DEVELOPING AND ASSESSING RESEARCH-BASED TOOLS FOR TEACHING QUANTUM MECHANICS AND THERMODYNAMICS

by

Benjamin R. Brown

B.S. Physics, Carnegie Mellon University, 2001

M.S. Physics, University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Benjamin R. Brown

It was defended on

July 7, 2015

and approved by

Chandralekha Singh, Dietrich School of Arts and Science, Physics & Astronomy

Larry J. Shuman, Swanson School of Engineering

Robert P. Devaty, Dietrich School of Arts and Science, Physics & Astronomy

Adam K. Leibovich, Dietrich School of Arts and Science, Physics & Astronomy

Russell J. Clark, Dietrich School of Arts and Science, Physics & Astronomy

Dissertation Director: Chandralekha Singh, Dietrich School of Arts and Science, Physics & Astronomy

DEVELOPING AND ASSESSING RESEARCH-BASED TOOLS FOR TEACHING QUANTUM MECHANICS AND THERMODYNAMICS

Benjamin R. Brown, PhD

University of Pittsburgh, 2015

Research-based tools to educate college students in physics courses from introductory level to graduate level are essential for helping students with a diverse set of goals and backgrounds learn physics. This thesis explores issues related to student common difficulties with some topics in undergraduate quantum mechanics and thermodynamics courses. Student difficulties in learning quantum mechanics and thermodynamics are investigated by administering written tests and surveys to many classes and conducting individual interviews with a subset of students outside the class to unpack the cognitive mechanisms of the difficulties. The quantum mechanics research also focuses on using the research on student difficulties for the development and evaluation of a Quantum Interactive Learning Tutorial (QuILT) to help students learn about the time-dependence of expectation values using the context of Larmor precession of spin and evaluating the role of asking students to self-diagnose their mistakes on midterm examination on their performance on subsequent problem solving. The QuILT on Larmor precession of spin has both paper-pencil activities and a simulation component to help students learn these foundational issues in quantum mechanics. Preliminary evaluations suggest that the QuILT, which strives to help students build a robust knowledge structure of time-dependence of expectation values in quantum mechanics using a guided approach, is successful in helping students learn these topics in the junior-senior level quantum mechanics courses. The technique to help upper-level students in quantum mechanics courses effectively engage in the process of learning from their mistakes is also found to be effective. In particular, research shows that the self-diagnosis activity in upper-level quan-

tum mechanics significantly helps students who are struggling and this activity can reduce the gap between the high and low achieving students on subsequent problem solving. Finally, a survey of Thermodynamic Processes and the First and Second Laws (STPFaSL) is developed and validated with the purpose of evaluating the effectiveness of these topics in a thermodynamics curriculum. The validity and reliability of this survey are discussed and the student difficulties with these topics among various groups from introductory students to physics graduate students are cataloged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

ACKNOWLEDGMENTS

The development and validation of STPFaSL was aided greatly through correspondence with Fred Reif, David Meltzer, Michael Loverude, Jessica Clark, Martin Vincent, and the input of many physics faculty at the University of Pittsburgh. R. P. Devaty was especially generous with both his attention and expertise. I am also grateful to E. Yerushalmi for useful discussions. I also thank the NSF for their support.

Also, I thank my advisor Chandralekha Singh, for giving me the opportunity to do this work, for being an attentive and conscientious mentor, and for being an exemplary role model.

Finally, I thank my parents Janet Brown and Raymond Brown, for introducing me to the universe, and supporting me throughout my seemingly perpetual education.

1.0 INTRODUCTION

Physics education research (PER) seeks to apply the scientific method to design and evaluate curricula to help students learn physics [99]. The design process necessarily requires pedagogical knowledge of student difficulties, and makes use of the fruits of cognitive science. To measure the effectiveness of a pedagogical intervention, actual measurements must be taken. *Student* performance must be evaluated (both before and after instruction) to establish the effectiveness of a particular curriculum or instructional approach.

First, cognitive science tells us a great deal about how humans learn, and the capacities and limitations of human cognition [100]. Second, in addition to knowledge of cognitive science, expertise in physics is essential for the design and evaluation of various pedagogies and curricula specific to the knowledge-rich domain of physics.

1.1 COGNITIVE SCIENCE

Physics courses are often touted as an excellent opportunity to conduct expertise research since depending upon a person's expertise, their ability to solve problems can differ significantly [101]. Cognitive science deals with modeling and measuring how people learn and demonstrates that how people organize their knowledge while learning and how they connect new and prior knowledge has consequences for whether the knowledge will be retrieved to solve problems later in situations in which the knowledge is applicable. There are several versions of learning theories which mainly differ in their focus on how they model learning [100].

1.1.1 Learning Theories

Learning theories which focus on how people learn can provide guidance on how to help students learn better.

1.1.2 Piaget's Optimal Mismatch Model

Piaget [102] introduced a model of learning in which he emphasized the importance of the internal mental state of "cognitive conflict." A student, when presented with an idea or observation that conflicts with his/her current understanding of some topic, is put into an uncomfortable internal state. Piaget's "optimal mismatch" model recommends that to help a student learn well, his/her current knowledge state must first be known. Then, to elicit productive cognitive conflict, the students must encounter a situation in which they will recognize that their observations are inconsistent with their prior conceptions. If the task is carefully designed, the cognitive conflict will end when students assimilate and accommodate the new and prior knowledge and develop a coherent understanding. Thus, an instructor should provide systematic tasks commensurate with a student's current knowledge to help them resolve inconsistencies and build a robust knowledge structure. Piaget is also known for his stages of cognitive development as shown in Table 1. The concrete operational and formal operational stages are particularly important for physics teaching and learning even at the college level because research shows that many students in college may not have transitioned from the concrete operational to formal operational stage and need concrete experiences to learn physics.

1.1.3 Zone of Proximal Development

An early contemporary of Piaget, Vygotsky, also introduced a learning theory. In Vygotsky's theory, one important concept is the "Zone of Proximal Development," or ZPD [103]. Vygotsky's ZPD is a zone described by what a student can do on his or her own vs. with the help of an instructor who is aware of his current knowledge and targets instruction above that current level to enable learning. The ZPD is defined as knowledge that is not too "close"

Table 1: Piaget's Stages of Cognitive Development. Piaget is most well known for characterizing the progression from infancy to adulthood of human cognition as a progression through four stages. This table presents the stages as described, their nominal ages, and a description of each [102].

Stage	**Age**	**Description**
Sensorimotor	0–2	Reflex Base, Coordinate Reflexes
Preoperational	2–6 or 7	Self-Oriented, Egocentric
Conrete Operational	6 or 7 – 11 or 12	More Than One View Point, No Abstract Problems
Formal Operational	11 or 12 +	Think Abstractly, Reason Theoretically, *Not all people reach this stage.*

to what the student knows, nor too 'far,' but in a zone near to the student's current understanding. Thus, learning is effective if the instructor is aware of the current knowledge state of a student, and uses an instructional design which is in the ZPD. To facilitate instruction which is in the ZPD, an instructor must provide scaffolding support (guidance and feedback with a focus on developing self-reliance) to stretch student learning, while taking advantage of a student's strengths. The implications of Vygotsky's ZPD model for my research are similar to that for Piaget's optimal mismatch model.

1.1.4 Preparation for Future Learning

Another theory of learning, advanced by Bransford and Schwartz [1], focuses on "Preparation for Future Learning (PFL)." One interpretation of this theory is that any instructional approach should be considered along two dimensions: innovation and efficiency. Any instructional method can be evaluated along these two axes. For example, if a teaching method is too innovative (see Figure 1), a student can become frustrated with struggling, and disen-

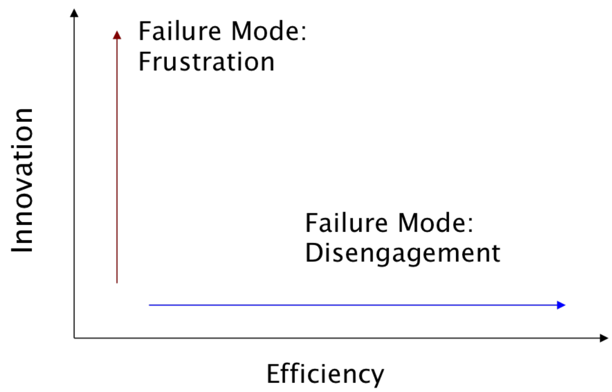


Figure 1: Innovation vs Efficiency, see [1]

gage with the learning process. If a method is too efficient, in the sense that it is too easy to follow and requires only low cognitive engagement, a student will become bored, and again, disengage from the learning process.

It is important to note that what is innovative and efficient is student-dependent. What is efficient for an expert can be innovative, or too innovative for a student. PFL recommends that instructors use techniques that balance innovation and efficiency. Again, the implication to my research is that it is important to develop learning tools based upon the prior knowledge state of the students to help them learn effectively.

As a side note, it is important for instructors to remember that their students are not mini-versions of themselves when designing a curriculum. A student would typically not have the same tolerance for remaining in a state of confusion as an expert. Instructors, especially at the university level, should not expect to be representative of the student population.

1.1.5 Cognitive Architecture

While learning theories can guide the design of curricula, it is important to keep in mind the constraints on the cognitive architecture when developing research-based learning tools [2, 104, 105]. From a PER researcher's point of view, at a coarse-grained level, human cognitive architecture has two components: working memory, and long-term memory.

1.1.5.1 Working Memory

The working memory, sometimes called short-term memory, consists of 7 ± 2 "slots"; each slot can hold a "chunk" of information [2]. The scarcity of slots is one of the main bottlenecks in learning and problem-solving. "Cognitive overload" involves a situation in which the capacity of the working memory is exceeded and one cannot make progress in solving a problem due to the lack of available slots with which to work [109].

1.1.5.2 Long Term Memory

While working memory is the immediate "working space" where information is processed, long-term memory is where knowledge is stored. The working memory gets information from long term memory and also from the sensory buffers. In a problem solving situation, each piece of information comes at the cost of occupying one of the precious few slots of working memory. Recall from long-term memory is cued by working memory and the struggle to make progress on a cognitive task. Long-term memory can be "primed" ahead of time, reducing the barrier to the triggering of certain memories, by cues in the environment, previous cognitive tasks, etc.

1.1.6 Chunking and Knowledge Structures

The number of slots available in the working memory is limited, but what can actually go *into* one of these slots is not limited. Current understanding is that the size of the chunk in a given domain increases with expertise [104]. Examples of a chunk include a piece of information, or a network of pieces of information which can be treated as a whole. A chunk can even consist of a network of other chunks... and so on. There is no meaningful upper bound to the amount of information that can ultimately be hierarchically nested within a single chunk.

Experts benefit greatly from past experience. One concrete way in which chunking takes place is that experts have larger, better organized knowledge structures in their domain of expertise so their knowledge chunks are larger. It is actually less cognitively demanding for an expert to solve the same problem compared to a novice due to the chunking of information that an expert does. This type of chunking has consequences for "Categorization and representation of physics problems by experts and novices"[3] and in how experts and novices reproduce good and bad boards of chess [105].

1.2 KNOWLEDGE STRUCTURES OF EXPERTS AND NOVICES

Research has been carried out with the goal of inferring the knowledge structures of physics experts and novices. The resulting knowledge structure of an expert is shown schematically in Figure 2. This structure was inferred from asking novices and experts to categorize a collection of textbook physics problems based upon how they would solve them. The investigators also used think-aloud interviews [106] to establish what subjects were thinking as they carried out the task.

In contrast, a novice's knowledge structure is shown in Figure 3. Note that a novice's structure is organized by superficial features, rather than the physics principles at the core of a physics problem.

1.3 PER RESEARCH FOCUSING ON CONTENT DEVELOPMENT

PER can take many forms. One type of research focuses on student difficulties while another type may evaluate research-based curricula and pedagogies in classes at various levels. Yet another type of PER research may focus on developing and evaluating new educational technologies for use in the physics classroom. This thesis concerns measuring the efficacy of different approaches on student learning and quantifying the relative prevalence of student difficulties with specific physics content in quantum mechanics and thermodynamics at the undergraduate (and even graduate) level.

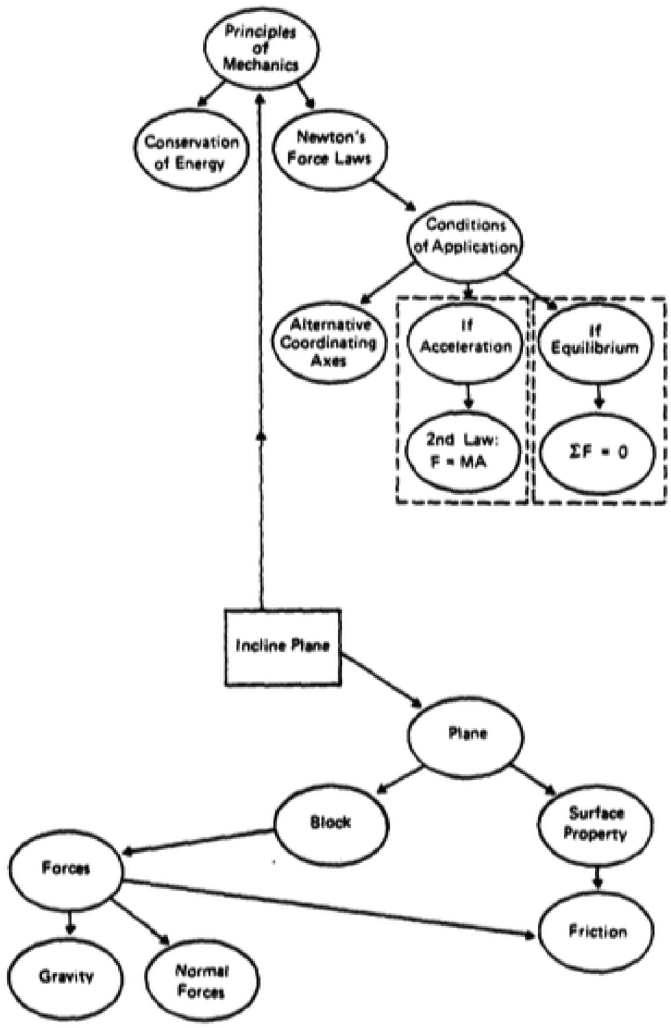


Figure 2: Expert Knowledge Structure(Figure 5 of ref[3]); Note the heirarchical organization, with high-level principles on top, and increasingly detailed concepts as one moves towards the bottom.

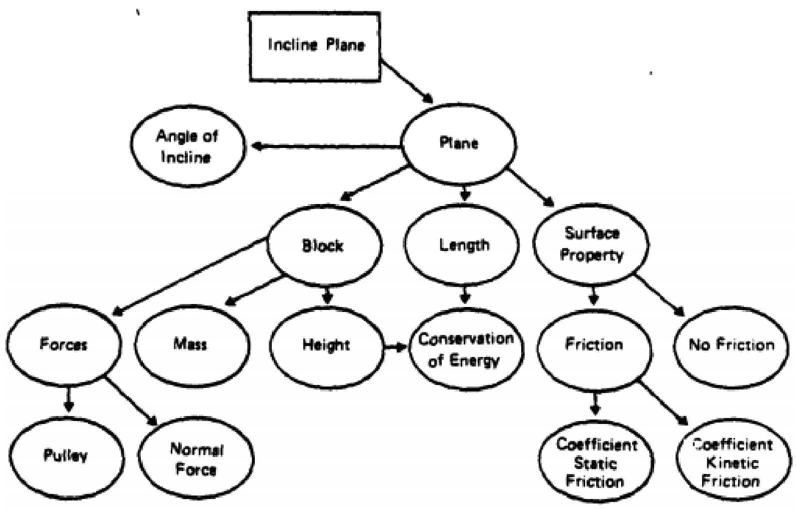


Figure 3: Novice Knowledge Structure (Figure 4 of ref [3]). Superficial features, rather than physics principles, are the organizing principle for the inferred knowledge structure of a novice. Here, for instance, we see a "Plane" knowledge structure, not a "Newton's Law's" knowledge structure. As we move downward, principles of physics are invoked, each as a special case to deal with the problem at hand.

1.3.1 Design and Evaluation of a Conceptual Survey

While developing and validating a conceptual survey of thermodynamic processes and first and second laws (STPFaSL), it was useful to think about it in the context of other standardized tests for university-level physics.

1.3.2 Electromagnetism

Two examples of surveys of electromagnetism are the Brief Electricity and Magnetism Assessment (29 items), or BEMA [4], and the Conceptual Survey of Electricity and (32 items), or CSEM [5]. Not only are the scores on the thermodynamic survey we developed and validated similar to BEMA and CSEM, our point biserials are also similar to BEMA's. Similar to the development process of these prior surveys, our development and validation process involved both experts and students. In addition, statistical validation was conducted to establish the overall usefulness of our survey similar to these surveys.

A review of difficulties that students have with upper-level electrodynamics [6] shows that upper-level undergraduates struggle not only with concepts in electrodynamics, but also have great difficulty with the mathematics, and in connecting the mathematics with their conceptual understanding of the topics [6, 7].

1.3.3 Classical Mechanics

At the introductory level, several studies (see, for example, research conducted by the University of Washington PER group [8]) have focused on difficulties with students' conceptual understanding of classical mechanics. Students' understanding of classical mechanics may have an influence on their learning of quantum mechanics[9].

1.3.4 Quantum Mechanics

Student difficulties in quantum mechanics in one dimension are investigated [10] by developing and validating a tool similar to the one discussed in this thesis for thermodynamics. Moreover, difficulties related to the time dependence of expectation values [11] connect well with the investigations related to the development and evaluation of a quantum interactive learning tutorial on the Larmor precession of spin discussed in this thesis.

1.4 DESIGNING AND EVALUATING RESEARCHED-BASED TOOLS FOR HELPING STUDENTS LEARN PHYSICS

PER is concerned with helping students learn physics at all levels; a great deal of research has been conducted at the post-secondary introductory level. From a researcher's point of view, this population is ideal in that it is the largest slice one could take of the enrollment pie in the physics courses. A large dataset can be generated in a single semester, at a single (large) university and useful inferences can be made.

In physics courses, instructors often start with a heterogeneous "student" population. The focus of research-based curricula and pedagogies is typically to help all students. Such curricula and pedagogies tend to reduce the spread among students. A stark example of this reduction is the result of applying the technique described and demonstrated in chapter 3.

The work presented in this thesis is especially geared towards advanced undergraduates, and the focus is on physics content in quantum mechanics and thermodynamics. The design and evaluation of a tutorial to help students better understand the time dependence of expectation values in quantum mechanics is the first chapter, 2. We develop and evaluate a tutorial on Larmor Precession of spin, a quantum phenomena somewhat analogous to the classical motion of a spinning top whose axis of rotation "precesses" as it spins on a floor. We then evaluate students' reasoning for correctness before and after students learn from the tutorial and also investigate whether students are more likely to adopt a quantum mechanical reasoning after working on the tutorial.

Secondly, in chapter 3, we present a simple, but effective method for helping students to learn from their own mistakes. Students (mostly physics majors) studying quantum mechanics are given the opportunity to earn back points on midterm questions, and their performance on the final exam is measured. Students given the explicit grade incentive to examine and correct their own mistakes are shown to perform markedly better on the same material when it appears on the final exam; students who were not given this incentive do not perform nearly as well.

The design of a conceptual thermodynamics survey cuts across many levels, from introductory to graduate levels, showing a remarkable consistency of reasoning amongst students at all levels. While the content of the survey is targeted at an introductory level (see chapter 4), we find that the students at all levels have great difficulty with the actual *concepts* of thermodynamics covered in the survey(see chapter 5).

2.0 IMPROVING STUDENTS' UNDERSTANDING OF TIME-DEPENDENCE OF EXPECTATION VALUES VIA LARMOR PRECESSION OF SPIN

2.1 INTRODUCTION

Quantum mechanics (QM) is a particularly challenging subject for undergraduate students. Based upon the research studies that have identified difficulties [11–15] and findings of cognitive research, we have developed a set of research-based learning tools to help students develop a good grasp of QM. These learning tools include the Quantum Interactive Learning Tutorials (QuILTs) [16, 17].

Here, we discuss the development and evaluation of a QuILT employing Larmor precession of spin as the physical system to help upper-level undergraduate students in quantum mechanics courses learn about time-dependence of expectation values. Through a guided approach to learning, the QuILT helps students learn about these issues using a simple two state system.

Generally, the expectation value of some observable Q evolves in time because the state of the system evolves in time in the Schrödinger formalism. The expectation value of an observable Q in a given quantum state at time t can be written as

$$\langle Q(t)\rangle \equiv \langle\psi(t)|\hat{Q}|\psi(t)\rangle \tag{2.1}$$

If the operator $\hat{Q}$ has no explicit time dependence, taking the time derivative of Eqn. 2.1 and substituting Schrödinger's equation where appropriate yields Ehrenfest's theorem:

$$\frac{d}{dt}\langle Q(t)\rangle = \frac{\langle\psi(t)|[\hat{Q},\hat{H}]|\psi(t)\rangle}{\hbar i}, \tag{2.2}$$

Thus, the question of whether an expectation value evolves with time is equivalent to asking whether or not the right hand side of Eqn. 2.2 is zero.

