

Learning and Conceptual Change in Thermal Physics Concepts: An Examination by Gender

Umairia Malik^a, Elizabeth J. Angstmann^b, Kate Wilson^c

Corresponding author: Umairia Malik (u.malik@student.adfa.edu.au)

^aSchool of Physical, Environmental, and Mathematical Sciences, UNSW Canberra, Canberra ACT 2612, Australia

^bSchool of Physics, The University of Sydney, Sydney NSW 2006, Australia

^cSchool of Engineering and Information Technology, UNSW Canberra, Canberra ACT 2612, Australia

Keywords: gender, TCS, heat transfer, normalised learning gain

International Journal of Innovation in Science and Mathematics Education, 27(1), 37–46, 2019

Abstract

Students enter physics classes often having an understanding of basic concepts which is incomplete or incorrect. In this study, we investigate the performance difference between males and females on one particular question of the Thermodynamics Concept Survey (TCS). The question evaluates student understanding of the heat (energy) that is transferred during change of state and change of temperature. We find that males outperform females, in the pre- and post-instruction tests. However, using transition matrices with the hierarchy of answer choices, reveals that females are more likely than males to improve their understanding from an incorrect answer to a better incorrect answer. This improvement is not captured in the facilities, or normalised learning gains. Hence while the gap in conceptual understanding is reduced after the instruction, this is not rewarded in the usual binary multiple-choice scoring system.

Introduction

Students have a natural tendency to relate scientific concepts with the physical world. During their interaction with everyday phenomena, students develop naïve concepts based on their observations and experiences and these naïve concepts may differ substantially from scientific concepts. Later, these prior concepts influence a student's learning which involves structural organisation of knowledge as a student fits new concepts into existing conceptual understanding. If the existing concepts are based on flawed understanding, it hinders a student's ability to integrate and understand what is being taught (Driver, 1989). Chi (2008) distinguishes between three types of learning. First, a student may have no knowledge. In this situation, prior knowledge is *missing*. Second, a student may have some correct knowledge, but that knowledge is incomplete. In this case, learning is considered as *filling* the gaps in the student's existing correct knowledge. Third, a student may have gained some ideas either during school education or from everyday life, which contradict scientific concepts. In this situation, learning is *changing* prior misconceptions ideas to correct scientific knowledge.

In order to assess students' conceptual understanding, various methods have been developed for identifying alternative concepts (or misconceptions), including, interviews (Gilbert, Osborne, & Fensham 1982), open-ended questions (Treagust 1986), and Concept Inventories (CIs) (Hestenes, Wells, & Swackhamer 1992). The most efficient method to test students' conceptual understanding is to apply CIs. These are multiple-choice assessments that are relatively easy to implement, statistically reliable, and provide quantitative results that may be

used for comparison of understanding before and after instruction, or of different student cohorts. Hence CIs can be used to evaluate the effectiveness of teaching methods and determine students' prior knowledge. For example, Hake (1998) used the Force Concept Inventory (FCI) to compare the effectiveness of interactive engagement methods versus traditional teaching methods in introductory physics courses. Since Hake's seminal paper, considerable work has been done on the development of inventories in physics education.

CIs have also been used to explore the performance difference between male and female students. A pattern of female students having lower pre-instruction test scores than their male peers in the CIs has been revealed in physics education research. For example, Docktor and Heller (2008) reported that on the FCI pre-instruction test scores, male students outperformed female students by 15%, and by 13% on the post-instruction test. However, there was no difference in course grade. One of the possible reason for females' underperformance is their lower self-efficacy compared to males (Lindstrøm & Sharma 2011).

Several studies have found that students from an early age hold misconceptions related to thermal physics (see for example Georgiou and Sharma (2010)). Paik, Cho and Go (2007) observed misconceptions related to heat and temperature among 4 to 11 year old Korean students. Many Korean students had the notion that the temperature of an object was related to its size, or material properties. However, these misconceptions diminished with student age. In another study, Harrison, Grayson and Treagust (1999) found that grade 11 students viewed temperature and heat as equivalent entities. Furthermore, Yeo and Zadnik (2001) found that when students were asked about heat and temperature many held the misconception that heat is proportional to temperature and a cold body does not contain any heat. They also found that some students believe that water can't exist at 0°C. These studies highlighted the difficulties faced by students of various age groups and countries in understanding thermal physics concepts.

