only to answer half the question correctly on average. In addition, as suspected from research early in the project, student performance suggests significant difficulty discriminating between those factors which increase the rate of heat transfer and those which increase the total amount of heat transferred in a given situation.

It will also be interesting to look at changes in student performance on the instrument before and after taking the relevant undergraduate heat transfer course. Such data for 228

engineering students at 6 institutions has been collected as part of this study, but has not yet been analyzed. These results will be available for presentation at the conference.

Conclusions

A concept inventory has been developed to assess engineering students' understanding of several important concepts in heat transfer. The current instrument assesses student understanding in four areas: (1) factors which influence the rate and amount of heat transfer, (2) distinctions between temperature and heat, (3) distinctions between temperature and perceptions of hot and cold and (4) the effect of surface properties on radiation. Each of these content areas was determined to be an important and difficult area of heat transfer by both an expert panel of engineering faculty and through initial testing with undergraduate engineering students.

The instrument has now undergone several tests and two formal phases of assessment. Results indicate that while the initial measures of reliability were marginally acceptable for the instrument as a whole and unacceptably low for some of the individual concept areas tested, the revised instrument now shows good performance with very good overall reliability numbers and much better reliability for each of the individual concept areas tested.

Initial results from student performance data indicate that undergraduate engineering students do have significant misconceptions in these areas. Future work will demonstrate the impact of classroom instruction on improving their understanding. In addition, work is currently underway to develop and test inquiry-based activities to repair misconceptions identified in this study.

References

Bransford, J., Brown, A., & Cocking, R. (2000). *How people learn: Body, mind, experience and school.*

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

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, *58*,357(25).

Hake, R. *Interactive-Engagement Vs. Traditional Methods: a Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses.* American Journal of Physics, 1998. **66.** 64.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141-158.

Jasien, P., & Oberem, G. (2002). Understanding of elementary concepts in heat and temperature among college students and K-12 teachers. *Journal of Chemistry Education, 79*, 889.

Laws, P., D. Sokoloff, and R. Thornton. *Promoting Active Learning Using the Results of Physics Education Research.* UniServe Science News, 1999. **13.**

Lightman, A., & Sadler, P. (1993). Teacher predictions versus actual student gains. *The Physics Teacher, 31*, p. 162.

Miller, R. L., Streveler, R. A., Olds, B. M., Chi, M. T. H., Nelson, M. A., & Geist, M. R. (2006). *Misconceptions about rate processes: Preliminary evidence for the importance of emergent conceptual schemas in thermal and transport sciences*. ASEE National Conference Proceedings, Chicago, IL.

Miller, R. L., (2008). Personal Communication. Thermal and Transport Science Concept Inventory, versions 2-3.04.

Mitchell, J., (2007). Personal Communication. This question was drawn from a 2004 draft version of the HTCI under development by John Mitchell and Jay Martin at the University of Wisconsin

Prince, M., & Vigeant, M. (2006). *Using inquiry-based activities to promote understanding of critical engineering concepts.* Paper presented at the ASEE National Conference, Chicago, IL.

Smith, J. P. diSessa, A. A., & Roschelle, J. (1993).Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning*

Sciences, *3*(2), pp. 115-163.

Streveler, R., Olds, B., Miller, R., & Nelson, M. (2003). *Using a Dephi Study to identify the most difficult concepts for students to master in thermal and transport science*. Paper presented at the annual meeting of the American Society of Engineering Education, Nashville, TN.

Suping, S. (2003). Conceptual change among students in science. *Eric Digest*.

Thomaz, M., Malaquiz, I., Valente, M., & Antunes, M. (1995). An attempt to overcome alternative conceptions related to heat and temperature. *Physics Education, 35*, 101.