There are two distinct general conditions when the expectation value is time-independent, i.e., its time derivative is zero. One case is when the commutator of operators $\hat{Q}$ and $\hat{H}$, $[\hat{Q}, \hat{H}]$, is zero. In this case, the operator $\hat{Q}$ commutes with the Hamiltonian. An operator which commutes with $\hat{H}$ corresponds to a conserved quantity. The other case that yields a time-independent expectation value is when the state vector $|\psi(t)\rangle$ is a stationary state. A stationary state is an eigenstate of the Hamiltonian operator. All observables (that correspond to operators with no explicit time dependence) have time independent expectation values in this case. The physical system used to develop and assess the effectiveness of a Quantum Interactive Learning Tutorial (QuILT) on time-dependence of expectation values is Larmor precession of a spin-$^1/_2$ system (with zero orbital angular momentum).

2.2 LARMOR PRECESSION OF SPIN

A particle with a magnetic moment will exhibit Larmor precession if an external magnetic field is applied. The system is analogous to the classical scenario of a rigid rotator (such as a child's spinning top) in the presence of a gravity field in which the precession of the angular momentum vector can be observed.

Working on the QuILT, students learn that a particle with a magnetic moment may exhibit Larmor precession if an external magnetic field is applied. For pedagogical purposes, the simplest case of a spin $^1/_2$ system, which is a two-state quantum system, is chosen in the QuILT to help students learn this challenging topic. In this case, spin operators in any basis are 2×2 matrices, which are least likely to cause cognitive overload to students. Despite the simplicity of the system, students learn to reason that the knowledge and skills they acquire in this context apply to a broad range of quantum systems.

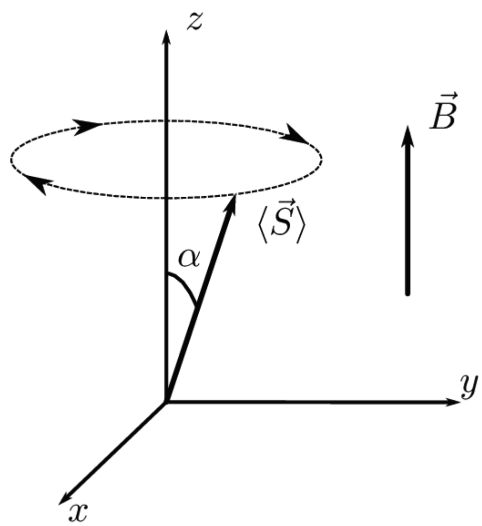


Figure 4: A magnetic moment, pointing in the direction of $\hat{\mathbf{S}}$, precesses in the presence of a magnetic field $\mathbf{B}$ directed upwards.

2.3 EXPLORING STUDENT DIFFICULTIES

Before the development of the QuILT, we investigated student difficulties with time dependence of expectation values using open-ended surveys and multiple-choice questions in order to address them in the QuILT. Here we summarize the findings from a written survey given to 39 students:

Difficulty with the relevance of the commutator of the operator corresponding to an observable and the Hamiltonian: A consequence of Ehrenfest's Theorem is that if an operator $\hat{Q}$ corresponding to an observable Q commutes with the Hamiltonian, the time derivative of $\langle Q\rangle$ is zero, regardless of the state. However, 46% of students did not realize that since the Hamiltonian governs the time-evolution of the system, any operator $\hat{Q}$ that commutes with it must be a constant of motion and its expectation value must be time-independent.

Difficulty recognizing the special properties of stationary states: If the magnetic field is along the z-axis, all expectation values are time independent if the initial state is an eigenstate of $\hat{S}_z$ because it is a stationary state. However, 51% of students surveyed incorrectly stated that $\langle S_x \rangle$ and $\langle S_y \rangle$ depend on time in this case. One common difficulty includes reasoning such as "since the system is not in an eigenstate of $\hat{S}_x$, the associated expectation value must be time dependent" even in a stationary state. Another very common difficulty is reasoning such as "since $\hat{S}_x$ does not commute with $\hat{H}$, its expectation value must depend on time" even in a stationary state.

Difficulty distinguishing between stationary states and eigenstates of operators corresponding to observables other than energy: Any operator corresponding to an observable has an associated set of eigenstates, but only eigenstates of the Hamiltonian are stationary states because the Hamiltonian plays a central role in time-evolution of the state. However, many students were unable to differentiate between these concepts. For example, for Larmor precession with the magnetic field in the z direction, 49% of students claimed that if a system is initially in an eigenstate of $\hat{S}_x$ or $\hat{S}_y$, the system will remain in that eigenstate. A related common difficulty is exemplified by the following comment from a student: "if a system is initially in an eigenstate of $\hat{S}_x$, then only the expectation value of S_x will not depend on time."

2.4 DEVELOPMENT OF A QUILT

Based upon a theoretical task analysis of the underlying knowledge from an expert perspective and common student difficulties found via research, a preliminary version of the QuILT with a sequence of guided questions along with pretests and posttests was developed. The pretest is to be administered after traditional instruction on Larmor precession and the posttest after students work on the QuILT. The QuILT strives to help students build on their prior knowledge and develop a robust knowledge structure related to the time-dependence of expectation values. Students have to actively think through the answers to each question in the guided approach to learning and the latter questions build on the preceding questions. The questions are designed such that the alternative choices deal with the common difficulties students have. The QuILT presents computer simulations [18, 19], adapted to represent Larmor precession, in which students can manipulate the Larmor precession setup to predict and observe what happens to time-dependence of expectation values for different observables, initial quantum states, and orientations of the magnetic field. In particular, students have to make predictions about whether the expectation values depend on time in various situations and then they engage via simulations of Larmor precession in learning whether their predictions are consistent with the observation. Then, students are provided with guidance and scaffolding support to help them reconcile the differences between their predictions and observations and to help them assimilate and accommodate relevant concepts. The components of the simulations students use in the QuILT to check their predictions and reconcile the differences include tools to orient the magnetic field in a particular direction, tools for state preparation, quantum measurement, and particle detection.

Different versions of the QuILT were iterated with five physics faculty members (experts) several times to ensure that they agreed with the content and wording of the questions. We also administered it individually to fifteen graduate students and upper-level undergraduate students in semi-structured think-aloud interviews to validate the QuILT for clarity, student interpretation, appropriateness of the guided sequence and in order to address any emergent issues. Students were first asked to think aloud as they answered the questions to the best of their ability without being disturbed. Later, we probed them further and asked them for clarification of points they had not made clear. Since both undergraduate and graduate students exhibited the same difficulties, we will not differentiate between the two groups further. Overall, the interviews focused on ensuring that the guided approach was effective and the questions were unambiguously interpreted. Modifications were made based upon each feedback.

2.5 RESULTS: ASSESSING EFFECTIVENESS OF THE QUILT

Once we determined that the QuILT was effective in individual administration, it was administered in class. Over several years, students in four junior-senior level quantum mechanics courses, totaling 95 students, were first given the pretest after traditional instruction. They then worked through the QuILT in class and were asked to complete whatever they could not finish in class as homework. Then, the posttest, which has analogous questions to the pretest (except, e.g., S_x and S_y are swapped in various questions) was administered in the following class. One class with 18 students (included in the 95 students) was also given four matched questions on their final exam on time-dependence of expectation values to investigate retention of these concepts more than two months later. Most data was collected by other physics education researchers in the Pitt PER group.

2.5.1 Reasoning-based Rubric

To assess student learning, both the pretest and posttest were scored based on two rubrics developed by the two researchers jointly. The "Strict" rubric only gives credit if the student provides a completely correct answer. "Completely correct" answer is defined as an answer in which both the correct answer is given and it is supported with correct reasoning. Correct reasoning can be either quantum mechanical or classical in nature (discussed later). Good performance on this strict scoring is likely to reflect an expert-like performance.

A second rubric, referred to as "Partial" was developed to give partial credit in scoring students' pretest and posttest performances. This rubric gives half credit for a correct answer (regardless of the reasoning) and half credit for correct reasoning (again, either correct quantum mechanical or classical reasoning is considered valid). This rubric is consistent with the common traditional methods of grading student performance.

More than 20% of the sample was tested for inter-rater reliability (scores of the two raters agree to better than 90%). For assessing retention of the concepts related to the time-dependence of expectation values, the same rubric is applied to the matched questions on the final exam.

2.5.2 Pretest vs Posttest Scores

Table 2 shows the average scores on each question on the pretest and posttest using both the strict grading scheme and partial grading scheme. The average pre-/posttest performances

Table 2: The average pre-/posttest scores of 95 students on each of the six items. For item 6, students were asked if there was precession in a given case and asked for an example in which precession does not take place. For this question, there is no "Partial" grading scheme.

	Strict			Partial		
Item	Pre (%)	Post (%)	p-value	Pre (%)	Post (%)	p-value
1	53.8	80.4	<0.001	57.1	81.4	<0.001
2	53.8	82.5	<0.001	64.8	87.1	<0.001
3	44.0	70.1	<0.001	49.5	78.4	<0.001
4	41.8	67.0	<0.001	47.8	72.7	<0.001
5	20.9	58.8	<0.001	29.1	62.4	<0.001
6	40.4	72.6	<0.001	*	*	*
Combined	42.5	71.9	<0.001	49.7	76.4	<0.001

were compared using a t-test; all improvements (using both strict and partial rubrics for each question) were statistically significant. Table 2 shows that the pretest scores were lower under the strict grading scheme than under the partial grading scheme. We note that although some students struggled on the posttest, the average posttest performance is superior to the baseline data from a group of 18 first year physics graduate students, indicating that time-dependence of expectation values is a challenging concept.

Items 1, 2, 3, and 5 on the pre-/posttests all ask questions of the following form for a given system: Given a specific initial state, does the expectation value of a certain observable Q change with time? Students were asked to explain their reasoning in each case. These questions are equivalent to asking whether or not the right hand side of Eqn. 2.2 is zero in that case. Items 3 and 5 are specifically designed to probe the knowledge of the two distinct general conditions when the expectation value is time-independent (when $\hat{Q}$ commutes with the Hamiltonian $\hat{H}$ or when the initial state is a stationary state). On Item 6, students were given a system initially in an eigenstate of $\hat{S_x}$ and asked whether precession occurs in this case (it does) and why, which prompts students to connect classical and quantum mechanical explanations.

Below, we discuss how student difficulties were reduced after working on the QuILT.

2.5.3 Items 1 and 2

Items 1 and 2 asked the student to consider a system initially in an eigenstate of $\hat{S}_x$, then asked whether $\langle S_x \rangle$ and $\langle S_y \rangle$ depended on time. The time evolution of each is a sinusoidal oscillation of amplitude $\frac{\hbar}{2}$, with $\langle S_x \rangle$ leading $\langle S_y \rangle$ by a phase of 90°. Performance on these two items was similar. Using the strict rubric, the scores improved from 54% to 80%, and 54% to 83% for items 1 and 2, respectively. One student states:

> "There's no ... there's no oscillation going on here you just have a pure S_x state."

Similarly for Item 2:

> "I'm trying to think this through, um ... So the only difference is... you've got the down states are imaginary in this case, and that's just... that'll go away once you take the expectation value, so I'm gonna say 'no', it's not gonna depend on time for the state... It's

> just these expectation values... we're gonna get what we're gonna get, as long as it's in that eigenstate."

While scores for these two items are approximately equal, with approximately equal improvements, the reasoning provided in the case of incorrect answers were different. In the case of 1, in which the measurement is made in the same eigenbasis as the measurement, 53% of students who responded "time independent" stated that the system was in a stationary state. However, when asked about the time dependence of $\langle S_y \rangle$, only 12% of students state that the system was in a stationary state. Instead, 72% of these students responded with reasoning based on the non-commutation of $\hat{S_x}$ and $\hat{S_y}$. For instance, a student responded "[since] $[Sx, Sy] \neq 0, [Sy]$ depends on time".

Another student states "No. Because it is then in an eigenstate of $\hat{S}_x$. $\hat{S}_x$, $\hat{S}_y$ are mutually incompatible observables."

These findings have led us to characterize a substantial fraction of students' understanding as follows: A stationary state is not a property of the system, but states are stationary *with respect to* a particular observable in question.

2.5.4 Item 3

Item 3 asked the student to consider the same system, initially prepared in an eigenstate of $\hat{S}_x$. The student is asked whether $\langle S_z \rangle$ depends on time. This quantity is constant with time. A response that cited and applied Ehrenfest's theorem, noting that $\hat{S_z}$ commutes with the Hamiltonian, was given a score of "1", and tabulated among answers with correct quantum reasoning. A response that $\hat{S_z}$ is constant due to precession not affecting the z component was given a score of "1", and tabulated among answers with correct classical reasoning. Under the strict rubric, this score improved from 44% to 70%.

2.5.5 Item 4

Item 4 is designed to assess student knowledge of the time evolution of the quantum state itself. The student is asked to consider a system which is prepared in a non-stationary eigenstate, and then state whether or not the system evolves away from this state, or remains in it until an external perturbation is applied. The alternate points of view are presented in conversation form. This item is designed to probe whether students suffer from over-generalization of the properties of a stationary state to any eigenstate. With the strict rubric, this score improved from 42% to 67%.

2.5.6 Item 5

Item 5 asked the student to consider instead a system initially in an eigenstate of $\hat{S}_z$, then asked whether $\langle S_x \rangle$ depends on time. In this case, no quantities are time dependent because the system is in a stationary state. Under the strict rubric, scores improved from 21% to 59%. Possible explanations of these atypically low scores are discussed later.

2.5.7 Item 6

Item 6 asked the student to consider again a system initially in an eigenstate of $\hat{S_x}$. The student is asked a pair of questions: First, whether precession occurs in this case (it does), and second, to provide an example in which precession does not occur. (A complimentary question is asked in the case that the student responds negatively to the first question). This question is designed to probe overall student knowledge of Larmor precession. Under the strict rubric, scores improved from 40% to 73%.

2.5.8 Retention of Learning

In addition to pre- and posttests, matched questions were administered as part of the final exam for 18 students in the junior-senior level quantum mechanics course in 2013.

Table 3 indicates that students' average score on questions on the final exam which were very similar to the questions on the pretest and posttest related to the time-dependence of expectation values are comparable to the posttest result. It is encouraging that the performance on the posttest and final exam are comparable overall because competing effects such as degradation of memory vs consolidation over the intervening half semester affect the performance on final exam.

Table 3: Assessment of $N = 18$ students at the end of the course shows that students retain understanding of Larmor Precession.

	Strict	Partial
Pre Test	32%	42%
Post Test	78%	82%
With Final Exam	83%	86%

2.5.9 Classical and Quantum Reasoning

The nature of most questions posed to students about the time-dependence of expectation values in the context of Larmor precession allowed either a quantum mechanical or semi-classical explanation. This presents an opportunity to investigate and compare the rate at which students adopt and correctly use the new "Quantum" framework learned, versus their reliance primarily on "Semi-Classical" reasoning.

For all questions, a rubric was developed that assigned values to statements students made. Statements were categorized as either "Quantum" or "Semi-Classical". They were assigned a score of "+1" if that statement was correct reasoning relevant to the question. A score of "-1" was assigned to incorrect statements. A common example reflecting student difficulties would be "$\hat{S}_x$ does not commute with the Hamiltonian, so $\langle S_x \rangle$ must depend on time," in a situation in which the system is in a stationary state. The overwhelming majority of incorrect statements of this type reflected known student difficulties. Absence of a statement is assigned a score of "0." A student's response can independently be characterized along two dimensions, "Quantum" and "Semi-Classical" reasoning, each of which can take on a value of "-1", "0", or "+1". Change in rate of correct statement between two assessments can therefore range from "-2" to "+2", indicating extremal negative progress to extremal positive progress, respectively.

Table 4 presents the results of broad adoption of correct quantum mechanical reasoning

Table 4: Change in method of reasoning between pretest and posttest for $N = 95$ students. In each of the five questions that could be answered with either "Quantum" or "Semi-Classical" reasoning, the rate of correct statement is tabulated.

Question Number	Quantum	p-value	Classical	p-value
1	+0.53	<0.001	+0.01	0.861
2	+0.51	<0.001	+0.06	0.319
3	+0.28	<0.001	-0.01	0.870
4	+0.44	<0.001	-0.01	0.856
5	+0.49	<0.001	+0.08	0.266

in student responses between pretest and posttest. Semi-Classical reasoning statements evaluated in this manner show either no change, or small positive progress.

Table 5 presents the comparison of the reasoning on posttest right after working through the QuILT with reasoning on the final exam several months later.

Comparison of student reasoning approaches in Tables 4 and 5 shows significant gain in "quantum mechanical" reasoning and retention of these gains. For the 2013 class, in which retention was measured, almost no semi-classical reasoning was used in the final exam. Although semi-classical reasoning was fine in this context, the change from posttest to the final exam can be interpreted as students acquiring expert-like fluency with quantum mechanical reasoning.

The average scores on the final exam compared to the posttest right after the QuILT suggest retention of knowledge for the duration of the course. Degradation of memory over time is a well studied phenomenon. Since the course does not continue instruction specific to Larmor Precession, degradation of memory might be expected to drop the score with time. However, we find that the scores on the final exam are the same (or even slightly higher). We hypothesize that a competing effect, i.e., consolidation of knowledge specific

Table 5: Retention in method of reasoning between posttest and final exam on matched questions for 18 students.

Question number	Quantum	p-value	Classical	p-value
1	+0.06	0.777	-0.06	0.560
2	-0.06	0.738	-0.06	0.560
3	+0.06	0.684	0.00	1.000
5	+0.22	0.429	-0.06	0.715

to quantum mechanics is taking place during the entire semester which helps retain the knowledge about time-dependence of expectation values acquired earlier. This is reflected both in final exam performance and the fact that almost no "semi-classical" statements are being made in the final exam on these questions involving time-dependence of expectation values in the context of Larmor precession of spin. This fact may be interpreted as students acquiring more expert-like fluency in problem solving.

2.6 SUMMARY

Students demonstrate an average improvement of 27% on posttest vs pretest, and evidence from all semesters supports that these gains are on average robust and retained. Other published work [11] supports the difficulty that students have with the abstract nature of quantum mechanics.

Additionally, the interactive simulation portion of the QuILT has been better integrated with the paper-pencil portion of the QuILT by leveraging technical improvements to enforce notational consistency to minimize cognitive load associated with changing tasks.

It is likely that these gains can be improved further, e.g., if students are asked to turn in the tutorials themselves for credit. Therefore, to improve engagement and participation rate, students should be given a grade incentive to complete the tutorial.

2.7 THE PRETEST AND POSTTEST FOR LARMOR PRECESSION

Here we present version A of the pre- and posttest. Version B is identical, except that all instances of $\hat{S}_x$ and $\hat{S}_y$ have been swapped. Versions A and B were each used as either pre- or posttest. No statistically significant difference was found between performance on Version A versus Version B ($p = 0.67$ that these were drawn from the same distribution via t-test).

All of the following questions refer to this system:

An electron is in an external magnetic field B which is pointing in the z direction. The Hamiltonian for the electron spin is given by $\hat{H} = -\gamma B \hat{S}_z$ where γ is the gyromagnetic ratio and $\hat{S}_z$ is the z component of the spin angular momentum operator. *Notation:* $\hat{S}_z \left|\uparrow\right\rangle_z = \frac{\hbar}{2} \left|\uparrow\right\rangle_z$, and $\hat{S}_z \left|\downarrow\right\rangle_z = -\frac{\hbar}{2} \left|\downarrow\right\rangle_z$. *For reference, the eigenstates of* $\hat{S}_x$ *and* $\hat{S}_y$ *are given by:*

$\left|\uparrow\right\rangle_x = \frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle_z + \left|\downarrow\right\rangle_z)$, $\left|\downarrow\right\rangle_x = \frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle_z - \left|\downarrow\right\rangle_z)$

$\left|\uparrow\right\rangle_y = \frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle_z + i\left|\downarrow\right\rangle_z)$, $\left|\downarrow\right\rangle_y = \frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle_z - i\left|\downarrow\right\rangle_z)$

1. If the electron is initially in an eigenstate of $\hat{S}_x$, does the expectation value of $\hat{S}_x$ depend on time? Justify your answer.
2. If the electron is initially in an eigenstate of $\hat{S}_x$, does the expectation value of $\hat{S}_y$ depend on time? Justify your answer.
3. If the electron is initially in an eigenstate of $\hat{S}_x$, does the expectation value of $\hat{S}_z$ depend on time? Justify your answer.
4. Consider the following statements from Andy and Caroline when the electron is initially in an eigenstate of $\hat{S}_x$ (the x component of the spin angular momentum):

   Andy: The electron will NOT be in an eigenstate of $\hat{S}_x$ forever because the state will evolve in time.

   Caroline: I disagree. If a system is in an eigenstate of an operator corresponding to a physical observable, it stays in that state forever unless a perturbation is applied.