As with mechanics CIs such as the FCI (Docktor & Heller, 2008), gender differences have also been observed on thermodynamics CIs. Nottis et al. (2017) found that in the Heat and Energy Concept Inventory (HECI) developed by Prince, Vigeant and Nottis (2012), in the overall test, males got higher average scores than females after instruction. In addition, Nottis, Prince and Vigeant (2017) investigated the performance difference between male and female students by categorising the HECI questions into four sub-sets: rate vs amount of heat transferred; energy vs temperature; temperature vs perception of hot or cold; and thermal radiation. They found that in the pre-and post-test scores, there was no significant difference between male and female students' performance on the rate vs amount of heat transferred questions. However, the average score of male students was significantly higher than those of females in the sub-sets of energy vs temperature and thermal radiation questions on the pre-instruction and post-instruction tests. In addition, there was no difference between the average male and female scores on the temperature vs perceptions of hot or cold in the pre-instruction test but male students outperformed females in the post-instruction test. Furthermore, Nottis et al. (2017) found that a question which deals with the heat transfer concept (question 22), the proportion of females giving the correct answer decreased by 7.9% after instruction while the proportion of males giving the correct answer increased by 8.3%.

In this study, we used the Thermodynamics Conceptual Survey (TCS) developed by Wattanakaswich, Taleab, Sharma and Johnston (2013) and it was applied in an introductory calculus-based Physics course at UNSW Sydney in 2016. The purpose of this study is to investigate the use of the TCS in measuring conceptual change in understanding

thermodynamics concepts and performance difference between male and female students on the TCS. We use a sample the students who took the TCS during their Physics 1A course at UNSW Sydney in 2016. In particular, we are interested in exploring the performance difference on question 4 of the TCS, which deals with the heat transfer concept.

The thermodynamics conceptual survey and question of interest

Among other CIs that have been developed over time, Wattanakasiwich et al. (2013) introduced the TCS. The TCS is an instrument which requires minimal computation and succinctly assesses a student's conceptual understanding in thermal physics. The TCS can be used to determine the impact of teaching on students' conceptual understanding of thermodynamics. The TCS comprises 35 multiple-choice questions and can be divided into two parts. The first part covers the concepts of temperature, heat transfer, and the ideal gas law while the second part covers the first law of thermodynamics.

The question we are considering in this study is Question 4, as shown in Figure 1, which comes under the first part of the TCS. This question is of interest because there is a large gender gap in the proportion of correct answers on both the pre-instruction and post-instruction test. The answer choices of males and females can tell us about their different conceptual understanding, and which misconceptions they hold. Comparing answer choices before and after instruction can tell us how students' conceptual understanding changes as a result of instruction, and whether this is different for males and females.

Question 4

If 100 grams of ice at 0°C and 100 grams of water at 0°C are put into a freezer, which has a temperature below 0°C . After waiting until their temperature equals to the freezer temperature, which one will eventually lose the greatest amount of heat?

- a. The 100 grams of ice.
- b. The 100 grams of water.
- c. They both lose the same amount of heat because their initial temperatures are the same.
- d. There is no answer because ice does not contain any heat.
- e. There is no answer because you cannot get water at a temperature of 0°C .

Figure 1: Question 4 of the TCS

The question evaluates students' understanding of the heat (energy) transfer during change of state and change of temperature, by asking students to compare the heat lost when ice and water start at the same temperature, and the temperature is then reduced to well below the freezing point of water. Looking at answer choices and conceptions related to them, we find that option B, the correct answer, represents an understanding of energy changes involved in both cooling and phase changes. Option C indicates an understanding of temperature, heat and energy when cooling is occurring, but not when a phase change is taking place. Option E indicates a "common sense" idea of water as existing only above 0°C , as seen by Yeo and Zadnik (2001), and a lack of understanding of both temperature as energy per particle, and of energy as being gained/lost in a phase change. Option D shows the lack of understanding of the relationship between heat and temperature, and is the same misconception identified by Yeo and Zadnik

(2001) that a cold body contains no heat. There is no specific misconception associated with option A.

Our students and physics course

UNSW Sydney provides an introductory calculus-based physics courses called Physics 1A. In this paper, we investigate the cohort who took the course in semester 1 2016, which numbered approximately 1600 students (23% females, 77% males). Most of these students (75%) are taking an engineering course, and the proportion of males and females in the engineering cohort matches the overall cohort distribution (i.e. there is no gender bias in the cohort by degree stream). Anecdotally, one quarter of the cohort are international students but due to privacy constraints we do not have demographic data by this measure. The Physics 1A course runs over twelve weeks. It consists of six contact hours each week, three one-hour lectures, a two-hour laboratory session and one-hour problem solving workshop.