With whom do you agree? Explain.

a. Andy

b. Caroline

5. If the electron is initially in an eigenstate of $\hat{S}_z$, does the expectation value of $\hat{S}_x$ depend on time? Justify your answer.

6. If the electron is initially in an eigenstate of $\hat{S}_x$, is there any precession of $\langle\vec{S}\rangle$ about the z axis? If your answer is yes, explain why and give an example of a situation where there will be no precession of $\langle\vec{S}\rangle$ about the z axis. If your answer is that there is no precession for the given case, explain why.

3.0 AN EASY-TO-IMPLEMENT INTERVENTION CAN SUBSTANTIALLY REDUCE THE PERFORMANCE GAP BETWEEN LOW- AND HIGH-ACHIEVING PHYSICS STUDENTS

3.1 INTRODUCTION

Helping students learn to think like a scientist is an important goal of most courses for science and engineering majors at all levels [20–27]. Meeting this goal is also critical to prepare an additional one million STEM professionals in ten years according to the recommendations of the United States President's Council of Advisors on Science and Technology (PCAST) report. However, it may be beneficial to many students if they are provided incentives to help them learn to think like scientists.

Two characteristics of science experts are that they have learned how to learn and they use problem solving as an opportunity for learning. In particular, experts automatically reflect upon their mistakes in their problem solution in order to repair, extend and organize their knowledge structure. Unfortunately, for many students, problem solving is a missed learning opportunity. Without guidance, students often do not reflect upon the problem solving process after solving problems in order to learn from them nor do they make an effort to learn from their mistakes after the graded problems are returned to them. However, instruction can explicitly prompt students to learn from their mistakes by rewarding them for correcting their mistakes. This type of activity can also help them learn to make use of problem solving as a learning opportunity.

Prior research has focused on how introductory physics students differ from physics experts [3, 28–31] and strategies that may help introductory students learn to learn [32–36]. By comparison, few investigations have focused on the learning skills of advanced physics students, although some investigations have been carried out on the difficulties advanced students have with various advanced topics, e.g., quantum mechanics [12, 14, 15, 37, 38].

It is commonly assumed that most students who have made it through an entire undergraduate physics curriculum have not only learned a wide body of physics content but also have picked up the habits of mind and self-monitoring skills needed to build a robust knowledge structure [3]. Instructors take for granted that advanced physics students will learn from their own mistakes in problem solving without explicit prompting, especially if students are given access to clear solutions. It is implicitly assumed that, unlike introductory students, advanced students have become independent learners and they will take the time out to learn from their mistakes, even if the instructors do not reward them for fixing their mistakes, e.g., by explicitly asking them to turn in, for course credit, a summary of the mistakes they made and writing down how those mistakes can be corrected [39, 39–45].

However, such assumptions about advanced students' superior learning and self-monitoring skills have not been substantiated by research. Very little is known about whether a physics professor develops these skills in a continuous or discontinuous manner from the time they are introductory students. There may be some discontinuous "boosts" in this process for many students, e.g., when they become involved in graduate research or when they ultimately independently start teaching and researching. There is also no research data on the fraction of students who have gone through the "traditional" undergraduate or graduatephysics curriculum and have been unable to develop sufficient learning and self-monitoring skills, which are the hallmark of a physicist. These issues are particularly important considering the diversity amongst STEM majors has increased significantly.

Moreover, investigations in which advanced physics students are asked to perform tasks related to simple introductory physics content do not properly assess their learning and self-monitoring skills [3, 30]. Advanced students may possess a large amount of "compiled knowledge" about introductory physics and may not need to do much self-monitoring or learning while dealing with introductory problems. For example, when physics graduate students were asked to group together introductory physics problems based upon similarity of solution, their categorization was better than that of introductory physics students [3]. While such tasks may be used to compare the grasp that introductory and advanced students have of introductory physics content, tasks involving introductory level content do not provide much insight into advanced physics students' learning and self-monitoring skills.

The task of evaluating advanced physics students' learning and self-monitoring skills should involve advanced-level physics topics at the periphery of advanced students' own understanding. While tracking the same student's learning and self-monitoring skills longitudinally is an extremely difficult task, taking snapshots of advanced students' learning and self-monitoring skills can be very valuable.

In an earlier investigation, Mason and Singh [46] investigated the extent to which upper-level students in quantum mechanics learn from their mistakes. They administered four problems in the same semester twice, both on the midterm and final exams in an upper-level quantum mechanics course. The performance on the final exam shows that while some students performed equally well or improved compared to their performance on the midterm exam on a given question, a comparable number performed poorly both times or regressed (i.e., performed well on the midterm exam but performed poorly on the final exam). The wide distribution of students' performance on problems administered a second time points to the fact that many advanced students may not automatically exploit their mistakes as an opportunity for repairing, extending, and organizing their knowledge structure. Mason and Singh also conducted individual interviews with a subset of students to delve deeper into students' attitudes toward learning and the importance of organizing knowledge. They found that many students focused on selectively studying for the exams and did not necessarily look at the solutions provided by the instructor for the midterm exams to learn, partly because they did not expect those problems to be repeated on the final exam and/or found it painful to confront their mistakes.

Giving incentives to students for learning from their mistakes, e.g., by explicitly rewarding them for correcting their mistakes can be an excellent learning opportunity both for learning content and developing useful skills. Students may gain a new perspective on their own deficient solutions by asking themselves reflective questions while solving the problems correctly making use of the resources available to them.

Here, we discuss a study spanning several years in which advanced undergraduate physics students taking a quantum mechanics course were given the same four problems in both the midterm exam and final exam similar to the Mason and Singh study but approximately half of the students were given incentives to correct their mistakes in the midterm exam and could get back up to 50% of the points lost on each midterm exam problem. Most data was collected by other physics education researchers in the Pitt PER group. The solutions to the midterm exam problems were provided to all students but those who corrected their mistakes were provided the solution after they submitted their corrections to the instructor. The performance on the final exam on the same problems suggests that students who were given incentives to correct their mistakes significantly outperformed those who were not given an incentive. We find that the incentive to correct the mistakes on the midterm exam had the greatest impact on the final exam performance of students who did poorly on the midterm exam, which is very encouraging.

3.2 METHODOLOGY

Students were directed to work on their own, but were free to use any resources, homework, notes, and books to help them with this correction opportunity. Of course, students in either the comparison group or incentivized group were free to use these resources to study at any time prior to the final in-class exam. It is worth noting that these questions were in-class exam questions: short enough to be answered during the exam and similar to homework and quiz questions that students previously worked on.

Our study took place over four years (the data were collected in four separate but identical courses) in the first semester of a two-semester honors-level undergraduate quantum mechanics course sequence taught by the same physics instructor at the University of Pittsburgh. The honors-level course sequence is mandatory only for those students who want to obtain an honors degree in physics. It is often one of the last courses an undergraduate physics major takes. Most students in this course are physics or engineering physics majors in their senior year (but some are in their junior year and there are also a few first year physics graduate students, who typically did not take a full year quantum mechanics sequence as undergraduate students). The four years in which the data were collected in the course were not consecutive years because typically a physics instructor at that university (University of Pittsburgh) teaches a course for two consecutive years and then another instructor teachers it. Therefore, the data were collected from four classes that spanned a six year period.

The classes were primarily taught in a traditional lecture format but the instructor had the students work on some preliminary tutorials that were being developed. Students were assigned weekly homework throughout the fifteen-week semester. In addition, there were two midterm exams and a final exam. The homework, midterm and final exams were the same in different years. The midterm exams covered only limited topics and the final exam was comprehensive. Students had instruction in all relevant concepts before the exams, and homework was assigned each week from the material covered in that week. Each week, the instructor held an optional class in which students could ask for help about any relevant material in addition to holding office hours. The first midterm took place approximately eight weeks after the semester started, and the second midterm took place four weeks after the first midterm. For our study, two problems were selected from each of the midterm exams and were given again verbatim on the final exam along with other problems not asked earlier. The problems given twice are listed in Appendix B.1. In the second and fourth year in which this study was conducted, the data were collected from classes in which students were asked to self-diagnose their mistakes on both their midterm exams in the course and could earn a maximum of 50% of the points lost on each problem for submitting the corrected solution to each of the midterm exam problems. These classes formed the experimental or incentivized group. Including both years, there were 30 students in the incentivized group. There were no self-diagnosis activities in the first and third year in that students were not provided any grade incentive to diagnose their mistakes and submit the corrected solution to the midterm exam problems. The students in these two years formed the comparison group. There were 33 students in the comparison group including both years. In all of these four years, there were two midterm exams and a final exam and the same midterm and final exam questions were given. The instructor was the same and the textbook and homework assignments from the textbook were the same.

All students were provided the solution to each midterm examination on the course website. However, for the experimental group, the midterm solutions were posted on the course website after students self-diagnosed their mistakes. Moreover, written feedback was provided to all students in both the experimental and comparison groups after their midterm exam performance, indicating on the exams where mistakes were made and how they can be corrected. Students in the experimental group were asked to submit the corrected solution to each problem on the midterm exam on which they did not have a perfect score. Students were given four days to diagnose and correct their mistakes. They were asked to do the diagnosis of their mistakes and correct their mistakes individually but were allowed to use their books, notes or online material. The corrected solution submitted by most students after the self-diagnosis was almost perfect so it was easy for the instructor to reward students with 50% of the points lost. The final exam in the course was cumulative and in addition to other questions, two problems from each midterm exam were repeated. Our goal was to evaluate how students performed in subsequent problem solving based upon whether they diagnosed and corrected their mistakes on the midterm examination when provided with a grade incentive.

Three of the problems chosen (problem 1 which will also be called the expectation value problem for convenience, problem 2 or measurement problem and problem 3 or momentum problem in Appendix B.1) were those that several students had difficulty with; a fourth problem (problem 4 or harmonic oscillator problem) which most students found straightforward on one of the two midterm exams was also chosen. The most difficult of the four problems (based upon students' performance) was the momentum problem in Appendix B.1 that was also assigned as a homework problem before the midterm exam but was perceived by students to be more abstract in nature than the other problems. The easiest of the four problems, the harmonic oscillator problem, was an example that was solved within the assigned textbook. Thus, students in both the experimental and comparison groups had the opportunity to learn from their mistakes before they encountered the four problems selected from the midterm exams on their final exam (as noted earlier two problems were selected from each midterm exam).

A scoring rubric, developed jointly with E. Yerushalmi and E. Cohen to assess how well the students in introductory physics courses diagnose their mistakes when explicitly prompted to do so, was adapted to score students' performance on each of the four quantum mechanics problems on both the midterm and final exams. The scoring was checked independently by two scorers for at least 20% of the students and at least 80% agreement was found on the scoring for each student on each problem in each attempt (on midterm and final exams).

The scoring rubric has two sections: one section scores students on their physics performance and the other section scores how well they presented their solution. The rubric for the presentation part was somewhat different from the corresponding part for introductory physics because quantum mechanics problems often asked for more abstract answers (e.g., proving that certain energy eigenstates are equally probable) as opposed to finding a numerical answer. Therefore, some categories in the introductory physics rubric (e.g. writing down units) were omitted from the presentation part of the quantum mechanics rubric and other categories were adapted to reflect the nature of the quantum problems better (e.g., checking the answer was adapted to making a conceptual connection with the results).

Although the grading rubric allows us to assign scores to each student for performance on physics and presentation parts separately, these two scores are highly correlated with the regression coefficient between the two scores being R=0.98. The reason for this high correlation is that students' presentation of the problem depended upon whether or not they understood the physical content. If a student did not know the relevant physical concepts, he/she could not set up an appropriate problem solving strategy to score well on the presentation part. We therefore only focus on students' physics scores on each of the four questions given on the pretest and posttest.

3.3 RUBRICS AND SCORING

Appendix B.2 demonstrates the scoring rubric for physics for all the four problem along with the scores of some students on both the midterm and final exams. Below, we first describe the symbols used for scoring and then give an explanation of how a quantitative score is derived after the initial scoring is assigned symbolically for each sub-part of the rubric. The symbol "+" (worth 1 point) is assigned if a student correctly completes a task as defined by the criterion for a given row. The symbol "-" (worth 0 points) is assigned if the student either fails to do the given task or does it incorrectly. If a student is judged to have gotten something partially correct, then the rater may assign a combination of pluses and minuses (++/-, +/-, +/−) to reflect this behavior, with the understanding that such a combination represents an average score of pluses and minuses (e.g. ++/- translates to 2/3 of a point). If the student's solution does not address a criterion then "n/a" (not applicable) is assigned and the criterion is not considered for grading purposes at all. For example, if the student does not invoke a principle, the student will receive a "-" in the invoking row but will receive "n/a" for applying it in the apply row because the student cannot be expected to apply a principle that he/she did not invoke.

An overall or cumulative score is tabulated for each of the physics and presentation parts for each question. For the cumulative physics score, the average of the scores for each subcategory (e.g., invoking appropriate physics concepts, invoking inappropriate concepts and applying concepts) can be used in each column for each student on a given problem on the midterm or final exams. Similarly, the cumulative score for the presentation or problem solving part can be computed by averaging over the scores for each of the subcategories (organization, plan, and evaluation) in each column for each student on a given problem on the midterm or final exams.

3.4 RESULTS

We looked for correlation between the midterm exam score (which we call pretest score) and final exam score (which we call posttest score) on the four common problems for each group (comparison and incentivized groups). In particular, by comparing the performance of incentivized students with the comparison group, we examined the effect of giving grade incentives to correct one's mistakes on the pretests on subsequent performance on the posttest on four problems repeated from the pretest.

The comparison group and incentivized group had nearly identical average performance on the pretest as shown in Table 6 (analysis of variance or ANOVA shows p=0.972). However, the distribution of posttest performance of the incentivized group is significantly better than the comparison group with a p-value of 0.001. Table 6 also shows that generally the students with lower performance on the pretest benefit more from the explicit incentive to correct their mistakes in the pretest. In particular, the comparison group students' average posttest performance on these problems is comparable to their performance on the pretest but the average posttest performance on these problems is significantly better than their performance on the pretest for the students in the incentivized group. Overall, the students with lower performance on the pretest benefit more from the explicit incentive to correct their mistakes in the pretest.

Before we focus closely on the change in each student's performance from the pretest to posttest on problems repeated a second time, we note that the fact that the average score on the posttest is comparable to the pretest in the comparison group suggests that the assumption that the senior-level physics majors will automatically learn from their mistakes may not be valid.

Examination of the performance of students in the comparison group shows that some students did well both times or improved in performance but others did poorly both times or deteriorated on the posttest. Moreover, students struggled the most on the momentum problem both on the pretest and posttest and regressed the most from pretest to posttest on the measurement problem. Moreover, examining the students who regressed from the pretest to posttest, we observe a pattern in which students who answered a question correctly on the pretest employed a different procedure for the same question on the posttest. The procedure used was often a technique learned in the second half of the course and was not relevant to solving the problem. This is suggestive that some students are applying memorized procedures, rather than trying to actually understand the problem they are solving.

In the comparison group, in many of these cases in which students performed poorly, students wrote extensively on topics that were irrelevant to the question asked. It is hard to imagine that the students did not know that the things written by them were not relevant to the questions asked. It is possible that the students thought that if they wrote anything that they could remember about the topic (whether relevant or not) they may get some points on the exam for trying. Often, the irrelevant writings of a student on a particular question were different in the pretest and posttest. The poor performance of the students both times suggests that when each pretest was returned to them and the correct solutions were provided, they did not automatically use their mistakes as an opportunity for learning.

Figure 5 shows the average question gain (defined as the arithmetic difference between posttest and pretest averages for the questions; gain can therefore range from -100% to +100%) vs pretest score (which can range from zero to 100%) for each student on the four questions repeated from the pretest to posttest in the comparison group and incentivized group. Red-filled triangles are for each student from the comparison group and the blue-filled squares are used for each student in the incentivized group. Students whose score improved are above the horizontal axis at zero gain; students whose performance deteriorated are below the horizontal axis. The regions of possible gain are shaded according to posttest score performance categories: green for High posttest performance, yellow for Medium posttest performance and orange for Low posttest performance. The performance categories are defined as follows: "High," for scores from 85% to 100%; "Medium" for scores from 50% to 85%; and "Low" for scores from zero to 50%. The 50% cutoff was chosen somewhat arbitrarily and the 85% cutoff was chosen so that roughly one third of the students scored in the High performance category on the pretest.

Figure 5, and Tables 6,7 all show that students with poor performance on the midterm exam were likely to benefit from self-diagnosis activities in which they submitted the corrected midterm exam solutions for 50% of the points lost on each problem. Therefore, the gap between the High and Low performers on the midterm exam was reduced for this incentivized group on the repeated problems on the final exam. On the other hand, for the comparison group in which students did not correct their mistakes but were given the correct solutions to all of the midterm exam problems, the gap remained, i.e., scores did not substantially improve, and remained diverse.

The data were analyzed by breaking the students into three groups based on their pretest performance as shown in both Figure 5 and Table 7. The initially high-performing students from both the comparison and incentivized groups (scoring 85% and higher on the pretest) generally performed very well on the posttest regardless of the intervention (see Figure 5). Most of these students who start in the High pretest category stay in that category. Students who initially performed at a Medium level on the pretest (scoring between 50% and 85%) in the incentivized group perform better on the posttest than the corresponding students who were in the comparison group. In the comparison group, students in the Medium performance category on the pretest were as likely to improve on the posttest (above the horizontal axis in Figure 6A) as they were to deteriorate (below the horizontal axis). In contrast, in the incentivized group, almost all of the students in the Medium category on the pretest improved on the posttest (see Figure 6B). Furthermore, about half of these students in the incentivized group improved as much as possible on the posttest, saturating the boundary for maximal improvement (see Figure 6B). Among the initially Low performing students (pretest scores less than 50), many students in both comparison and incentivized groups improved on the posttest. However, the degree to which these struggling students performed on the posttest is highly dependent on whether or not they received a grade incentive to improve. The students in the Low category on the pretest in the incentivized and comparison groups had an average gain of 45.0% and 16.2%, respectively (Table 7). In summary, the gains are much larger for the incentivized group, bringing the average of the Low category to the level of the Medium category, and the Medium category to the High category (see Figure 5 and Table 7).

3.5 DISCUSSION AND SUMMARY

A common hypothesis is that advanced students will benefit from solutions to exams provided after the exams because they would want to learn from their mistakes. However, we find that many students who do not receive an incentive to learn from their mistakes, in fact, do not learn from their mistakes. More encouragingly, students who are given an incentive (earn back points for correcting their mistakes), perform substantially better.

Prior research on problem solving and self-monitoring skills of introductory physics students demonstrates that the introductory students do not learn these skills automatically, e.g., by listening passively to lectures and having access to solved examples [39–45]. Moreover, many introductory physics students are "captive audiences" – they may not buy into the goals of the course and their main goal becomes getting a good grade even if their learning is superficial [32]. Research suggests that these introductory physics students can benefit from explicit guidance and feedback in developing problem solving and learning skills and alignment of course goals with assessment methods [32–36, 39–45, 47, 48]. It is commonly assumed that the learning skills of students in advanced physics courses are superior to those of students in introductory courses. One reason offered is that advanced physics majors have chosen this major and are therefore eager to learn the material. They will make every effort to repair, organize and extend their knowledge because they are intrinsically motivated to learn and are not grade driven. For example, instructors often believe that physics seniors in an honors-level quantum mechanics course can monitor their own learning and they will automatically take the time out to learn from their mistakes. Contrary to these beliefs, our earlier investigation [46] found that advanced students in the honors-level quantum mechanics sequence did not automatically improve their performance on identical questions given in a midterm and on the final exam. The students were provided the correct solutions and their own graded exams. Even then, there was an apparent lack of reflective practice by supposedly mature students and many students did not take the opportunity to repair and organize their knowledge structure. In individual interviews, we probed students' attitudes and approaches towards problem solving and learning, and also asked them to solve the same problems again. The statistical results were consistent with students' "self-described" approaches towards problem-solving and learning. In the interviews we also find evidence that even in these advanced courses there are students who do not use their mistakes as an opportunity for learning and for building a robust knowledge structure; they resort to rote learning strategies for getting through the course. One interviewed student alluded to the fact that he always looked at the correct homework solutions provided but did not always look up the correct midterm exam solutions partly because he did not expect those questions to be repeated on the final exam. This tendency to learn the problems that may show up on the exam without making an effort to build a knowledge structure is typically not expected of physics seniors. Individual discussions with some physics faculty suggests that sometimes their incorrect inferences about advanced physics students' learning and self-monitoring skills is based on the fact that they feel that all physics majors are like them.

They may not appreciate the large diversity in the population of physics majors today and may not realize that those who become college physics faculty consist of a very select group of physics majors. While longitudinal research is needed to investigate the differences between those advanced students who become physics faculty and those who do not, it is possible that those students aspiring to be physics faculty are intrinsically motivated to learn and make more effort to learn from their own mistakes on their own. Similar to introductory students, advanced physics students who do not automatically use their mistakes as a learning opportunity may benefit from explicit scaffolding support and guidance to help them become independent learners. Students will automatically use problem solving as an opportunity for reflecting and learning if they are intrinsically motivated to learn the content and to extend and organize their knowledge [49–52]. However, students who are not intrinsically motivated may need extrinsic motivation, e.g., explicit reward for developing higher order thinking and self-monitoring skills. One strategy is explicitly asking students to fix their mistakes in homework assignments, quizzes and exams. Such strategies have been used for introductory courses with different levels of scaffolding provided to students for diagnosing their mistakes [39–43].

We investigated how well students in a junior-senior level undergraduate quantum mechanics course learn from their mistakes on physics problems on midterm examinations when provided with time and grade incentive for written corrected solutions to midterm-exam problems (after the midterm exams with comments were returned to them) and the effects of this self-diagnosis activity on subsequent problem solving on some of the same problems repeated later on the final exam.

In particular, we administered four quantum physics problems in the same semester both in the midterm and final exams to students enrolled in the honors-level quantum mechanics. In the first and third year of the study, students were not provided any incentives to learn from their mistakes on the midterm exam (comparison group) but in the second and fourth years, students were given grade incentives to correct their mistakes (experimental group). Solutions to all of the midterm questions were available to students on a course website except that the experimental group was provided the solutions after the students had submitted their corrected midterm exam solutions.

Our goal was to explore the extent to which these advanced physics students use their mistakes as a learning opportunity [53] and whether their performance on problems administered a second time on the final exam is significantly better than the performance on those problems in the midterm exams. We looked for correlation between the midterm exam score and final exam score on the four common problems for each group (experimental and comparison groups). Our study suggests that students with poor performance on the midterm exam benefit the most from self-diagnosis activities in which they submitted the corrected midterm solutions for 50% of the points lost on each problem. In particular, the gap between high and low performers on the midterm exam shrinks for the experimental group on the repeated problems on the final exam. On the other hand, for the comparison group in which students did not correct their mistakes, the gap remains.

Our research suggests that students may be more motivated to engage with instructional material in a more meaningful way if they are provided a grade incentive to correct their mistakes. Considering the relative ease with which instructors in physics courses at all levels can implement the intervention in which students are given grade incentives to correct and learn from their mistakes, instructors at all levels should consider giving students this opportunity to learn and develop their problem solving, reasoning and higher order thinking skills. Asking students to correct their mistakes in several courses may also help students understand the importance of learning from mistakes and the role of appropriate struggle in learning physics. Providing students with incentives to learn from their mistakes may positively impact the preparation of an additional one million STEM professionals in ten years according to the PCAST report [54] without a significant effort on the part of the instructors.

Many research-based methods require a substantial initial or recurring capital investment of some kind. For example, training for instructors, instructor preparation time, class time, and materials are typical costs associated with the introduction of an effective technique. However, this method does not require exceptional effort on the part of instructors, nor does it require class time. Like many effective research-based methods, the student is the active agent. It has been found that for introductory students, the corrections themselves take some time to evaluate, incurring some ongoing instructor labor. However, for upper level students we find that students seize the opportunity and produce expert-like solutions. Therefore, the additional instructor labor is minimal, due to the high quality of the solutions. In any case, the additional labor is no more than an additional round of exam evaluations.

While these results are encouraging, caution is urged in interpreting improvement. Our findings support the claim that students improve on problems administered a second time when expert-like behavior (self-diagnosing and correcting their mistakes) is explicitly incentivized. It does not necessarily follow that students have become adept at self-monitoring skills from just two such interventions in the two midterm exams in quantum mechanics. In particular, we compared the performance of students in the incentivized and comparison groups on four other problems on the same final exam for which incentivized group students did not diagnose their mistakes. We find that while the incentivized group scored higher than the comparison group, the results are not statistically significant.

3.6 CONCLUSION

Offering grade incentives to diagnose and correct mistakes can go a long way to close the performance gap between struggling and high-performing students. If this type of easy-to-implement intervention is implemented routinely in all STEM courses, students are likely to use their mistakes as a learning opportunity and may even develop better self-monitoring skills over time. We note that we are advocating for instructors to give students incentives to correct their mistakes in all STEM courses; we are not suggesting that they repeat the same questions on a final exam as was necessary for the purposes of this research.

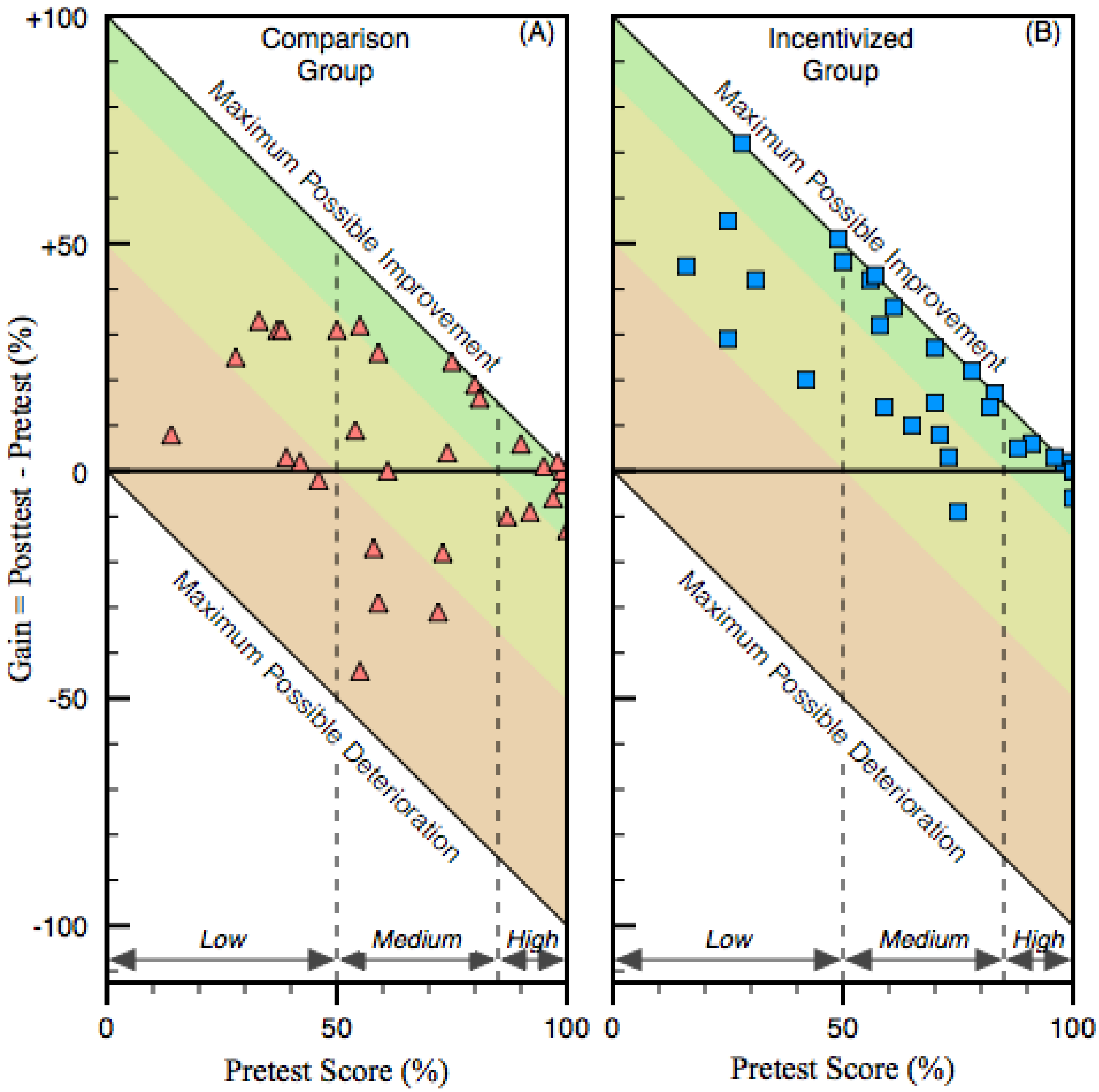


Figure 5: Gain (defined as the difference between posttest and pretest score) vs pretest score for each student on the four questions repeated from the pretest to posttest in the comparison group (A) and incentivized group (B). Red-filled triangles are for each student from the comparison group and the blue-filled squares are for each student in the incentivized group in which students received an explicit grade incentive to correct their own mistakes in pretest before answering the same questions on the posttest. Students whose scores improved are above the horizontal axis; students whose performance deteriorated are below the horizontal axis.

Table 6: Pretest, posttest, and average gains for students in the comparison and incentivized groups broken down into low, medium and high performance categories based on students' pretest score and for all students. While the pretest scores are comparable for the comparison and incentivized groups, the posttest scores are significantly higher for the incentivized group.

	Comparison Group ($N = 33$)			Incentivized Group ($N = 30$)		
	Pre	Post	Gain	Pre	Post	Gain
Low	34.6	50.8	+16.2	33.2	78.2	+45.0
Medium	64.7	66.3	+1.6	67.3	88.5	+21.3
High	96.0	93.0	-2.9	96.9	98.1	+1.2
All	**67.9**	**71.5**	**+3.6**	**67.6**	**88.4**	**+20.8**

Table 7: Average pretest and posttest scores and normalized gain, G, for each question for the comparison (left) and incentivized (right) groups for each of the four questions(G is the change in performance divided by the maximum possible improvement). Possible positive normalized gains range from zero (no gain) to +1.00, representing maximal improvement. In all cases, the incentivized group showed much higher gains. We note that the normalized gain of the comparison group on problem 4 is negative, but negative normalized gains are not meaningful so we have not reported it in this table.

	Comparison Group ($N = 33$)			Incentivized Group ($N = 30$)		
	Pre(%)	Post(%)	G	Pre(%)	Post(%)	G
Question 1	63.1	74.7	+0.31	88.4	99.2	+0.93
Question 2	62.7	65.2	+0.07	67.3	92.4	+0.77
Question 3	36.0	48.3	+0.19	33.1	65.5	+0.48
Question 4	88.3	80.0	*	81.7	94.6	+0.71
Overall	67.9	71.5	+0.11	67.6	88.4	+0.64

4.0 SURVEYING STUDENT UNDERSTANDING OF THERMODYNAMIC PROCESSES AND FIRST AND SECOND LAWS I: DEVELOPMENT AND VALIDATION OF A SURVEY

4.1 INTRODUCTION

The goal of Physics Education Research (PER) is to improve the effectiveness of physics teaching and learning. Evaluation of student learning based solely upon tests made by the instructors is subjective, and is not necessarily a good measure of the effectiveness of a curriculum or instruction. If well-designed pretests and posttests are administered before and after instruction in relevant concepts, the effectiveness of a curriculum and instruction in a particular course can be objectively measured. A standardized test can help to collect data on student performance; these data can be analyzed to evaluate the effectiveness of instruction using a particular instructional approach. This type of analysis can inform us about the effectiveness of various types of pedagogies and curricula in helping students from different demographics learn physics.

Research-based conceptual multiple-choice surveys are useful tools for evaluating student understanding of various topics and carefully developed and validated surveys can play an important role in measuring the effectiveness of a curriculum and instruction. When compared to free response, multiple choice is free of grader bias and such tests can be graded with great efficiency. Furthermore, the results are objective and amenable to statistical analysis so that different instructional methods or different student populations can be compared. The Force Concept Inventory (FCI) is a conceptual multiple-choice survey that helped many instructors recognize that introductory physics students were often not developing a functional understanding of force concepts in a traditionally taught course even if they could solve quantitative problems assigned to them by using a plug-and-chug approach [110]. Other conceptual surveys at the introductory physics level have also been designed for many physics topics, including surveys for electricity and magnetism [4, 5], kinematics (represented graphically) [55], energy and momentum [56], rotational and rolling motion [57], circuits [58], and Gauss's law [7]. These surveys reveal that traditional instructional approaches are not effective in helping a majority of students develop a good understanding of these topics and students in introductory physics courses have many common conceptual difficulties about different topics. Research-based instructional strategies have been successful in significantly improving students' conceptual understanding of some of these topics.

In order to choose appropriate distractors for alternative choices in a multiple-choice survey (so that students who do not know the correct answer are unlikely to select the correct answer by guessing it), interviews with students should be conducted and free response questions should be administered as part of a written test to appropriate student populations in various relevant courses. The frequency of most common responses can then be used to design incorrect choices in the multiple-choice questions. In addition, to ensure that the questions are consistent with the goals of a course and the wording is unambiguous, expert validation is also essential.

4.2 TEST DEVELOPMENT AND VALIDATION

To explore the effectiveness of curriculum and instruction related to fundamentals of thermodynamic processes and the first and second laws that are typically covered in both algebra-based and calculus-based introductory physics courses, we developed and validated a 33-item conceptual multiple-choice test which we call the Survey of Thermodynamic Processes and First and Second Laws (STPFaSL). We wanted the survey to be basic enough that the content in the survey is appropriate for both algebra-based and calculus-based introductory physics courses in which these topics in thermodynamics are covered. The survey development started by consulting with many instructors who teach introductory courses in which these topics in thermodynamics are covered and asking them about the goals they have when teaching these topics and what their students should be able to do. In addition to carefully looking through the coverage of these topics in several algebra-based and calculus-based introductory physics textbooks, we also browsed over homework, quiz and exam problems that faculty in introductory algebra-based and calculus-based courses at the University of Pittsburgh (Pitt) had given to their students in the past before starting the development of the survey.

While designing the STPFaSL, we paid particular attention to the issues of reliability and validity. Test reliability refers to the relative degree of consistency between the scores if an individual immediately repeats the test, and validity refers to the appropriateness of interpreting the test scores. We note that the survey is appropriate for making interpretations about the effectiveness of instruction on relevant topics in a particular course and it is not supposed to be used for high stakes testing of individual students. Also, although the survey focuses on topics that are typically covered in introductory thermodynamics and appropriate for introductory students, it is also appropriate for administering to advanced students, e.g., undergraduate students in upper-level thermodynamics course in which the same topics are generally repeated and to physics graduate students.

Before developing the survey questions, we first developed a test blueprint to provide a framework for deciding the desired test attributes. The specificity of the test plan helped us to determine the extent of content covered and the complexity of the questions. The preliminary distribution of questions from various topics was discussed and agreed upon with several introductory physics course instructors at Pitt.

In developing good alternatives for the multiple-choice questions, we took advantage of prior work on student difficulties with relevant topics in thermodynamics (see chapter 5 for detailed discussion and references). To investigate student difficulties further, we administered a set of free-response questions in which students had to provide their reasoning. The answers to these open-ended questions were summarized and categorized, and this helped us develop good alternatives for the questions in the survey based upon common difficulties. We note that the alternative or incorrect choices for each question often had distractors which reflected students' common alternative conceptions to increase the discriminating properties of the question. In particular, having good distractors as alternative choices is important so that the students do not select the correct answer for the wrong reason. Statistical analysis was conducted on the preliminary versions of the multiple-choice questions that were developed which helped refine the questions further.

We also interviewed individual students using a think-aloud protocol at various stages of the test development to develop a better understanding of students' reasoning processes when they were answering the free-response and multiple-choice questions. Within this interview protocol, students were asked to talk aloud while they answered the questions so that the interviewer could understand their thought processes. Individual interviews with students during development of the survey were useful for an in-depth understanding of the mechanisms for common student difficulties and to ensure that students interpreted the questions appropriately. In addition to these advantages, interviews offer unique insight into students' attitudes that affect learning and the use of thermodynamics principles and concepts to solve problems. During the think-aloud interviews, some previously unnoticed difficulties were revealed. These common difficulties were incorporated into the wording of the multiple-choice questions in the survey developed.

During the initial stage of the development process, 15 students in various algebra-based and calculus-based physics courses and a few undergraduates who had learned these concepts in an upper-level thermodynamics course participated in think-aloud interviews. The purpose of involving some upper-level students in these interviews was to compare the thought processes and difficulties of the upper-level students in these courses with introductory students. As discussed in detail in the next chapter(chapter 5, the students' reasoning difficulties across different levels are remarkably similar except in a few instances, e.g., upper-level students were more facile at reasoning with PV diagrams than introductory students. Nine additional interviews, drawn from a pool of students in introductory courses who were currently enrolled in a second semester course after finishing the first semester course in which mechanics and thermodynamics were covered, were conducted with the STPFaSL when it was close to its current form (see appendix C) to refine and tweak the wording of the questions.

As noted, in addition to developing good distractors by giving free-response questions to students and interviewing students, ongoing expert feedback is essential. We not only consulted with faculty members initially before the development of the questions, but also iterated different versions of the open-ended and multiple-choice questions with several instructors at Pitt at various stages of the development of the survey. Four professors at Pitt reviewed the different versions of the STPFaSL several times to examine its appropriateness and relevance for introductory algebra-based or calculus-based courses and to detect any possible ambiguity in item wording. Two professors and a graduate student from other universities who have been extensively involved in physics education research also provided extremely valuable comments and feedback to fine-tune the survey.

Since we want STPFaSL to be administered in one 50 minute long class period, the final version of the survey has 33 multiple-choice questions. We find that most of the students are able to complete the survey in one class period. Each question has one correct choice and four alternative or incorrect choices.

In addition to student difficulties with content discussed in chapter 5, interviews also suggested that students had attitudes and epistemology about learning thermodynamics that can impact their learning. For example, unlike instructors who felt that they assigned homework and exam problems to students that can only be solved if students had a solid grasp of the concepts, interviewed students often felt that the laws were not particularly useful in solving problems. For example, one interviewed student from an introductory physics course noted, "I just remember the laws being like, obvious, self-explanatory to the point where like they weren't useful for problems like actual problems... so I didn't really memorize them. They were like really... generalizations about heat transfer and like natural processes, but I didn't actively use them when I was solving problems, so I used one of the equations." In response to an explicit prompt to explain the first and second laws of thermodynamics, another introductory student who had difficulty internalizing how the laws of physics could apply to real world noted, "Physics to me is just basically sorcery."

Other students who had learned some thermodynamics in both physics and chemistry courses often thought that there is a chemistry thermodynamics and a physics thermodynamics and what they had learned in each course may only be applicable in that particular course. For example, one student who correctly pointed out that entropy increases in a spontaneous process questioned whether what he said made sense in response to a physics question about a spontaneous process and wondered whether what he had stated about change in entropy for spontaneous reactions from a chemistry course would be valid in a physics context: “delta S is greater than zero for spontaneous reactions... oh, that’s just chemistry ugh.” In the reasoning related to a similar problem involving change in entropy, another student attempted to scribble the following equation but then stopped questioning whether what he was writing was physics related: “delta(H) = -T*del(S) + something else ... not physics related”. These types of dilemma related to whether thermodynamics learned in a chemistry course should be used to solve problems in physics and vice versa reflect students’ epistemological difficulties that can impact their learning.

4.2.1 Relevant knowledge of introductory students before instruction

Motivated by responses during the think-aloud interviews that suggested that students had some knowledge of thermodynamics from high school physics and chemistry courses, college chemistry courses or medical school entrance exam preparatory materials, we administered the following brief survey as bonus questions in an exam to students in a first semester of algebra-based physics course in which the instructor had started discussing thermodynamics introducing concepts such as temperature, heat capacity, thermal expansion and heat transfer but there was absolutely no instruction in the first and second laws of thermodynamics:

1. Describe the first law of thermodynamics in your own words.
2. Describe the second law of thermodynamics in your own words.
3. Describe other central or foundational principles of thermodynamics (other than first and second laws).

Of the 207 students, 134 chose to respond to at least some of these bonus questions (65%). Their responses about the laws of thermodynamics and the difficulties they reveal are shown in Table 8. In particular, in the brief knowledge survey administered before instruction in the first and second laws of thermodynamics, we find that while about half of the students stated that energy is conserved, e.g.,“Energy cannot be created or Destroyed” (52%), only 5% made a statement that includes heat transfer as part of the conservation law. Another frequent response was that heat itself is conserved, with 15% of students making statements such as “There is no loss of total heat in the universe.” These responses further give support to the interview findings that many students in introductory physics courses have learned about the first and second laws of thermodynamics before instruction in the physics course.

4.3 FURTHER VALIDATION OF THE SURVEY

Validity of the STPFaSL was established by two broad classes of parallel efforts: one is expert design using effective methods of development already discussed and the other is its validation *for a specific purpose*. For the STFaSL, the goal for the development of the survey was to measure the effectiveness of traditional and/or research-based techniques of teaching thermodynamics concepts included in the survey. Specifically, the survey is designed to be a low stakes test to measure the effectiveness of instruction overall in helping students in a particular course reason about the first and second laws of thermodynamics and thermodynamic processes after learning these topics in thermodynamics (which are typically covered in introductory physics courses but then repeated in upper-level undergraduate courses).