In the thermal physics course, students complete two laboratory exercises. One of these exercises is related to specific and latent heat of water. Students come up with their own method for measuring the latent heat of fusion of ice. In the other exercise, students use a syringe with a temperature and pressure sensor to observe effects on volume and temperature of changing the pressure of air. Students also take air through a cycle, adiabatic compression, isovolumetric cooling then isothermal expansion, recording measurements and sketching a PV plot. Furthermore, problem solving workshops consists of two classes in which students go through worksheets based on “Tutorials in Introductory Physics” (McDermott & Shaffer, 2002).

Method

Students were given the TCS as the pre-instruction test in the sixth week of lectures, just before they started the thermal physics topic. The TCS consists of 35 questions. In the pre-instruction test, only 15 questions were given from the TCS; those that dealt with concepts with which students may have already been familiar. Questions about concepts to which students had not been exposed, for example, P-V plots for gasses, or which covered the same concepts as other questions, were excluded from the pre-instruction test. The post-instruction test was conducted at the end of 12th week of lectures. In the post-instruction test, students were given all 35 questions of the TCS. A bonus point was awarded to students who completed both the pre-and post-test. This had a minimal impact on final grades of students, giving them less than 0.3 marks towards their final grade. In our analysis, we included only those students who gave us permission to use their responses.

In the Physics 1A course, there was a total of 1071 students who submitted the pre-instruction test, and 825 students who submitted the post- instruction test. The analysis included only those students who took both the pre-instruction and post-instruction tests. In the matched data set of the pre-instruction and post-instruction tests, there were 672 students, 499 (74%) males, and 173 (26%) females, closely matching the total cohort mentioned earlier. The proportion of engineers in the matched dataset remained effectively the same (76%) with a gender split equivalent to the entire matched dataset (i.e. matching the data in this way has not introduced a gender bias or a bias by degree stream).

For the comparison of overall performance in the pre-and post-instruction tests, we considered only the 15 questions that appeared on both tests. We calculated the facility (fraction of cohort answering correctly), gap ($[\text{male facility}] - [\text{female facility}]$), and Normalised Learning Gains

(NLGs) (Hake, 1998) for all 15 questions combined. We also calculated facility, gap, and NLGs for each question individually.

To investigate the change in answer choices in question 4 from the pre-instruction test to post-instruction test, we used transition matrices (Morris, Walter, Seeks, & Schwartz 2017).

Results

In this analysis, only matched data was used from students who completed the 15 questions during the pre-instruction and post-instruction test. In the pre-instruction test, the average facility for males was 0.62 and for females was 0.55. In the post-instruction test, the average facility for both males and females increased, for males to 0.67 and for females 0.60. Although the average facilities increased from the pre-instruction test to post-instruction test, the average gap, 0.06, remained constant. The average facilities for males are higher than for females in the pre-and post-test which indicates that females are underperforming in the TCS overall.

Question of interest

In the pre-instruction test, the complete cohort facility on question 4 was 0.39 and the overall facility for males was 0.42 and for females was 0.31, giving a gap of 0.11. In the post-instruction test, the overall facility of the cohort increased to 0.61, with male facility increasing to 0.65 and female facility to 0.55, giving a gap of 0.10 as shown in Table 1. Hence the facility on this question increased with instruction by 23% for males and 24% for females: effectively the same absolute gain, resulting in the performance gap being maintained. This question showed a large and significant gap, which was maintained in spite of instruction. The NLGs for males was 0.39 and for females was 0.35, so by this measure the instruction was no more effective for females than for males (values are rounded to the two significant figures for presentation purpose).

Table 1: Results of pre-test and post-test of question 4

Test	Male		Female		Gender Gap
	Facility	Fraction Answering C	Facility	Fraction Answering C	
Pre	0.42	0.46	0.31	0.49	0.12
Post	0.65	0.20	0.55	0.31	0.10

To answer this question correctly, a student must apply knowledge of temperature, heat, and energy when cooling and change of phase occur. Among all other answer choices, option C was the common choice for both males and females in the pre-instruction test and post-instruction test. However, females chose option C more frequently than males, causing the significant, large gender gap.

We explored answer choices of the pre-instruction and post-instruction matched data set of students responding to question 4 of the TCS. A transition matrix is a useful tool which provides insight into students' conceptual change due to instruction. We categorise answer choices based on associated misconceptions and rank the answer choices accordingly. Through the categorisation, we can explore how students' conceptual understanding is changing by their changes in answer choices.