4.3.1 Topics

As noted, the selection of topics for the problems was made systematically and included consultation with faculty members who teach introductory thermodynamics (some of whom had also taught upper-level thermodynamics) about the types of problems they expected their students in introductory physics courses to be able to solve after instruction. The topics were chosen drawing from the existing literature regarding student difficulties in thermodynamics, input from student written responses and interviews and input from physics faculty members who teach these topics. The conceptual problems on the survey are at all levels of Bloom’s Taxonomy [59]. However, since most questions on the survey require reasoning, there are very few questions which can be answered at the lowest levels of Bloom’s Taxonomy (recall and comprehension). As part of the validation of the survey, we verified that experts

Table 8: Responses of introductory physics students in an algebra-based course about the laws of thermodynamics before instruction in these laws. The percentages are determined by only taking into account the students who answered the bonus questions.

Topic	Claim	Frequency (%)
First Law	Energy is conserved (no mention of heat)	52
	Energy is conserved, with heat somehow incorporated	5
	Heat is conserved	15
	Total first law-like responses	72
Second Law	Entropy increases always	14
	Entropy increases under some conditions	4
	Energy becomes unusable	10
	Heat flows from warmer objects to cooler objects	23
	Total second law-like responses	51

agreed on the content involved in the problems on the STPFaSL. The concepts involved in the problems have been independently determined by several physics faculty members who regularly teach thermodynamics and members of the Physics Education Research (PER) group at Pitt. A consensus was reached for each question and an excellent agreement was found between the responses of the traditional physics faculty members and members of the PER group at Pitt regarding the concept involved in each question.

Table 9 shows that the broad categories of topics covered in the survey are "'Processes,"' "'Systems,"' "'Quantities & Relations,"' "'Representation,"' the "'First Law of Thermodynamics,"' and "'Second Law of Thermodynamics"'. The Processes category includes items which require understanding of thermodynamic constraints such as whether a process is reversible, isothermal, isobaric or adiabatic. Also included are problems involving irreversible and cyclic processes. The Systems category includes items involving knowledge of the distinction between a system and the universe, items involving subsystems or isolated system. The Systems category also includes items in which a student could make progress by making use of the fact that the system is an ideal gas (e.g., for an ideal gas, the internal energy and temperature have a simple relationship which can be used to solve a problem). Quantities and Relations included survey items specific to a quantity such as internal energy, work, heat, entropy, and their quantitative relationships. For example, the relationship between work and the area under the curve on a PV diagram is tested in several problems. The Representation category includes items in which a process is represented on a PV diagram. Finally, the last two categories include items requiring the first law and second law of thermodynamics. We included questions about heat engines in with Second Law of Thermodynamics category (although heat engines involve both first and second law of thermodynamics) due to the particular focus of the only two problems on the survey that touched upon heat engines.

Table 9: Topics by item number. Pitt physics faculty members and the Pitt PER group independently reached a consensus on identifying topics involved in each problem. A yellow "M" indicates that a concept was mentioned but not required to solve the problem in the opinion of content experts. A pink "R" indicates that a concept is required to solve the problem. An "I" (also pink) indicates that a topic is implicitly required though not explicitly asked for or mentioned. For instance, generally, if a student must reason about a PV diagram to infer the heat transfer to a system, the concept of "work" is implicitly required.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Processes	Reversible	R	R	R	R											R	R			R									R					
	Irreversible											R	R	R	R		R							R	R	R				R			R	R
	Cyclic					R	R	R	R							R	R		R	R														
	Isothermal			R	R					R									R		R						M	R	R					
	Isobaric									R								M			R	R												
	Isochoric																	M				R								R	R			
	Adiabatic	R	R																R		R						M				R			
Systems	Systems & Universe	R										R	R	R	R					R				R	R	R			R			R	R	
	Isolated System											R	R		R		R							R	R							R	R	
	Ideal Gas			R	R													R										R	R		R	R	R	
Quantities & Relations	State Variables					R	R	R	R		R					R							R											
	Internal Energy		R	I		R				I		R			R	R		I				R	R	R				I	I	R	I	R		
	Relation to T		R	R						R								R										R	R		R			
	Heat				R				R	R	R			R		R			R	R		R						R						
	Work		I	R			I		R	I	I				I	I						I	R	R			I	I		I	R	I		
	Relation to p, V		R				R			R	R				R	R						R		R			R	R		R	R	R		
	Entropy	R			R			R					R		R	R	R			R			R		R				R	R			R	
	Relation to Q, T	R			R								R							R					R				R	R				
Representation	*PV* Diagram					R	R	R	R	R	R					R		R				R					R				R			
First Law	1st Law		R	R					R	R	R	R		R	R	R						R		R		R		R		R	R	R		R
Second Law	2nd Law											R	R	R	R		R			R					R	R			R	R			R	R
	Engine Efficiency																R			R														

Below, we discuss some quantitative measures, including item analysis and reliability to ensure that the survey is reliable, which is necessary for validity. When the quantitative measures are considered along with expert validation regarding the content, a picture of overall validity emerges.

4.3.2 Item Difficulty

The item difficulty is simply the average score on a particular item. On a multiple choice test, only a 0 or 1 was scored on each item depending on whether the choice selected by a student was incorrect or correct, so:

$$P_i \equiv \frac{N_{i1}}{N_i} \tag{4.1}$$

where N_i is the number of students who responded to the ith item, and N_{i1} is the number of students who responded correctly to the ith item. In cases in which blanks are scored as zeroes, N_i can be replaced with the number of students who took the exam. P_i is always between zero and one. Results are shown in Figure 6.

4.3.3 Point Biserial Coefficient

The Point Biserial Coefficient, or PBC, is designed to measure how well a given item predicts the overall score on a test. It is defined as the correlation coefficient between the scores for a given item and the overall score. For a multiple choice test, in which the student either scores 0 or 1 on an item, the PBC for the ith item is reduced to:

$$r_{ix} \equiv \frac{\bar{x}_{i1} - \bar{x}}{\sigma_x} \sqrt{\frac{P_i}{1 - P_i}} \tag{4.2}$$

where σ_x is the standard deviation of the test scores, $\bar{x}_{i1}$ is the average test score among students who correctly answered the ith item, and $\bar{x}$ is the average test score. The PBC can take on values between -1 and 1; a negative value indicates that otherwise high-performing students score poorly on this item, and otherwise poorly-performing students do well on the item. A widely used criteria [60] is that it is desirable for this measure to be greater than or equal to 0.2. Results for STPFaSL are discussed in the following section. The point biserial coefficients themselves are shown in Figure 6.

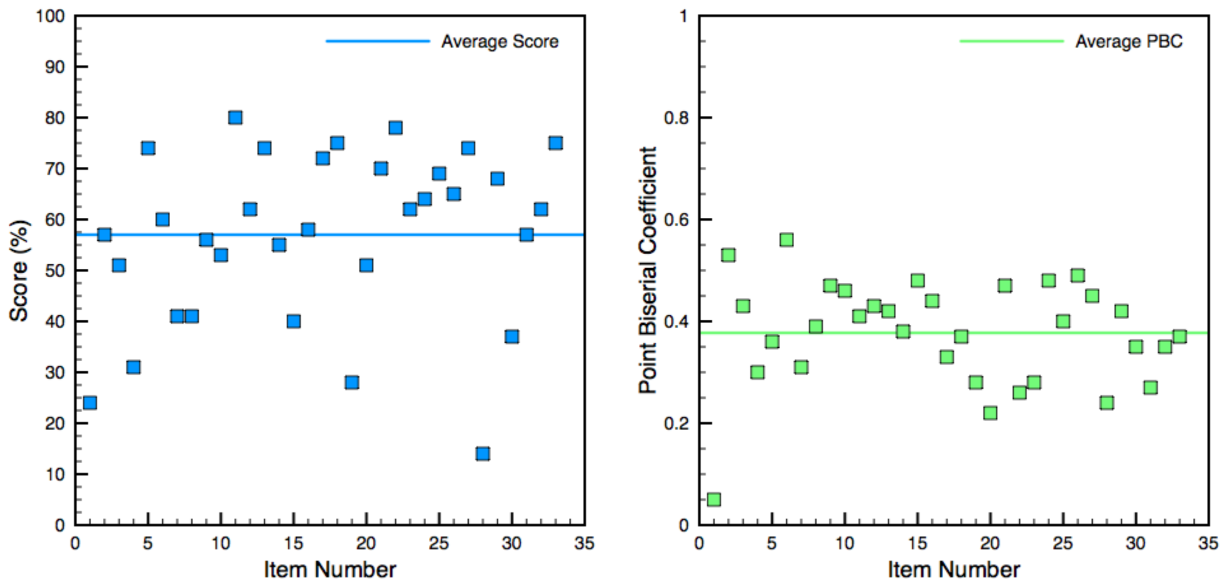


Figure 6: Item discrimination (left) and point biserial coefficient (right) for each item for upper-level students ($N = 147$). The PBC can range (in principal) from -1 to +1. The standard for a good item is that its PBC be 0.2 or higher, which has been exceeded for 32 of the 33 items on STPFaSL. The first item has low item discrimination because it exemplifies that even those students who perform well overall have difficulty distinguishing whether the change in the entropy of a system in a reversible adiabatic process is zero because the reversible process is adiabatic or whether the change in entropy of the system is zero in a reversible processes in general (confusion between the system and the universe).

4.3.4 Test Reliability

The overall reliability of a test as an evaluation tool is also important for establishing validity. One way to measure reliability of the test is to prepare an ensemble of identical students, administer the test to them, and analyze the resulting distribution of item and overall scores. Since this is generally impractical, instead, a method is devised to use subsets of the test itself, and consider this distribution. The KR-20 reliability index [61] takes this to its limit: a test of K items is notionally divided into K mini-exams, and the reliability between these scores is estimated.

$$R = \frac{K}{K-1} \frac{\sigma_x^2 - \sum_{i=1}^{K} \sigma_i^2}{\sigma_x^2} \tag{4.3}$$

KR-20

where σ_i^2 is the variance of the score across all students on the ith item. This is an approximation to KR-8:

$$R = \frac{\sigma_x^2 - \sum_{i=1}^{K} \sigma_i^2}{2\sigma_x^2} + \sqrt{\frac{\sum_{i=1}^{K} r_{ix}^2 \sigma_i^2}{\sigma_x^2} + \left(\frac{\sigma_x^2 - \sum_{i=1}^{K} \sigma_i^2}{2\sigma_x^2} \right)^2} \tag{4.4}$$

KR-8

where r_{ix} is the correlation between the ith item and the total score, the point biserial coefficient shown in eqn 4.2. The results for various groups are tabulated in Table 10.

4.3.5 Effect of Ordering Distractors on Student Performance

Another study was performed to assess one form of validity of the survey. In particular, the answer choices were re-ordered to determine if answer choice ordering had an effect on performance. One version was the original order, and three more versions differing only in answer choice order were administered to students in a calculus-based introductory physics course after instruction in relevant concepts. Performing a Kruskal-Wallis test for statistically significant difference between any of the four sets found no difference; and the p-value that differences between the four groups were due to chance alone was found to be 0.994. This study was performed with a total of 226 students who scored an average of 36.9%, which was typical of the performance of students in a calculus-based introductory course after instruction (posttest) (see Table10).

4.3.6 Performance and Statistical Reliability for Various Populations

The STPFaSL has been administered at Pitt and five other colleges or universities. Over two thousand students drawn from introductory and upper level undergraduate thermodynamic courses have participated (See Table 10). In addition, 45 entering physics graduate students in their first semester of the graduate program (who had not yet taken graduate level thermodynamics) participated in the survey. Although the survey is designed to be appropriate for students in introductory physics courses in which thermodynamics is covered, it is useful to benchmark the test by administering it to upper-level undergraduate students taking thermodynamics who learned the relevant material in that course and graduate students in physics.

In introductory courses, the pretest was administered before students learned about thermodynamics in that course and the posttest was administered after instruction in relevant concepts. We note that instructors generally administered pretests in their classes by awarding students bonus points as incentive to take the survey seriously but generally awarded students a small amount of quiz grade for taking it as a posttest. Moreover, since thermodynamics is covered after mechanics in the same course, some instructors teach it at the end of the first semester introductory physics course while others teach it at the beginning or in the middle of a second semester course. Moreover, some instructors only spent two weeks on these topics whereas others spent up to four weeks on these topics in the introductory courses. However, scores in introductory courses were essentially the same across courses taught by different instructors in different universities and regardless of whether the students were in an algebra-based or calculus-based introductory physics course and what the actual duration over which these topics were discussed.

The reliability of the test was also estimated via KR-20 and KR-8. The KR-8 is the exact form of the estimator; the KR-20 is an approximation which was designed (in the 1930's) to save computational labor[62]. The results of the two measures are typically very close, with KR-20 being designed to be a lower bound to the measure of reliability.

Another measure of test validity is that it closely correspond to models of expected performance. Particularly, we should expect introductory students to be out-performed by advanced undergraduates and graduate students. Similarly, pairs of graduate students working together should outperform those working individually (see Table 24). Similarly, we should expect introductory students in calculus-based or algebra-based courses to perform better post-instruction than pre-instruction; and we should expect their data to be more reliable (students are expected to be guessing on the survey items more pre-instruction than on the posttest). These data are tabulated in Table 10; and the expected trends are observed which serve as an additional measure of validity.

The mean performance across the 33 items and the number of students who participated in the survey at each level are shown in Table 10. "Upper" consists of both advanced undergraduates and entering graduate students in their first semester of the graduate program. Graduate pairs consist of small groups (20 pairs and one group with 3) of graduate students discussing and responding to the survey together. Faculty self-report that they perform nearly perfect, missing zero to two items. We note that as one considers levels from advanced to introductory, performance deteriorates. From pairs of graduate students to pretest scores for introductory students, the average performance drops from 75% to 29%. The reliability at each level also decreases. The survey reliability for upper-level students with KR-8 is 0.81, or estimated with KR-20 it is 0.79. The KR-8 reliability for graduate students alone and graduate student pairs is 0.88 and 0.89, respectively (or, as estimated with KR-20, 0.87 and 0.89, respectively). The reliability of the STPFaSL for introductory students post instruction is 0.72 (KR-20 is also 0.72). Reliability for introductory students pre-instruction is low and indicates that while they may, e.g., know *statements* of the first and second law (see 4.2.1), they have great difficulty *reasoning* about thermodynamics conceptual questions on the survey.

4.3.7 Detailed Analysis of Student Difficulties

Performance at the upper level (advanced undergraduate students and entering physics graduate students in their first semester of the graduate program) is shown in Figure 7. In addition to the individual interviews during the development process, since STPFaSL has been administered to nearly 2400 students in various courses, quantitative conclusions can be drawn about the prevalence of the many conceptual difficulties students at all levels have with these fundamental concepts in thermodynamics. (The prevalence of student difficulties is the topic of the next chapter, chapter 5.)

Some of these conceptual difficulties include difficulty reasoning with multiple quantities simultaneously, difficulty in systematically applying various constraints (for an isothermal, adiabatic, isochoric, reversible, or irreversible process, isolated system, etc.), difficulty due to oversimplification of the first law and overgeneralization of the second law. Many of these difficulties have already been documented [63–76]. Our findings demonstrate the robustness of the previous findings about student difficulties with these topics. Finally, detailed tabulation of the prevalence of these difficulties for each item can be found in appendix D.

4.4 DISCUSSION AND SUMMARY

We developed and validated a survey on thermodynamic processes and the first and second law at the level these topics are covered in an introductory physics course. The average individual scores on the survey from traditionally taught classes included in this study range from 30% to 60%, depending on whether the survey was administered as a pretest vs posttest and whether the students were in introductory or advanced courses. It is worth mentioning that scores for other conceptual standardized tests, e.g., the electromagnetism surveys listed below for traditionally taught classes, are similarly low:

- Brief Electricity and Magnetism Assessment [4]: 23% to 45%
- Conceptual Survey of Electricity Magnetism [5]: 25% to 47%

Both of these surveys on electricity and magnetism typically span topics covered in bulk of the second-semester introductory physics course. These low scores on the posttest indicate that traditional instruction methods are ineffective in helping students learn relevant concepts.

The survey can be used to measure the effectiveness of instruction in these topics in thermodynamics using a particular curriculum or pedgagogy. As a note of caution, we are not recommending that the survey be used as a high-stakes exam to evaluate the performance of a particular student, but as a measure of teaching effectiveness of a group of students. On the other hand, if students learn these topics using a research-based method and perform substantially higher than those reported here for traditionally taught courses, that instructional approach is likely to be superior to a traditional approach.

Table 10: The average performance of different groups on the 33 items or a subset of items and the number of students who participated in the survey at each level and reliability index. “Upper” consists of both advanced undergraduate students who had learned the relevant concepts in an upper-level thermodynamics course and entering physics graduate students in their first semester of the graduate program. Graduate pairs consist of small groups (20 pairs and one group with 3) of graduate students discussing and responding to the survey together. Faculty data are not shown here, as those data cannot be considered in a statistical manner. Faculty self-report that they perform near-perfect, missing zero to two items. Reliability is best measured by KR-8, and a lower bound is given by KR-20. Note that the reliability is significantly higher for post-instruction scores relative to pre-instruction scores, and for advanced students relative to introductory students, which provides validation to the content.

	Grad Pairs	Grad Ind.	Upper Post	Calc Pre	Calc Post	Algebra Pre	Algebra Post
N	21	45	147	705	507	218	382
Reliability (KR-8)	0.89	0.88	0.81	0.49	0.77	0.42	0.61
Reliability (KR-20)	0.88	0.87	0.79	0.43	0.77	0.34	0.57
Score (%)	75	55	57	29	37	30	37
First Law (%)	76	58	60	29	37	28	38
Second Law (%)	74	56	60	28	36	29	42
PV Diagram (%)	71	53	56	28	38	29	29
Reversible (%)	65	44	38	22	27	22	27
Irreversible (%)	79	62	66	32	40	32	45

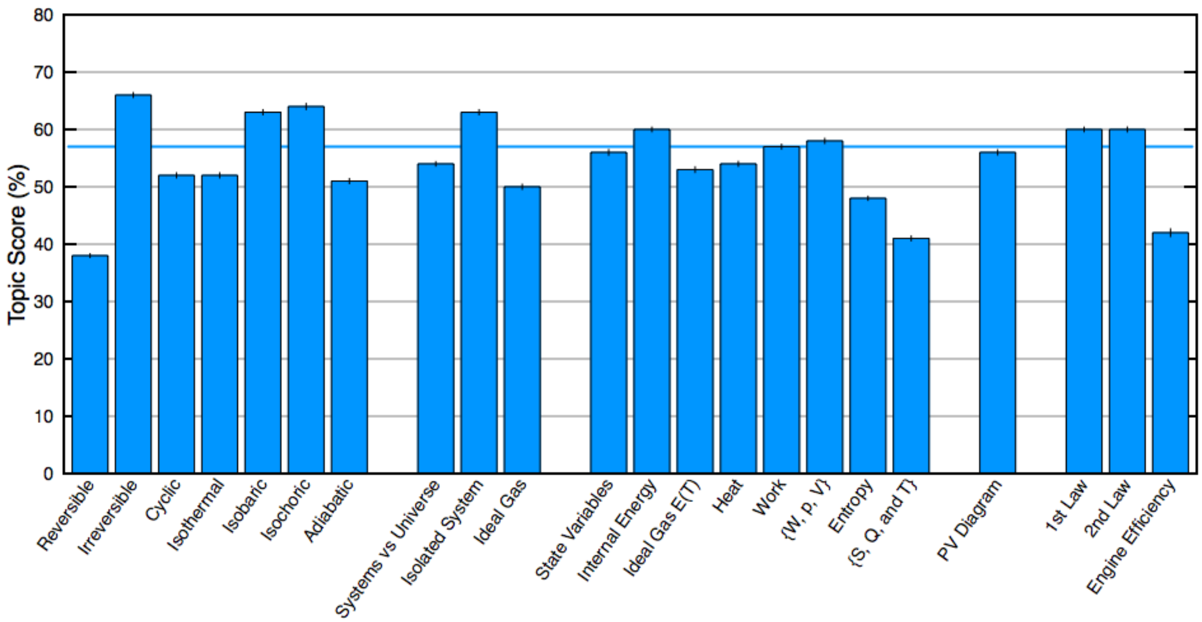


Figure 7: Average performance on the survey by topic for N=147 upper-level students. For completeness, error bars on each topic score (very small black vertical lines) indicate the sample error of the mean topic score, *assuming each topic is independent*. For comparing pairs of topics whose coverage on the STPFaSL is minimally overlapping, e.g., comparison of performance on problems involving "'Irreversible"' and "'Reversible"' processes, an assumption of independence may be appropriate, but otherwise the topics and their errors should not be directly compared within a population. Another pair of topics for which there is minimal overlap involves the "'second law"' problems and problems requiring knowledge of the "'state variables"'.