The categorisation of answer choices is based on the above discussion, where lack of information or misconceptions associated with each distractor are identified. Based on this, the answers are ranked from best to worst as follows.

1. Best answer: option B – correct answer.
2. Option C – correct but incomplete knowledge.
3. Options D and E – each represents a particular misconception as described by Yeo and Zadnik (2001); or option A – incorrect answer, no specific misconception.

Table 2a represents the transition matrix of question 4 for male students. A transition matrix shows how the selection of answer choices changes between pre- and post-instruction tests. The rightmost column shows the proportion of male students choosing each option in the pre-instruction test, the last row shows the proportion of male students choosing each option in the post-instruction test. Within the body of the table, each entry shows the proportion of male students choosing a particular combination of pre-instruction and post-instruction answer choices. In the transition matrix, answer choices are ranked as described above (noting that A, D and E are all considered of equal rank, but are presented separately in the table). If we look at the first row in the body of the table (these students who answered A in the pre-test), we can see the proportion of the cohort that answered each option in the post-test. For example, 2% of the entire cohort answered A in the pre-test then choose B in the post-test.

Table 2a: Transition matrix for question 4 for male cohort

Test	Post					
Pre	A(49)	E(15)	D(8)	C(103)	B(324)	Total
A(26)	1.4%	0.0%	0.2%	1.6%	2.0%	5.2%
E(32)	0.6%	1.2%	0.0%	1.4%	3.2%	6.4%
D(5)	0.0%	0.0%	0.4%	0.4%	0.2%	1.0%
C(225)	5.0%	1.6%	0.6%	13.2%	24.6%	45.1%
B(211)	2.8%	0.2%	0.4%	4.0%	34.9%	42.3%
Total	9.8%	3.0%	1.6%	20.6%	64.9%	100.0%

The cells highlighted green in table 2a show the proportion of males who moved from an incorrect answer choice to the correct answer (dark green) and those moved a step forward from an incorrect answer (A, D or E) to better incorrect answer (C) (light green). These students have gained knowledge, and for those that have answered D or E in the pre-instruction but choose option B or C in the post-instruction, they have also overcome their previous misconceptions.

The entries highlighted red in table 2a represent a step backwards in conceptual understanding, deep red for a change from fully correct answer (B) to any incorrect answer, and pink for partially correct answer (C) to incorrect answers (A, D or E). These students have moved from either the correct answer (B) or the correct and incomplete answer choice (C) to an answer which implies either a specific misconception (D or E) or lack of knowledge or confusion (A).

Entries highlighted yellow show students who have begun and ended up with same misconception and/or lack of knowledge or confusion. These students did not gain from instruction. It is interesting to note that students who began with one particular misconception did not change to the other identified misconception, i.e. there are no students who transitioned from D to E or vice versa.

Finally, the cell highlighted blue is those students who already understood the relevant concepts, and answered correctly in both pre-instruction and post-instruction tests (nothing to gain). The same colour coding is used in table 2b, which shows the transition matrix for female students.

As we can see, after instruction, the proportion of males and females answering correctly increased substantially. In the pre-instruction test, option C was a common distractor for both males and females, but females were also more likely to choose option E as shown in Table 2b. In the post-instruction test, the proportion of both males and females choosing options C and E decreased, in contrast to this, the proportion choosing option A more than doubled.

When we compare the transition matrices for male and female students, we can see that more males than females gave the correct answer on both the pre-instruction and post-instruction tests (34.9% compared to 19.7%). These students did not improve on this question because there was no room for improvement—they were already choosing the best answer.

Table 2b: Transition matrix for question 4 for female cohort

Test	Post					
Pre	A(16)	E(5)	D(1)	C(56)	B(95)	Total
A(7)	1.2%	0.0%	0.0%	1.2%	1.7%	4.0%
E(27)	1.7%	1.7%	0.0%	6.4%	5.8%	15.6%
D(1)	0.0%	0.0%	0.0%	0.6%	0.0%	0.6%
C(85)	4.6%	0.6%	0.0%	16.2%	27.7%	49.1%
B(53)	1.7%	0.6%	0.6%	8.1%	19.7%	30.6%
Total	9.2%	2.9%	0.6%	32.4%	54.9%	100.0%

The remaining 65.1% of males and 80.3% of females either improved, went backwards or failed to improve when there was a room for improvement.