5.0 SURVEYING STUDENT UNDERSTANDING OF THERMODYNAMIC PROCESSES AND FIRST AND SECOND LAWS II: STUDENT DIFFICULTIES

5.1 INTRODUCTION

In the previous chapter, we discussed the development and validation of a conceptual multiple-choice Survey on Thermodynamic Processes and First and Second Laws (STPFaSL) of thermodynamics. The survey can be used as a pre-/posttest to evaluate the effectiveness of instruction related to these foundational topics generally discussed both in introductory and advanced undergraduate courses in which thermodynamics is included. During the development of the conceptual survey, students were administered free-response and multiple-choice questions on relevant topics and a subset of them was interviewed individually to understand the reasons for their difficulties. The students who were administered the written survey as a part of their course include those in introductory calculus-based and algebra-based courses, advanced undergraduate students enrolled in an upper-level thermodynamics course in physics departments, and first year physics graduate students who had not yet taken a core course in thermodynamics and statistical mechanics. We find that many students have serious conceptual difficulties related to relevant topics across all levels after traditional instruction.

Here, we discuss the evidence of student difficulties with various topics in thermodynamics found during the development and validation of the survey. Many of these conceptual difficulties have been discussed in previous studies and are referenced in different sections below as appropriate. Since this study involves a large number of students at various levels, it provides further support to the difficulties that have been documented in previous studies on related topics. While the degree to which students had difficulties *generally* decreased from the pretest to posttest in the introductory courses and from introductory to advanced courses, the relative prevalence of the difficulties with different topics is remarkably similar across all groups.

Even before instruction, many students in the algebra-based and calculus-based introductory physics courses show a trend in responding to the survey questions that suggests that they have some knowledge of the laws of thermodynamics and thermodynamic processes even before they learn thermodynamics in an introductory physics course. Individual discussions with some students suggest that they have learned about the first and second laws and thermodynamic processes in a high school physics or chemistry or a college chemistry course. Also, a majority of students in the algebra-based courses are bio-science majors and/or premeds and some of them have also learned about these topics from preparatory materials for the future medical school exam in which physics is a component.

The robustness of the conceptual difficulties in the pretest indicates that students are not blank slates with regard to these topics even before they learn them in an introductory physics course. The conceptual difficulties revealed in the posttest after instruction across all levels indicates that traditional instruction is not successful in reducing student difficulties and helping them build a coherent knowledge structure (including after an advanced undergraduate thermodynamics and statistical mechanics course). In other words, many students, including those who become physics graduate students, continue to struggle with these foundational topics in thermodynamics.

5.2 CHARACTERIZING STUDENT DIFFICULTIES

While attempting to categorize student difficulties by topics, it is important to keep in mind that student categorization of the problems is often context dependent, unlike the categorization of experts. In particular, the students may view two problems with the same deep features involving the same concepts and principles differently due to their different contexts or "surface features" [3]. While solving a problem, an expert and a student, who is developing expertise, may analyze the affordances and constraints of a problem very differently based upon their current expertise [107]. An expert is likely to take into account the various affordances and constraints appropriately while solving a problem using a systematic approach and navigate through the problem space in a manner that is likely to yield a correct solution [107]. On the other hand, a student may navigate through the problem space in a haphazardous manner and may not make the correct inferences based upon the information provided in the problem.

5.2.1 Inconsistency Across Different Contexts

Student responses to similar topics across different contexts are often inconsistent because many students do not have a robust knowledge structure related to these concepts and their reasoning skills are not sufficiently developed. According to Simon's theory of bounded rationality, an individual's rationality in a particular context is constrained by his/her expertise and experience and all individuals choose only one of the few options consistent with his/her expertise that does not cause a cognitive overload [107]. Simon framed these expectations and limitations during problem solving from by individuals with different levels of expertise in a particular domain as "satisficing" [107]. Even from the perspective of a knowledge-in-pieces framework, a particular context may trigger activation of knowledge that students think is relevant in that context and they may run with it and reach a conclusion that makes sense to them but that is not correct [108]. On the other hand, on another problem, the students may invoke the relevant piece of knowledge appropriately to solve the problem correctly. In other words, they may show different facility in their ability to solve two problems with the same underlying concepts and principles but with different contexts. In this sense, this knowledge-in-pieces view is similar to Simon's theory of "satisficing" in which individuals will select only a few of the large number of possible paths in the problem space that are consistent with their expertise in the domain which satisfies them and does not cause a cognitive overload to solve a problem. If an individual is not an expert in a domain, it is likely that the path traversed in the problem space while solving a problem in a particular context is not the one that will lead to success. When students satisfice, there is no need to discern the deep similarity between problems and transfer their reasoning from one problem to another problem involving the same underlying principles and concepts, because within the bounds of their expertise, the solution strategy that comes to their minds for each problem may be different and makes sense to them for each problem.

In fact, research suggests that student responses depend not only on the physical context of a problem, but they are also sensitive to the wording of a question, particularly for multiple choice questions which include an explicit mention of a particular student difficulty. The context dependence of student responses should be kept in mind in the discussion of the difficulties below. For example, even if only 10% of the students show a certain type of difficulty while solving a problem in a particular context, it may indicate that a higher percentage will display the same difficulty while solving another problem with the same underlying concepts posed in a different context.

In thermodynamics, there are many physical quantities, processes, and types of systems that one must consider to solve a problem in addition to the principles and laws that govern each situation in the problem posed. It is possible that certain information given in the problem statement (e.g., whether a process is reversible or not) is actually relevant in one situation to solve the problem but not in another. As suggested both by the bounded rationality framework and the knowledge-in-pieces framework, experts are likely to make appropriate use of the information provided in order to reduce the problem space and make progress towards a solution unlike students, who are still developing expertise, and may get distracted by some of the information provided. The students may not appropriately recognize or distill the affordances and constraints in a problem appropriately or may not know how to systematically exploit them to navigate through the problem space effectively [107].

5.2.2 Ignoring Relevant Information and Focusing on Irrelevant Details

When experts reviewed the STPFaSL survey, they noted that certain types of information in a problem were not relevant for solving a particular problem. For example, in item 17 on the survey, since the problem is about the internal energy of an ideal monatomic gas, the only affordance that experts focus on is that the temperature differences between the initial and final states in the two processes is given. They often explicitly noted that the information such as the fact that the first process is isochoric and the second process is isobaric is not relevant for solving the problem. However, consistent with prior expertise research [107, 108], while experts immediately focused on the temperature differences between the initial and final states as the relevant features of the problem, many students got distracted by the "surface" features. In particular, many students did not realize that the internal energy of an ideal gas is proportional to the absolute temperature (even though this information was provided to students on the cover page of the survey). In other words, what experts considered to be irrelevant details turned out to be the focus of many students' attention while reasoning about the problem. Similarly, on item 26, students were given a PV diagram and asked to compare the work done by the gas for an adiabatic and an isothermal expansion which start at the same state and end at the same volume. When explicitly asked, many experts who solved the problem stated that in order to find the work done by the gas, one only needs to find the area under the curve in the PV diagram and since the area is larger for the isothermal process on the PV diagram, most of the other information about the process is irrelevant to solving the problem. However, student responses suggest that at least 35% of the students in upper-level courses and somewhat less than 40% of the students in the introductory courses even after instruction did not know how to infer the work done by the gas in a process using the PV diagram. Interestingly, out of the five choices in this multiple choice question, only the correct answer to item 26 stated that a larger work is done for the isothermal process, which made the identification of the correct response very easy for experts. On the other hand, this problem was very challenging for students who often focused on the irrelevant details of the problem.

Individual interviews also reveal that while solving conceptual problems on the survey, students often ignore information they do not want to consider or do not know how to consider to solve the problems, which leads to mistakes. While certain information in the problem may in fact be irrelevant, students often ignore important information and hope that they will still be able to solve the problem correctly. We find that in addition to the information given in individual questions that many students ignored, some also ignored the information provided on the cover page of the survey. For example, the cover page contains the relation between the internal energy and temperature of an ideal monatomic gas and the definition of the term "adiabatic". However, students at all levels answered various survey questions as though they had not read or did not remember or did not find relevant this type of information on the cover page of the survey that was useful for solving problems. For example, on item 18, which explicitly asks students to identify the processes that have no heat transfer between the system and surroundings, approximately two-thirds of the calculus-based and one-fourth of the algebra-based introductory physics students, even after instruction, did not use the information on the cover page that an "adiabatic" process fits this category. In fact, the average student performance at each level was very similar for the problems which required knowledge of the fact that there is no heat transfer between the system and the surroundings in an adiabatic process when this definition was given on the cover page of the survey compared to an earlier version administered in the past in which this information was not provided. It is possible that students, who are still developing expertise in thermodynamics, do not have the cognitive capacity to keep this type of information given on the cover page in their working memory while solving the survey problems [107, 109]. In particular, if students' knowledge structure related to thermodynamics is not robust and their knowledge is in pieces [107–109], they are likely to ignore the relevant information on the cover page to manage cognitive load during problem solving. As discussed later, students often ignored important information provided in individual items as well while solving the survey problems.

Below, we describe different types of common difficulties that prevented many students from solving various survey problems in an expert-like manner. In all of the tables in the following sections, if the number of students across different questions with a similar type of difficulty is averaged, the corresponding tables in appendix D provide the breakdown for the number of students with that difficulty for each question separately. These tables are useful because they display whether the percentage of students at a particular level with a particular type of difficulty is stable across different problems or whether it is sensitive to the context of the problems. We note that the item numbers in the tables which have a star (asterisk) on them are not included in tabulating the prevalence of a particular difficulty because it was not possible to tease out the exact percentage from the data due to the prevalence of other difficulties. We also note that in some of the tables below, for comparison, the percentages of students in a particular group with the correct answer are provided (in boldface) before the percentages of students who have a certain type of difficulty are presented.

5.3 DIFFICULTIES WITH STATE VARIABLES

Several prior investigations have focused on student difficulties with state variables [63, 67, 71]. Consistent with some of these prior studies, we find that many students had difficulty distinguishing state variables from thermodynamic quantities which are path-dependent. Some of the difficulties with the survey problems concerning cyclic processes and/or with heat engines are tied to student difficulties with state variables.

Table 11 summarizes student difficulties in all groups with the concept of state variables. On item 22 on the survey, students were explicitly asked if, for a given thermodynamic system, the values of the work done by the system, the internal energy of the system and the entropy of the system are determined by the state of the system and not by the process that led to that state. Approximately half of the introductory students in both calculus-based and algebra-based courses and one fourth of the upper-level students provided incorrect responses to this question. Items 5–8 on the survey ask about ΔE, W, Q and ΔS in a cyclic process and item 15 asks about changes in some of these quantities in a reversible cyclic process. On the other hand, item 10 is a version of the problem in which a PV diagram is provided [67], and students are asked to compare Q for two different processes that share the same initial and final states. Across all these different problems, one common difficulty is incorrectly treating all thermodynamic quantities E, W, Q and S as though they are state variables or treating the set $\{E, W, Q\}$ as state variables but assuming $\Delta S > 0$ in various processes. However, Table 11 shows that some students did not treat even the internal energy as a state variable.

Table 11: Prevalence of Difficulties Related to State Variables

Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Not treating E as a state variable.	*5, 15, 22*	13	19	26	29	22
Treating W as though it is a state variable.	*6, 10, 22*	9	31	41	36	38
Treating Q as though it is a state variable.	*8, 10, 15, 18*	16	36	46	41	45
Not treating S as a state variable.	*7, 15, 22*	39	37	38	41	32
Treating all three quantities $\{E, Q, S\}$ or $\{E, W, S\}$ as state variables.	*15, 22*	11	28	31	30	35

5.3.1 Difficulties with State Variables in Cyclic Processes

Several prior investigations have focused on student understanding of cyclic processes that start and end in the same states [63–65, 67]. Consistent with prior research, we find that many students did not know that in a general cyclic process (i.e., not necessarily reversible), only the entropy of the *system* returns to the initial value. In particular, the entropy of the environment (surroundings) does not return to its original value after a complete cycle in an irreversible process. Many students only focused on the system and ignored the possible changes in the environment and treated all cyclic processes as reversible because the initial and final states are the same in a cyclic process.

Items 5-8, 15 and 18 are related to cyclic processes in different contexts. Since the internal energy E and entropy S are state variables, they are unchanged (i.e., $\Delta E = 0$ and $\Delta S = 0$) since the initial and final states of the system are the same after a complete cycle. On the other hand, the work done by the system W and the heat transfer to the system Q are path dependent and will differ depending upon the actual path traversed even if the initial and final states of the system are the same. However, some students over-generalized what they had learned about the state variables and some concluded, e.g., that all thermodynamic quantities return to the same value after a cycle (so that $Q = 0$ and $W = 0$ in cyclic processes). Tables 11 and 12 show the prevalence of the difficulties with these concepts. These tables show that many students incorrectly thought that there is no change in any of the thermodynamic quantities in a cyclic process (including work and net heat transfer on items 6, 8 and 15) because they treated all of the thermodynamic variables as state variables.

In addition, Table 12 also shows that approximately one fourth of the upper-level students and up to 50% of the introductory physics students in algebra-based courses even after instruction did not know that the internal energy of the system does not change in a complete cycle. Also, approximately 14% of the upper-level students and half of the introductory physics students even after instruction did not know that W and Q are not zero in a cyclic process.

Others claimed that all of the thermodynamic variables would return to the same value after a complete cycle only for a reversible cyclic process, and even for state variables such as the internal energy and entropy, the values would be different if the cycle is not reversible (for example, see student responses in Table 12 on items 5 and 7 related to internal energy and entropy, respectively, in a cyclic process). In particular, Table 12 shows that a significant percentage of students at each level incorrectly thought that the internal energy E and the entropy S of the system return to their initial values after a cycle only if the cycle is reversible and those who incorrectly thought that W and Q return to their initial values only if a cycle is reversible. Moreover, Table 12 shows that even on item 15 on which the problem statement explicitly stated that the cyclic process was reversible, some students at all levels incorrectly concluded that E and S do not return to their initial values and some students incorrectly thought that W and Q return to their initial values in a complete reversible cycle.

However, what is perhaps more surprising is that the upper-level students performed the worst of all groups on item 7 related to entropy changes for the system, ΔS, in a cyclic process and 59% of them did not realize that $\Delta S = 0$ for the system in a complete cycle. In addition, the introductory students performed worse after instruction (on the posttest) on the same problem (item 7 is about $\Delta S = 0$ in a complete cycle) compared to before instruction (on the pretest). The interviews suggest that this unexpected trend on item 7 (and also a similar trend on item 15) related to ΔS in Table 12 is due to students' common alternative conception that the entropy of the system increases in a cyclic process (in interviews, some of these students claimed that the entropy of the system should increase in all processes).

Table 12: Prevalence of Difficulties Related to Cyclic Processes

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
$\mathbf{\Delta E = 0}$ after a cycle. (correct)	***5***	**74**	**68**	**61**	**50**	**63**
$\Delta E \neq 0$ after a cycle.	*5*	26	32	39	50	37
$\Delta E = 0$ only if a cycle is reversible.	*5*	9	18	15	27	16
When explicitly given that a cycle is reversible, $\Delta E \neq 0$	*15*	10	19	31	25	23
$\mathbf{W \neq 0}$ for a cycle. (correct)	***6***	**83**	**49**	**28**	**37**	**28**
$W = 0$ for a cycle.	*6*	14	43	64	53	61
$W = 0$ only if a cycle is reversible.	*6*	5	5	8	6	8
$\mathbf{Q \neq 0}$ for a cycle. (correct)	***8***	**76**	**48**	**38**	**34**	**33**
$Q = 0$ for a cycle.	*8, 18*	11	29	39	31	38
$Q = 0$ only if a cycle is reversible.	*8*	3	7	9	10	12
When explicitly given that a cycle is reversible, $Q = 0$	*15*	32	62	75	75	76
$\mathbf{\Delta S = 0}$ after a cycle. (correct)	***7***	**41**	**53**	**55**	**43**	**62**
$\Delta S \neq 0$ after a cycle.	*7, 19**	59	47	45	57	38
$\Delta S = 0$ only if a cycle is reversible.	*7*	9	13	10	22	5
When explicitly given that a cycle is reversible, $\Delta S \neq 0$	*15*	35	28	32	26	27

When we interviewed students and faculty members, we find that physics experts and students, who are struggling with these concepts, reason about changes of different thermodynamic quantities associated with a system in cyclic processes very differently. Experts start by considering the distinction between state variables (e.g., internal energy and entropy) and other quantities (e.g., work and net heat transfer). For them, questions about changes in state variables associated with a system in a complete cycle are easy, and for those quantities that are not state variables, e.g., net work done by the gas in the entire cycle, they focus on the direction of the arrows on the PV diagram (e.g., to determine whether total work done by the gas is positive or negative for the entire cycle) and then use the first law of thermodynamics (conservation of energy) to make inferences about the net heat transfer between the system and surroundings for the complete cycle. On the other hand, many students, who do not have a clear understanding of state variables, use incorrect heuristics that were discussed earlier to answer these types of questions.

5.3.2 Difficulties with State Variables in Processes That Start and End at the Same States

As noted earlier, Table 11 shows student difficulties with state variables across various items on the survey regardless of whether the process is cyclic or not. For example, many students treated two processes which start at the same point and end at the same point on a PV diagram (same initial and final states) using reasoning similar to the one they used for cyclic processes. For example, on item 10, some students thought that the same amount of work is done in both processes that start and end at the same points. In fact, the performance of students on item 10 in the calculus-based courses is worse than those in the earlier studies by Meltzer on similar questions. The percentages reported by Meltzer [67] are closer to the performance of the upper-level students in this study. This difference may be due (at least partly) to the fact that the questions in this study had explicit distractor choices for the multiple-choice questions which were not given in the Meltzer study [67].

Table 13: Prevalence of Difficulties with Internal Energy (E)

Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Not recognizing that E is proportional to T for an ideal gas.	*17, 27, 30, 9*, 28**	27	49	51	56	51
$\Delta E \neq 0$ for an ideal gas that undergoes an isothermal process.	*27, 28*	24	42	43	40	40
$\Delta E \neq 0$ for an isolated system.	*11, 14, 23, 31*	16	26	32	28	40
$\Delta E = 0$ in an isochoric process	*17, 21, 29*	16	27	35	30	33
Incorrect sign of ΔE	*2*	27	47	52	40	58
Not treating E as a state variable.	*5, 15, 22*	13	19	26	29	22
$\Delta E \neq 0$ after a cycle.	*5*	26	32	39	50	37

5.4 DIFFICULTIES WITH THE INTERNAL ENERGY

Several prior investigations have focused on student difficulties with the internal energy of a thermodynamic system [63, 64, 66, 69, 77, 78]. Below, we discuss some of the difficulties we found in our research, many of which are consistent with prior studies (also see Table 13).

5.4.1 Confusing the Internal Energy With Free Energy

Interviews suggest that some students confuse internal energy with free energy although there is no question on the survey that explicitly asks about the free energy. Students often learn about the concept of free energy in their chemistry classes even if it is not necessarily discussed at length or not discussed at all in an introductory physics course in which thermodynamics is covered. During interviews, students with this confusion claimed that in order to reach equilibrium *internal energy* of the system tends to decrease. For example, when working on the survey problems while thinking aloud in a one-on-one interview, one introductory physics student after instruction in thermodynamics in the physics course commented, "System tends towards greater entropy and a lower internal energy state, which is why I was suspicious about all the ones that said it was going to a higher energy state." Another student noted that "The 2^{nd} law of thermodynamics states that the universe will proceed in a way that is energetically preferable, lower energy overall." Another student claimed, "Matter is most stable at the lowest energy" which shows confusion between "energy" and "free energy". Some of the students used the terms "energy" and "internal energy" interchangeably. Some students even confused free energy with potential energy. For example, one student noted that "Entropy $(S)-$ prefer a state of disorder, highest entropy is lowest potential energy so it is preferred."

5.4.2 Confusing the Internal Energy With Mechanical Energy

We also find evidence that some students are not clear about the difference between the internal energy of the system and the mechanical energy of the system. For example, in individual discussion, one student claimed that in a cyclic process, the work done by the system is zero, and therefore, the internal energy of the system is conserved. When asked further, the student reasoned that when the potential energy increases, the kinetic energy decreases such that the total internal energy remains unchanged. The student was drawing analogy with the fact that when the work done by the non-conservative forces on a system is zero, the total mechanical energy of the system is conserved with the fact that since the work done by the system is zero in a cyclic process, the internal energy of the system is conserved. The fact that often the same symbol E is used to denote both the internal energy and mechanical energy of the system can add to this type of confusion.

5.4.3 Not Recognizing That E Is Proportional to T for an Ideal Gas

Several prior investigations have focused on student difficulties with an ideal gas system [63, 68–70, 79–82] Some prior research studies have focused on student difficulties with the relation between the internal energy and temperature of an ideal gas [63, 77].

Many questions on the survey only require knowledge of the fact that for an ideal monatomic gas, the internal energy of the system E is proportional to the absolute temperature T. These questions are generally extremely easy for experts because as soon as they realize that the question is about the internal energy of an ideal gas and the temperatures of the initial and final states are provided, they are able to solve the problem without getting distracted by any other information. On the other hand, many students did not know (or did not recall) the relation between the internal energy and temperature of an ideal gas and they often focused on redundant features of the problems and their problem solving thereby often got derailed due to it. We also note that the relation between the internal energy and temperature of an ideal gas is provided on the cover page of the survey, but many students do not remember it or think it is as relevant when solving problems that require this information.

5.4.4 Not Recognizing That $\Delta E = 0$ for an Ideal Gas Undergoing an Isothermal Process

Table 13 shows that one common difficulty on items 27 and 28 on the survey related to a monatomic ideal gas undergoing an isothermal process was that the internal energy of the system will change in the process. This difficulty was often due to the fact that students thought that in an isothermal process, $Q = 0$. When these students made use of first law assuming $Q = 0$, they arrived at a non-zero value for ΔE in the isothermal process.

5.4.5 Assuming That $\Delta E \neq 0$ for an Isolated System

Table 13 also shows that across many questions on the survey, students often incorrectly thought that the internal energy is not conserved for an isolated system.

5.4.6 Assuming That $\Delta E = 0$ for an Isochoric Process

Furthermore, table 13 shows that across many questions on the survey, some students claimed that the internal energy is conserved in a constant volume process. Some interviewed students reasoned that if there is no change in volume, then $W = 0$ which implies that $\Delta E = 0$. This type of reasoning suggests that these students did not recognize the first law as a relation between three thermodynamic quantities and they ignored Q in their reasoning.

5.4.7 Incorrect Sign of ΔE

This difficulty will be discussed in detail later in the context of the difficulties with the first law in section 5.7.3, which delineates difficulties with inferring an incorrect sign of a physical quantity while using the first law (see Table 13).

5.4.8 Not Treating E as a State Variable

This difficulty has already been discussed earlier (see Tables 11, 12, and 13).

5.5 DIFFICULTIES WITH WORK IN THE CONTEXT OF THERMODYNAMICS

Several prior investigations have focused on the difficulties with work in the context of thermodynamics [63, 65, 67–69, 71–73, 75, 78, 83–86]. Most of our findings discussed below are consistent with prior research.

5.5.1 Difficulties in Applying Mathematical Representation of Work in Different Situations

Previous research has focused on student difficulties with the relation between work, pressure and volume [63, 67, 68, 83, 84]. We find that students had common difficulties in putting to use their mathematical definition of work. Sometimes, during interviews, students attempted to use a quantitative description of work. For example, after instruction in relevant concepts, one interviewed student stated, "Work equals P delta V....[after a pause] Negative P delta V." However, students with this type of a response in a mathematical representation often had difficulty relating the definition of W to a reasoning about W in a thermodynamic process (including cases in which a PV diagram was provided). In addition, the confusion of this student with the sign of W in his reasoning with the equation for W was common among interviewed students and often led to difficulties on problems involving the first law.

In particular, students who analyzed the problem using conceptual reasoning, in general, did better than those who focused on signs in a memorized formula when using the first law.

5.5.2 Assuming $W \neq 0$, Even Without a Change in Volume in a Spontaneous Process

Written responses (see Table 14) and interviews suggest that many students thought that the work done by the gas is positive in free expansion or in spontaneous processes such as heat transfer from one sub-system to another or mixing of gases in two chambers even if the total volume of the system does not change (questions 14, 23, 31). Some students argued that in a free expansion process and spontaneous mixing of gases in two adjacent chambers, the work done by the gas is non-zero because the gas molecules have more available volume to occupy. In interviews, some students argued that in these cases since "things are spreading out" (where 'thing' could even refer to heat transfer from one sub-system to another), work done by the gas is positive even if the volume of the system does not change. In the interviews, explicitly asking students about the work done by the gas in terms of pressure and change in volume was usually not productive in getting students to correctly consider the relevance of the pressure and volume to the work done by the gas in these situations.

5.5.3 Assuming $W \leq 0$ in an Isothermal Expansion of an Ideal Monatomic Gas

Table 14 shows that approximately one fourth of the students in both the algebra-based and calculus-based introductory physics courses thought that in an isothermal expansion of an ideal monatomic gas, the work done by the gas $W \leq 0$.

5.5.4 Assuming $W = 0$ in an Isobaric Process

Some prior investigations have focused on student difficulties with isobaric process [63, 78] Table 14 shows that on the survey, on item 21, some students confused a constant pressure process with a constant volume process. In particular, on item 21, in which a PV diagram is provided, some students thought that $W = 0$ in a constant pressure process (see Table 14).

Table 14: Prevalence of Difficulties with Work done by a system (W)

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
$W \neq 0$ for various spontaneous processes where $\Delta V = 0$.	*14, 23, 31*	33	41	47	40	50
$W \neq 0$ for a spontaneous process with only particle transfer and $\Delta V = 0$	*14, 31*	36	45	51	42	53
$W \neq 0$ for a spontaneous process with only heat transfer and $\Delta V = 0$	*23*	27	32	39	35	45
$W \leq 0$ in an isothermal expansion	*27*	10	22	33	25	39
$W = 0$ in an isobaric process	*21*	8	18	33	21	28
Not recognizing W as area under the curve in a PV diagram	*9, 26, 6*, 10*, 15*, 21**	29	45	59	49	58
When comparing an isothermal expansion to an adiabatic expansion on a PV diagram where both start in the same state and both have the same ΔV:						
$\mathbf{W_{isothm.} > W_{adi}}$ **(correct)**	***26***	**65**	**41**	**17**	**31**	**17**
$\lvert\Delta P_{adi}\rvert > \lvert\Delta P_{isothm.}\rvert \Rightarrow W_{adi} > W_{isothm.}$	*26*	11	19	25	26	27
$\lvert\Delta T_{adi}\rvert > \lvert\Delta T_{isothm.}\rvert \Rightarrow W_{adi} > W_{isothm.}$	*26*	12	15	23	20	27
$\Delta V_{adi} = \Delta V_{isothm.} \Rightarrow W_{adi} = W_{isothm.}$	*26*	10	18	22	14	18
Incorrect sign of W	*2, 6, 27**	24	38	48	54	49
Treating W as though it is a state variable.	*6, 10, 22*	9	31	41	36	38
$W = 0$ for a cycle.	*6*	14	43	64	53	61

It appears that these students did not know how to use the PV diagram to interpret W in the process. During discussion, one student stated that in a constant pressure process, the weights on the piston prevents the piston from moving so that the pressure remains constant.

5.5.5 Difficulties With Work as Area Under the Curve on a PV Diagram

Table 14 shows that many students did not correctly recognize W as the area under the curve on a PV diagram. In interviews, some students who state that work done by the system W is the integral of PdV (where P is the pressure and dV is the infinitesimal change in volume) or $P\Delta V$ (in an algebra-based introductory physics course) do not necessarily recognize W as the area under the curve on a PV diagram. In other words, they have difficulty connecting different representations of knowledge. In interviews, some students who knew the mathematical definition of work had difficulty reasoning about the relative amount of work done in different processes including in the context of processes for which a PV diagram for the process was provided. For example, some students did not even know that in a constant volume (isochoric) process depicted on a PV diagram (see [63, 78]), $W = 0$.

Even in situations in which consecutive problems related to PV diagrams primed interviewed students to think about work done by the system mathematically as PdV or $P\Delta V$, some students struggled. Those who realized that W is the area under the curve on a PV diagram did not necessarily associate the arrows on the process showing expansion or compression with work done by the system or work done on the system. In a cyclic process, some were confused about how to find the work done for an entire cycle. They did not know how to interpret whether W is positive or negative in the lower or upper parts of the cycle depicted on a PV diagram when the cycle was clockwise or counterclockwise. Discussions with students suggest that sometimes their instructors had only focused on the mathematical representation and sometime they had only focused on the diagrammatic representation but they had not necessarily focused on providing scaffolding support and helping students learn to connect those different representations of W. Thus, many students had difficulty connecting these different representations to solve the survey problems correctly.

5.5.6 Difficulties With Comparison of W in an Isothermal and Adiabatic Expansions

On item 26, in which an isothermal expansion is compared to an adiabatic expansion on a PV diagram, both processes start in the same state and both have the same final volume, and therefore the same change in volume ΔV. The correct response is that W is larger in the isothermal expansion because, for each small change in gas volume between the initial and final points, the pressure is higher. However, Table 14 shows that students at all levels have great difficulties with this question and many students claimed that W is larger for the adiabatic expansion because the change in pressure between the initial and final points is larger in that process or W is larger in the adiabatic expansion because additional work must be done to change the temperature or W is equal for both processes because the final and initial volumes are the same.

5.5.7 Incorrect Sign of W

This difficulty will be discussed later in the context of difficulties with the first law in section 5.7.3, which characterizes difficulties with inferring the correct sign of a physical quantity using the first law (see Table 16).

5.5.8 Treating W as Though It Is a State Variable

This difficulty has already been discussed earlier (also see Tables 11, 12 and 14).

5.6 DIFFICULTIES WITH HEAT TRANSFER BETWEEN THE SYSTEM AND THE SURROUNDINGS

Several prior investigations have focused on difficulties students have with heat transfer in thermodynamics [63–65, 67, 70–75, 78, 85, 86]. Many of the findings below are consistent with prior research.

5.6.1 $Q \neq 0$ in an Adiabatic Process

Several prior investigations have focused on adiabatic processes [63, 68–70]. Our findings are consistent with previous studies. When the earlier versions of the survey were administered to the students, they had difficulty with the meaning of the word adiabatic. Therefore, the later versions of the survey explicitly included the meaning of the word adiabatic on the cover page with other instructions. However, throughout the survey, many students do not make use of the definition of an adiabatic process consistently despite the fact that they are explicitly given that an adiabatic process is one in which there is no net heat transfer to or from the system (see data from item 18 in Table 15 in which students were explicitly asked to identify processes in which $Q = 0$). In fact, there is no statistically significant difference between student responses to questions requiring the knowledge of the fact that $Q = 0$ in an adiabatic process in the earlier version when the definition of an adiabatic process was not provided on the cover page vs. the later versions in which this information was provided.

5.6.2 $Q = 0$ in an Isothermal Process

Several prior investigations have focused on the confusion between an isothermal and adiabatic process [63, 64, 77–79]. Consistent with prior studies, we found that many students could not clearly distinguish between an isothermal and an adiabatic process and thought that there is no heat transfer in an isothermal process (see Table 15). Some of the interviewed students directly inferred that an isothermal process must have $Q = 0$ whereas others first noted that since the temperature does not change in an isothermal process, constant temperature implies no heat transfer between the system and the surroundings. For example, on one problem, a student stated "Same temperature. Net heat absorbed... there shouldn't be any because the temperature sticks to be the same." Another student who used a similar reasoning on a problem involving an isothermal process noted, "Same temperature? so there shouldn't be any net heat". When questioned, some of these students admit that they do not interpret isothermal as meaning "$Q = 0$" in all problems requiring this knowledge and their answers to some later survey questions are inconsistent with other answers that they had given earlier.

Table 15: Prevalence of Difficulties with Heat transfer to a system (Q)

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
$\mathbf{Q = 0}$ in an adiabatic process *(correct)*	***18***	**88**	**64**	**51**	**76**	**65**
$Q = 0$ for gas heated in a closed container	*20, 28**	23	39	38	31	40
$Q = 0$ in an isothermal processes	*3, 4*, 18, 27,* 28*	20	49	49	46	43
$Q = 0$ in a *reversible* isothermal processes	*3,* 4*, 28*	18	38	24	33	21
$Q = 0$ in a *reversible* isothermal process in which gas is heated in a closed container	28*					
$Q = 0$ in a *reversible* cycle	8*, *15*	32	62	75	75	76
$Q = 0$ in an isochoric process	*21*	12	26	32	24	29
Incorrect sign of Q	*3, 8, 9*, 10**	39	50	59	54	58
Treating Q as though it is a state variable.	*8, 10, 15, 18*	16	36	46	41	45
$Q_1 = Q_2$ if two process share initial and final states on a PV diagram.	*10*	22	53	67	63	70
$Q = 0$ for a cycle.	*8, 18*	11	29	39	31	38

We also find that the incorrect heuristic that an isothermal process implies a process with no heat transfer between the system and the surroundings led many students to conclude that the internal energy of an ideal gas should change in an isothermal process. For example, on the survey item 27, the most common difficulty was assuming that, in the isothermal expansion of the ideal gas, the internal energy of the gas changes and there is work done by the gas, but there is no exchange of heat between the system and surroundings. On a positive note, these students applied the first law of thermodynamics correctly after making the incorrect assumption that $Q = 0$ in an isothermal process.

Another example from the survey is the pair of items 3 and 4 which are about the net heat transfer to the system and the change in entropy of the system for a reversible isothermal compression of a monatomic ideal gas. Both these items are extremely difficult for all students including the upper-level students. On both problems, introductory students score approximately 30% after instruction whereas the upper-level students obtained roughly 50% and 30% on items 3 and 4, respectively. On item 3, a very common difficulty among students at all levels (approximately 30-40% at the introductory level and 20% in the upper-level) is that $Q = 0$ in the isothermal compression. On item 4, 30% of students answered the question correctly regardless of whether they were introductory or upper-level students and the most common difficulty was assuming that the entropy of the system remains the same in a reversible isothermal compression. In the interviews, some students explained this type of reasoning by noting that since the temperature is constant in an isothermal process, there is no heat transfer between the system and the surroundings, and therefore, the entropy of the system which depends on heat transfer cannot change. Other students provided different reasoning and claimed that the entropy of the system does not change because the process is reversible and 'entropy does not change in a reversible process'. These students were confusing change in the entropy of the universe with the change in entropy of the system. A third group of students thought that each of these terms "isothermal" and "reversible" mean $Q = 0$. This type of confusion was prevalent in some other problems as well. In fact, items 3 and 7 on the survey are two of the very few problems on which the performance of introductory and upper-level students are comparable because of the conceptual difficulties with reasoning related to changes in the entropy of the system in a reversible isothermal compression and in a cyclic process, respectively.

Even on item 28, which concerns an isothermal reversible expansion of an ideal monatomic gas, the fact that the process is isothermal and reversible reinforces some students' confidence that $Q = 0$ in the process (as noted, to some students isothermal means $Q = 0$, to others reversible process means $Q = 0$, and to some other students, each of these terms "isothermal" and "reversible" mean $Q = 0$).

Item 18 explicitly asked students to identify processes (among adiabatic, isothermal and cyclic processes) in which there is no exchange of heat between the system and the surroundings (see Table 15). About 25% of the upper-level students and more than 50% of the introductory students, even after instruction, claimed that either isothermal or cyclic processes or both of these types of processes involve no heat transfer between the system and the surroundings. Sometimes, when students claimed that $Q = 0$ for both isothermal and adiabatic processes and they were asked about the difference between the two processes, they stated that they did not remember the exact difference between these two processes but remember that both of these processes have $Q = 0$ while others claim that in an isothermal process both $Q = 0$ and $\Delta E = 0$ since the temperature is constant ($\Delta E = 0$ for an isothermal process involving an ideal monatomic gas). This latter group of students did not take into account the fact that there is work done by the ideal monatomic gas in an isothermal process but if both $Q = 0$ and $\Delta E = 0$, the first law of thermodynamics gives an erroneous conclusion about W.

5.6.3 $Q = 0$ for a Reversible Isothermal Process

This difficulty was discussed in the preceding section in the context of isothermal reversible compression or expansion of an ideal monatomic gas for items 3, 4, and 28 on the survey (also see Table 15).

5.6.4 $Q = 0$ in an Isochoric Process

Prior research has focused on student difficulties with $Q = 0$ in an isochoric process [63, 78]. Student responses to item 21 on the survey suggest that 12% of the students in an upper-level course and about one fourth to one third of the students in both algebra-based and calculus-based introductory physics courses even after instruction thought that $Q = 0$ in an isochoric process (see Table 15). Interviews suggest that this type of difficulty is sometimes due to applying the first law incorrectly by ignoring the changes to the internal energy of the system and thereby incorrectly claiming that $W = 0$ in an isochoric process implies $Q = 0$.

5.6.5 $Q = 0$ for a Gas Heated in a Closed Container

We find that some students have confusion between an isolated system vs. a closed container and assume that $Q = 0$ for a gas heated in a closed container (see Table 15). This difficulty can be due partly to the fact that sometimes instructors use the term "closed" system to describe an "isolated" system and students overgeneralize the phrase "closed system" to be synonymous with "closed container". On the survey, items 20 and 28 both involve a closed container in which there is net heat transfer to the gas. What is interesting about these systems is that while item 20 explicitly states that "the gas is heated with a burner" and item 28 states that "there is net heat transfer to the gas during the process", many students thought that there is no exchange of heat between the system and surroundings in these cases partly because the container is closed in each case (see Table 15). In item 28, the choice $Q = 0$ may be even more appealing since it involves a reversible process (in which some students incorrectly claim that $Q = 0$). In other words, the two phrases "closed container" and "reversible" sometimes reinforce incorrect thinking that $Q = 0$ for this process. Item 28 is the most difficult problem on the survey, on which students from all groups (including those in the algebra-based and calculus-based courses after instruction and upper-level students) perform extremely poorly (the correct response rate of all of these groups is between $9\% - 14\%$). On the other hand, some faculty members, who provided feedback during the development of the survey, have noted that the information "there is net heat transfer to the gas during the process" is not needed for item 28, but this information make the problem easier for students. Although we deliberately kept this information that some experts considered redundant in order to prime students to realize that the entropy of the gas increases in the process if it is being heated, it does not help a majority of the students. The most common incorrect answer for item 28 is that the entropy of the gas remains constant because the process is reversible.

5.6.6 Multiple Factors Can Increase Student Confidence That $Q = 0$ in a Process When $Q \neq 0$

Interviews suggest that when the problem has more than one factor that makes students think incorrectly that $Q = 0$ in a given situation, those factors often reinforce each other and give students confidence in the incorrect inference that $Q = 0$. For example, we have already discussed earlier that on items 3, 4, and 28, the fact that the processes were both reversible and isothermal made some interviewed students even more confident that $Q = 0$ in the given situations since they felt that an isothermal process should have $Q = 0$ and a reversible process should have $Q = 0$. Similarly, on item 15, the fact that the process is reversible and cyclic made some interviewed students more confident that $Q = 0$ for the process because they felt that a reversible process should have $Q = 0$ and a cyclic process should have $Q = 0$. Also, as noted in the preceding section, on item 28, which is the most difficult item on the survey with the maximum correct percentage from any group being 14%, the process is a reversible isothermal process in which the gas is in a closed container. As noted, in interviews, some students claimed that $Q = 0$ on item 28 and hence there is no change in the entropy of the gas. Some of these students thought that each of the factors (or at least two of these factors) "reversible", "isothermal" and "closed system" about the process ensured that $Q = 0$. In fact, even in the written responses, more than 25% of the students from upper-level, introductory algebra-based and introductory calculus-based courses claimed that the entropy of the gas does not increase, consistent with the interview finding regardless of the fact that the problem statement explicitly states that there is net heat transfer to the gas during the process. Thus, when multiple factors are present that make students think incorrectly that $Q = 0$, they may even miss the explicit statement that $Q \neq 0$ because there is net heat transfer to the system.

5.6.7 Incorrect Sign for Q

This difficulty will be discussed below in the context of difficulties with the first law in section 5.7.3, which is related to inferring an incorrect sign of a physical quantity using the first law (also see Table 15).

5.6.8 Treating Q as Though It Is a State Variable

This difficulty has already been discussed in an earlier section (also see Tables 11, 12 and 15).

5.7 DIFFICULTIES WITH THE FIRST LAW OF THERMODYNAMICS

A number of difficulties related to the first law of thermodynamics have been documented in prior research [63–76].

In the brief thermodynamic knowledge survey (see 4.2.1) discussed in more detail in chapter 5, in response to questions about the first and second laws of thermodynamics before instruction on these topics in an introductory algebra-based physics course, about half of the students had responses along the lines that energy is conserved although they did not specify the details of energy conservation. Another common response was that "heat" itself is conserved with 15% of the students making statements such as "There is no loss of total heat in the universe." While only mentioning energy conservation does not provide a good picture of student understanding of the first law and the statements that students made about the conservation of heat are not accurate, such responses suggest that students in the introductory physics courses have had previous instruction regarding the laws of thermodynamics (e.g., from high school physics courses, high school and college chemistry courses, and MCAT preparatory materials) before learning it in the introductory physics course. Some instructors who administered the survey noted that since the first law of thermodynamics is a conservation of energy law similar to those students have encountered in earlier courses, it is likely to be easier for students than the second law. They noted that the second law was difficult partly because it was the first time students had encountered a quantity that was not necessarily conserved in a process for an isolated system. These instructors predicted that, applicable to diverse situations, the form of the second law that states that the change in the entropy of the universe is always greater than or equal to zero in a process will be more challenging for students than applying the first law in different situations. Contrary to those instructors' prediction, students struggled with both laws. Moreover, we find that while many students indeed struggle with the second law and incorrectly assume, e.g., that entropy entropy of various systems or sub-systems is conserved across a variety of problems, others use the opposite heuristic and incorrectly assume that $\Delta S > 0$ in all situations for the individual system or sub-systems, and/or the combined system or universe. The difficulties with the second law will be discussed after the first law.