Summing the values in the green cells shows that 43.4% of females and 33.4% of males improved their conceptual understanding, moving either from any other answer to the correct answer or from any incorrect answer (A, D or E) to the partially correct answer (C). This improvement for females compared to males was significant by using Z-test at 5% level (p -value ≤ 0.01)

In comparison, 14.6% of males and 16.2% of females changed their answer either from the correct answer (B) to any answer choice or from partially correct (C) to one of the incorrect answers (A, D or E). Hence substantially more students, of both genders, have gone forwards

in their conceptual understanding than have gone backwards. Of these, were 7.4% of males and 11% of females who moved from the correct answer to an incorrect answer following the instruction.

The remaining 17% of males and 20.8% of females, do not appear to have either improved or decreased their conceptual understanding, based on their answer choices for this question.

Discussion

The results of this study show that males outperform females on the overall TCS. This finding is consistent with the Nottis et al. (2017) study that found that males performed better than females in the thermodynamics concept inventories. One of the possible reasons for this may be that multiple-choice assessments or questions with a lack of context may disadvantage female students (Hazel, Logan, & Gallagher, 1997).

In question 4, students were asked about the amount of heat loss by two materials (water and ice) during their temperature equalizes to a freezer's temperature. The correct answer requires understanding of heat transfer as water changes state from liquid water to ice and during cooling. We found that from the pre-instruction test to post-instruction test, the proportion of correct answers increased for both males and females. This finding was different from the Nottis et al. (2017) study, in which an analysis of a question which assessed the same concepts (question 22 of the HECI), showed that the proportion of females answering correctly decreased after instruction.

In our study, almost half the male and female students in the pre-instruction test chose option C (46% of males and 49% of females). This shows that students were not thinking about heat transfer during changes of state (freezing), but only during temperature change (cooling). One of the possible explanations for this, as Rozier and Viennot (1991) presented, is that in questions which require incorporation of a relationship between multiple variables, students either ignore some variables or combine them into a single-variable relationship. In this case, it appears that students are ignoring the heat transfer associated with the state-change. Alternatively, students may not be ignoring heat transfer during change of state, but may simply be unaware that there is heat transfer during change of state. In this case, choosing option C indicates a correct knowledge of heat, temperature, and energy during the cooling, but a gap in knowledge of phase changes. Thus, *filling* the gap of this concept through instruction can bring them to the correct answer.

In addition to option C, female students were also more likely to choose option E compared to males in the pre-instruction test. This is a common and 'common sense' misconception and was also found by Yeo and Zadnik (2001). This view of matter assumes that a material can be in only one state at a time, rather than an equilibrium of two states - in this case water and ice - at a temperature below the phase transition. It also assumes constant (normal) atmospheric pressure. It is not clear why female students would be more likely to hold this misconception. However, it is consistent with typical high school explanations of state changes of water as happening at a constant temperature, with freezing happening at a constant 0°C.

The results of question 4, that male students have higher facility than females both before and after instruction, indicate that they have better conceptual understanding of heat, temperature and energy during cooling and state changes. In addition, the NLGs for males were slightly higher than females implying that males improved more than females due to instruction.

However, the use of transition matrices and a hierarchy of answer choices reveals that females were *more* likely than males to improve their conceptual understanding when we include in our analysis changes from an incorrect answer (A, D or E) to a better incorrect answer (C). Overall, 33.4% of males were found to have improved conceptual understanding, compared to 43.4% of females.

The analysis using facilities and NLGs can be used to determine the effectiveness of the teaching and learning for the concept being examined. However, an analysis based only on the proportion of correct answers does not provide a complete picture of the learning gains. Where there is a hierarchy of answers, but a binary marking system is used, then information on the learning is being lost. In the typical binary system marks are given for a correct answer *only* and no marks are given for *any* incorrect answer, so there is no distinction made between answers which indicate different levels of understanding other than “wrong” and “right”. The binary marking system does not capture gains which result in students moving from an incorrect answer to an answer which is still incorrect but is a better incorrect answer, i.e. a move upwards in the answer hierarchy. This is the value of transition matrices. They not only provide information about the proportion of a cohort giving correct answers, but show in details how students’ answers change with instruction which gives information on the progress of their understanding.

The loss of information associated with using a simple binary marking system has two significant drawbacks. First, the instructor does not receive adequate feedback on their teaching effectiveness, and may believe that they have not achieved valuable learning outcomes when in fact many students have improved their conceptual understanding. Second, students do not receive the best possible feedback on their learning (or reward in terms of marks for their learning) – they may have an improvement in understanding but not an improvement in scores. This may discourage them, and may also make them doubt their new learning.

In this study, transition matrices allowed us to see that both males and females benefitted from the instruction. The gender gap in conceptual understanding pre-instruction has been reduced, although this is not reflected in the calculated gaps in facility. This information would not have been available to the instructor had only the facilities and NLGs been considered. Therefore, it is important to explore the students’ answer choices and how they change, to determine whether a cohort (or sub-cohorts) benefitted (or benefitted equally) from instructions.

Acknowledgements

The corresponding author acknowledges the financial support of a UNSW Canberra Ph.D. scholarship, and Dr. David Low and Dr. Zlatko Jovanoski for their supervision.

References

- Chi, M. T. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S.Vosniadou (Ed.), *Handbook of research on conceptual change*, (pp. 61-82): Hillsdale, NJ: Erlbaum.
- Docktor, J., & Heller, K. (2008). Gender differences in both force concept inventory and introductory physics performance. *AIP Conference Proceedings*, 1064(1), 15-18. doi:10.1063/1.3021243
- Driver, R. (1989). Students’ conceptions and the learning of science. *International Journal of Science Education*, 11(5), 481-490. doi:10.1080/0950069890110501
- Georgiou, H., & Sharma, M. D. (2010). A report on a preliminary diagnostic for identifying thermal physics conceptions of tertiary students. *International Journal of Innovation in Science and Mathematics Education*, 18(2), 32-51. Retrieved from <https://www.scopus.com/inward/record.uri?eid=2-s2.0->

78650828011&partnerID=40&md5=859787084e224bc7b73528f57c564dbb

- Gilbert, J. K., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education*, 66(4), 623-633. doi:10.1002/sce.3730660412
- 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. doi:10.1119/1.18809
- Hazel, E., Logan, P., & Gallagher, P. (1997). Equitable assessment of students in physics: Importance of gender and language background. *International Journal of Science Education*, 19(4), 381-392. doi:10.1080/0950069970190402
- Harrison, A. G., Grayson, D. J., & Treagust, D. F. (1999). Investigating a grade 11 student's evolving conceptions of heat and temperature. *Journal of Research in Science Teaching*, 36(1), 55-87. doi:10.1002/(SICI)1098-2736(199901)36:1<55::AID-TEA5>3.0.CO;2-P
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141-158. doi:10.1119/1.2343497
- Lindstrøm, C., & Sharma, M. D. (2011). Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter? *International Journal of Innovation in Science and Mathematics Education*, 19(2), 1-19. Retrieved from <https://openjournals.library.sydney.edu.au/index.php/CAL/article/view/4770>
- McDermott, L. C., & Shaffer, P. S. (2002). *The Physics Education Group at the University of Washington, Tutorials in introductory Physics*: Prentice Hall, Upper Saddle River, NJ.
- Morris, G. A., Walter, P., Skees, S., & Schwartz, S. (2017). Transition matrices: A tool to assess student learning and improve instruction. *The Physics Teacher*, 55(3), 166-169. doi:10.1119/1.4976661
- Nottis, K. E., Prince, M. J., & Vigeant, M. A. (2017). Undergraduate engineering students' understanding of heat, temperature, and energy: An examination by gender and major. *US-China Education Review*, 7(3), 125-143. doi:10.17265/2161-623x/2017.03.001
- Paik, S. H., Cho, B. K., & Go, Y. M. (2007). Korean 4-to 11-year-old student conceptions of heat and temperature. *Journal of Research in Science Teaching*, 44(2), 284-302. doi:10.1002/tea.20174
- Prince, M., Vigeant, M., & Nottis, K. (2012). Development of the heat and energy concept inventory: Preliminary results on the prevalence and persistence of engineering students' misconceptions. *Journal of Engineering Education*, 101(3), 412-438. doi:10.1002/j.2168-9830.2012.tb00056.x
- Rozier, S., & Viennot, L. (1991). Students' reasonings in thermodynamics. *International Journal of Science Education*, 13(2), 159-170. doi:10.1080/0950069910130203
- Treagust, D. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice items. *Research in Science Education*, 16(1), 199-207. doi:10.1007/bf02356835
- Wattanakasiwich, P., Taleab, P., Sharma, M. D., & Johnston, I. D. (2013). Construction and implementation of a conceptual survey in thermodynamics. *International Journal of Innovation in Science and Mathematics Education*, 21(1), 29-53.
- Yeo, S., & Zadnik, M. (2001). Introductory thermal concept evaluation: Assessing students' understanding. *The Physics Teacher*, 39(8), 496-504. doi:10.1119/1.1424603