Individual interviews suggest that even after instruction, many introductory physics students only recalled fragments of the first law instead of a coherent statement of energy conservation involving work done by the system W, heat transfer to the system Q and change in the internal energy of the system ΔE. Evidence of the prevalence of difficulties with the first law is widespread throughout the written responses to the STPFaSL survey. In particular, 14 of the 33 problems can be solved using the first law (see the companion paper for a complete table of the thermodynamic concepts involved in various problems), although some of them can also be solved using an alternative reasoning not requiring the first law. For example, on items 14, 23 and 31 on the survey involving different types of spontaneous processes (free expansion, spontaneous heat transfer and spontaneous mixing, respectively), one can use the first law to correctly conclude that if there is no heat transfer between the system and the surroundings ($Q = 0$) and the work done by the system is zero ($W = 0$), the change in the internal energy, ΔE, is zero as well. However, the fact that the change in the internal energy must be zero can also be inferred without the use of the first law by simply noting that the internal energy of an isolated system cannot change. Similarly, on the same items 14, 23 and 31 on the survey, if a student recognizes that $Q = 0$ and $\Delta E = 0$, that student can use the first law to correctly conclude that work, $W = 0$. However, the use of first law is not mandatory to reason correctly that $W = 0$ in these problems if one recognizes, e.g., that there is no work done in any of these situations because the overall volume of the system is fixed. We also note that three other items on the survey (items 13, 25 and 33) related to why spontaneous processes do not occur in the opposite direction, include possible "violation of the first law" as a choice although this choice is incorrect. The average scores for students in different groups on these problems that touched upon the first law in some manner or another ranged from 29% to 59%.

5.7.1 Challenges Due to the Fact That the First Law Is a Relation Between Three Quantities

We find that the fact that the relation between three thermodynamic quantities is involved in the first law makes it challenging for students, who are not adept at reasoning about three quantities simultaneously. For example, when prompted about how they would describe the first law of thermodynamics to another student, some students only recalled one of the three quantities involved in the first law. For example, one introductory student after instruction stated: "Heat can't appear and disappear out of nowhere it's gotta come from somewhere. Is that conservation of heat?" The above student did not know the first law and omitted work and internal energy in his description of the first law of thermodynamics. In an individual interview, we also find that the same students reasoned about the first law by considering two of the three quantities in the first law and ignoring the third one, but depending upon the question, a different quantity is ignored in the "two-quantity at a time" reasoning involving the first law. In particular, many interviewed students only mentioned two of the three quantities involved in the first law as in the following response from an introductory student after instruction in relevant concepts: "Conservation of Energy when there's no work being done on a system." Even after further prompting, this introductory physics student omitted Q from the discussion and continued to claim that first law says that "energy" does not change when no work is done on a system. Another introductory student who ignored the changes in the internal energy of the system stated: "I still don't know how you would connect the work and the heat done." When prompted further about what else the student could say about the first law, the student added: "I know that all the heat that you have in a system can't go into doing work for the system." This type of statement shows confusion between the first and second laws.

Based upon the bounded rationality and knowledge in pieces frameworks, it is not surprising that students are using this type of "two-quantity at a time" reasoning. It can help them manage their cognitive load during problem solving and make the problem solving and reasoning process more manageable even though they are likely to arrive at an incorrect conclusion having left out one of the three quantities in the first law.

5.7.2 Inappropriate Use of a Compensation Heuristic

Many items on the survey have an explicit description of a process and students must use the first law of thermodynamics to reason about a target quantity (W, Q, or ΔE) using the information inferred from the description of the problem situation. As noted earlier, in these types of problems, students often ignored a quantity, especially if it is not explicitly mentioned in the problem. Moreover, interviews suggest that in some cases, even if all the three quantities involved in the first law are mentioned in the problem in some form or another (e.g., items 21 and 27), some students use a compensation heuristic to determine the target variable using only two of the three quantities. The students with this type of reasoning generally claim that an increase in one quantity must lead to a decrease in another quantity and vice versa. The various cues in the problem determine how a particular student may proceed, which quantity the student focuses on first and then which quantity the student makes an inference about based upon the compensation heuristic.

While it is not possible to obtain exact percentages of the use of compensation heuristics in written responses, Table 16 provides an estimate of the prevalence of the responses in which an increase in one quantity is paired with a decrease in another quantity incorrectly.

5.7.2.1 Difficulties in Situations in Which the Relevant Quantities Are Not Zero

On item 9, both processes involve an expansion in which positive work is done by the system. The problem asks students to make a comparison between W_1 and W_2 and between Q_1 and Q_2, which are the work done by the system W and heat transfer to the system Q for two different processes. For this problem, students are provided both a PV diagram (Figure 8)

Table 16: Prevalence of Incorrect Use of Compensation Heuristic in First Law Problems

Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
An increase in one quantity implies a decrease in another quantity.	*2, 9, 10, 14**	21	23	22	20	17
Two of the three quantities in the first law are zero in a process.	*17, 21, 27, 29**	16	28	34	30	30

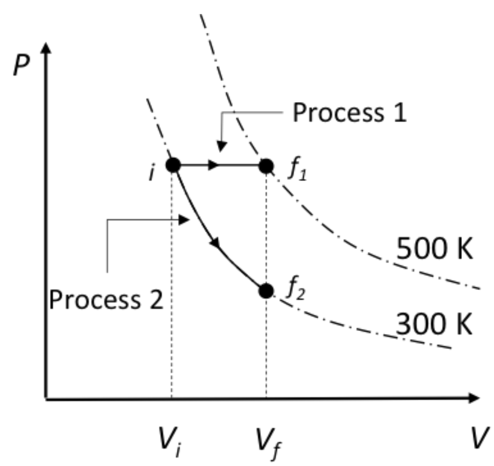


Figure 8: Example PV Diagram from STPFaSL from item 9.

and a verbal description.

In order to solve this problem, a student must make use of the first law correctly. However, interviews suggest that after determining a relation between one pair (e.g., inferring that $W_1 > W_2$ correctly by discerning the areas under the curves on the PV diagram), some students use a compensation heuristic instead of using the first law to determine the remaining relationship. For example, in this case, they may incorrectly conclude that since $W_1 > W_2$, then $Q_1 < Q_2$ to compensate for it. On the other hand, some students who first correctly reasoned that $Q_1 > Q_2$ likely used a compensation heuristic to incorrectly infer that $W_1 < W_2$ (this latter group of students generally did not know how to infer the work done by the system using the area under the curve in a PV diagram). Thus, there is a pattern to their reasoning.

Another situation in which some interviewed students used a compensation heuristic is item 10, which is one version of the Meltzer [67] problem, in which two expansion processes start and end at the same initial and final states on a PV diagram. The most common incorrect response from upper-level students was $Q_1 < Q_2$ since $W_1 > W_2$. Such a response was also quite common among introductory students. Similarly, on item 2, which asks about the change in the internal energy and work done by the gas in an adiabatic expansion, some interviewed students used a compensation heuristic to incorrectly reason that the internal energy of the system must increase and the work done by the gas must be negative after they decided on the increase or decrease in one of those variables incorrectly.

5.7.2.2 If One Quantity Is Zero, Another Quantity Is Zero A special case of the compensation heuristic, which was as prevalent in interviews as the cases in which students reasoned that a higher value of one quantity must be compensated by a lower value of another quantity in a process, is one in which students conclude that if one thermodynamic quantity is zero, then another thermodynamic quantity is zero as well (see Table 6). This difficulty was often apparent in the interviews and individual discussions when students reasoned about some of the conceptual problems that involve the first law. For example, on both items 17 and 21, one common difficulty was concluding that in an isochoric process, since there is no change in the volume of the system and therefore no work is done, and hence there is no change in the internal energy of the system (we note that some students directly concluded that no change in the volume of the system implies no change in the internal energy even without reasoning about the fact that $W = 0$ in an isochoric process). Many of these students in one on one discussion did not realize that they should also be considering the net heat transfer to the system in this process and if two of the variables (internal energy of the system and work done by the system) are zero then the net heat transfer will also be zero in the isochoric process as described by them. In fact, on question 21, almost all of the interviewed students with this type of reasoning thought that only the work done by the gas and the change in the internal energy of the system are zero but the net heat transfer between the system and the surroundings is not zero. Even when some of them were explicitly asked about the net heat transfer in the isochoric process, they did not know how to reason about it and some did not seem to be very concerned about the fact that the first law of thermodynamics is a relation between three quantities in which only two of the three quantities were zero according to their reasoning. Similarly, on question 21, another very common difficulty was that in the isochoric process, there is no work done by the system and there is no net heat transfer between the system and the surroundings but there is a change in the internal energy of the system. In individual discussions, many of these students also used the reasoning that if the work done by the system is zero then the net heat transfer should be zero as well and they generally did not remember that the first law is a relation between three quantities in which only two of the three quantities cannot be zero for a process. We note that 52% of introductory students in an algebra-based course pre-instruction noted an energy-conservation principle as the first law. However, the use of a compensation heuristic concluding that two quantities are zero in a process suggests that many students do not have a solid grasp of the first law.

We note that on item 27 related to an isothermal expansion of an ideal monatomic gas, one common incorrect response is that in an isothermal process, there is no heat transfer between the system and surroundings, the internal energy does not change but the work done by the gas is positive. Unlike the responses of students who claimed that the internal energy must change in this isothermal process because $Q = 0$, the type of response in which students claim that $Q = 0$ and $\Delta E = 0$ in an isothermal process shows a lack of understanding of the first law of thermodynamics as a relation between three thermodynamic quantities. If the students knew how to reason about the first law correctly using all the three quantities involved, they would have realized that it is impossible to have a situation in which two of the quantities in the first law are zero while the third quantity is non-zero.

5.7.3 Inferring an Incorrect Sign of a Physical Quantity Using the First Law

When students use an algorithmic approach and attempt formula fitting to solve problems that can be solved using the first law, they often get the sign of a physical quantity wrong. For example, on item 2 related to a reversible adiabatic expansion, the internal energy of the system must decrease in the expansion. However, Table 13 shows that the most common incorrect answer among students at all levels was that the internal energy should increase and the work done by the system should also be positive (correct) in this adiabatic expansion. Interviews suggest that students with this kind of reasoning often thought that since the change in the internal energy of the system is equal to the work done by the gas in this case, both changes must be positive. These students did not think conceptually about the fact that the work done by the gas should decrease the internal energy of the system in this adiabatic expansion. They often focused on an algorithmic approach and claimed that both the left hand side and right hand side of the formula they were using should have the same sign.

Another common incorrect response for item 2 from students at all levels is that the internal energy of the gas must increase and the work done by the gas must be negative (both signs incorrect). In individual interviews and discussions, many students did not think conceptually about the work done by the gas and they had difficulty in distinguishing between the work done by the gas and the work done on the gas. In particular, many students in both introductory and upper-level courses could not reason conceptually that the work done by the gas will decrease the internal energy of the gas and the work done on the gas will increase the internal energy of the gas. These interviewed students, who did not reason conceptually about these issues and made use of some equation form of the first law of thermodynamics (correct or incorrect form), were very likely to get the sign of the work done by the gas wrong. The difficulty in conceptual reasoning about the work done by the system can also lead students to incorrectly answer questions related to heat transfer between the system and the surroundings because these two quantities are intertwined in many problems through the conservation of energy (for example, in items 3, 8, 10, 15, 27 for the relation between work and heat transfer).

Interviews suggest that on item 3, which is about a reversible isothermal compression of a monatomic ideal gas, some students made an incorrect inference about the sign of Q (whether net heat transfer is to or from the system) using the first law often because they were not thinking conceptually about the process (see Table 15). Even those students who correctly reasoned that $\Delta E = 0$ for the process did not realize that in a compression, work is done on the system by the surroundings so to keep the internal energy constant there must be net heat transfer from the gas.

On item 27, some interviewed students incorrectly noted that there is no heat transfer between the system and the surroundings for an isothermal process. However, some of them used a correct chain of reasoning to conclude that if the work done by the gas is positive and there is no heat transfer, the internal energy of the gas must decrease. The students who stated that the internal energy decreases essentially provided the correct reasoning using the first law for an adiabatic expansion (their mistake was assuming that a constant temperature process implies there is no heat transfer between the system and the surroundings). On the other hand, some interviewed students who did not reason conceptually and made use of the first law equation, incorrectly inferred that the work done by the gas should be positive and the internal energy of the gas should increase as well since there is no heat transfer between the system and surroundings for the isothermal process in item 27.

These types of conceptual difficulties may at least partly be responsible for student difficulties with the relation between work, heat transfer, and internal energy changes in items 9, 21, 29, 30.

5.7.4 Using the First Law as the First Line of Reasoning When Other Reasoning Should Precede It

Similar to previous findings of expertise research, we find that many students immediately focus on whatever the question is asking about instead of making an overall plan to systematically approach the problem solving process [107]. In many of the problems, reasoning using the first law of thermodynamics is not the first line of reasoning students should use. However, in many situations, some students often invoke the first law right away when they are asked for a relation between any combination of E, W and Q. For example, on item 10, even before invoking the first law of thermodynamics, an expert would first focus on the fact that both processes depicted on the PV diagram have the same initial and final states and conclude that the change in internal energy of the system is the same for both processes. Next, the expert would note that the work done by the system is larger for process 1 than for process 2 (as given by the area under the curve for each process on the PV diagram). Then, they would invoke the first law of thermodynamics to reason that from the conservation of energy, there should be larger net heat transfer to the system in process 1 than in process 2 if both systems start and end in the same states. Individual interviews and discussions suggest that some students immediately start thinking about the net heat transfer to the system using some version of the first law (not necessarily correct) and do not reason systematically about the fact that the changes in the internal energy are the same in both processes and work done by the system is larger in process 1 before invoking the first law.

5.8 DIFFICULTIES WITH THE SECOND LAW OF THERMODYNAMICS

Previous investigations have shed light on student difficulties with the second law [63–66, 76, 77, 79, 87–89]. Many of the findings in our investigation are consistent with the earlier findings. Interviews and written responses suggest that many students do not remember the various forms of the second law of thermodynamics correctly (Clausius statement, Kelvin-Planck statement, and the change in the entropy of the universe should be greater than or equal to zero in any thermodynamic process, etc.) [90]. One common difficulty was that students often overgeneralized the second law. For example, in individual interviews, many students after instruction echoed the statement of this introductory physics student: "Second law says that the entropy always increases". This type of common overgeneralization of the second law suggests that many students do not have a clear understanding of this form of the statement of the second law and they are confusing the system with the universe and/or are unable to distinguish between reversible and irreversible processes.

On the brief survey of thermodynamics given before instruction on relevant topics for bonus points on an exam in an algebra-based introductory physics course, students were less likely to make a correct or incorrect statement about entropy and the second law (28%) than they were to make statements about the first law (52%). Even in these responses from students before instruction, the most common statements about the second law were the well-known overgeneralizations of the law as conveyed by the following statements from students: "Entropy increases always (and/or everywhere)" or "Energy is always moving toward disarray" (the latter statement suggests confusion between energy and entropy). Examples of the prevalence of the difficulties with the second law can be found in the written responses to the survey questions.

Since many of the survey problems that involved the second law were framed in terms of entropy, difficulties with those problems were often associated with difficulties with the concept of entropy. Many prior investigations have focused on student understanding of entropy [63, 65, 66, 71, 77, 79, 86–89]. Many of our findings are consistent with earlier studies. Confusion about the system and the universe is often responsible for student difficulties with the second law as it relates to entropy. Some students claimed, e.g., that the entropy of the system (not the universe) increases for irreversible but stays constant for reversible processes. Another common difficulty with the second law as it relates to entropy is that the entropy cannot change for an isolated system. Students often incorrectly claimed that even in an irreversible process, the entropy of an isolated system does not change. For reversible processes, some students claimed that entropy is conserved for the system (or both the system separately and the universe). On the other hand, some students claimed that the entropy of the universe increases even in a reversible process since they did not know the relevance of the word "reversible" in this context. These difficulties will be discussed in more detail in the following sections.

5.8.1 Difficulties With Heat Engines

Several prior investigations have focused on student understanding of heat engines and some have also focused on learning tools that can help students develop a coherent understanding of heat engines [63–65]. We find that most students in all groups had great difficulty with the two problems about heat engines on the survey and many answered them by resorting to memorized information about heat engines. On item 16, students had to recognize the correctness of the following two versions of the second law that were related to heat engines (directly or indirectly): "No heat engine can be more efficient than a reversible heat engine between given high temperature and low temperature reservoirs." and "The net heat transfer to the system from a hot reservoir cannot be completely converted to mechanical work in a cyclic process". Students had great difficulty with these. Interviews suggest that some students did not have a clear understanding that in heat engines, as the cyclic process is repeated, there is net heat transfer to the working substance (system) from the hot reservoir, some of this energy is used to perform work by the system and the rest transfers to the cold reservoir in each cycle. All of the difficulties discussed earlier in the context of a cyclic process were observed in interviews (e.g., not recognizing that $\Delta S = 0$ for both the system and the environment in a complete cycle if the process is reversible but $\Delta S = 0$ only for the system and that $\Delta S > 0$ for the environment so that the entropy of the universe increases consistent with the second law of thermodynamics for an irreversible complete cycle). Some interviewed students had difficulty with articulating how the operation of the heat engine involves both the first and second laws of thermodynamics consistent with earlier findings.

5.8.2 Difficulties With ΔS in a Reversible Process

The common difficulties with the concept of reversibility of a process can be classified into two categories. One difficulty was overgeneralizing the statement that reversibility implies that $\Delta S = 0$ for the universe to reversibility implies $\Delta S = 0$ for the system. The second

Table 17: Prevalence of Difficulties with Entropy change in Reversible Processes

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Reversible Adiabatic Process						
$\Rightarrow \mathbf{\Delta S_{system} = 0}$ ***(correct)***	***1***	**24**	**28**	**19**	**29**	**24**
Reversibility (alone) $\Rightarrow \Delta S_{system} = 0$	*1*	39	29	16	36	19
$\Delta S_{system} > 0$	*1*	28	32	51	29	46
Reversible Isothermal Compression						
$\mathbf{\Delta S_{system} < 0}$ ***(correct)***	***4***	**31**	**28**	**37**	**32**	**44**
$\Delta S_{system} = 0$	*4*	42	36	25	43	20
$\Delta S_{system} > 0$	*4*	23	26	23	12	19
Reversible Isothermal Expansion						
$\mathbf{\Delta S_{System} > 0}$ ***and*** $\mathbf{\Delta S_U = 0}$ ***(correct)***	***28***	**34**	**40**	**36**	**42**	**27**
$[\Delta S_{System} = 0]$ and $\Delta S_U = 0$	*28*	26	27	23	30	33
$\Delta S_{system} > 0$ and $\Delta S_U > 0$	*28*	40	33	41	28	40

difficulty was to disregard the word "reversible" in a question and let the other details of the problem dictate the reasoning in the problem (these students often claimed that entropy of the system and/or the universe increases in a reversible process). The analysis of the survey data suggests that students sometimes performed better on problems involving irreversible processes than they did on problems involving reversible processes partly because of these two difficulties.

Table 17 shows the prevalence of student difficulties with entropy in a reversible process across various items on the survey. The table shows that one of the most common difficulties on item 1 about a reversible adiabatic expansion of a gas is that the entropy of the system remains constant for a reversible process regardless of whether the process is adiabatic or not. Table 17 shows that the entropy of the system remains constant for a reversible process is one of the most common difficulties also on item 4 (isothermal reversible compression) and item 28 (in item 28, a gas undergoes an isothermal reversible expansion). As noted, this type of response is often due to the confusion between the system and the universe.

Table 17 also shows that another common conceptual difficulty with items 1 and 4 is that the entropy of the system increases in the reversible processes described in the problems but the entropy of the system does not change in item 1 and it decreases in item 4. Interviews suggest that students with this type of a response either incorrectly thought that the entropy of the system increases because it is an expansion process or that the entropy of the system increases in all processes.

On the other hand, on item 28, a common incorrect response was that the entropy of the universe increases (see Table 17). Interviews suggest that students with this type of difficulty often did not recognize the relevance of the word reversible and assumed that entropy increases in all processes.

5.8.3 Difficulties With ΔS in an Isolated System

Some prior studies have investigated student difficulties with an isolated system [63, 87]. Our findings are consistent with these studies. A number of the survey questions contain a system (sometimes consisting of two subsystems) which cannot exchange energy in any form with the surroundings (e.g, items 11, 12, 14, 16, 23, 24, 13, and 32 all require reasoning about isolated systems). The entropy of an isolated system increases if it undergoes an irreversible process. However, questions related to isolated systems for which neither particles nor energy are being exchanged with the surroundings, some students thought that the entropy of the system cannot change. Table 18 shows student difficulty with such questions. For example, on item 16, more than 80% of the students before instruction and 73% of students after instruction in introductory courses claimed that the entropy of any system that cannot exchange energy with its surrounding environment must remain unchanged (see Table 18). In upper-level courses, the corresponding average percentage is 42%. Some interviewed students with this difficulty claimed that if there is no exchange of energy, there is no heat transfer between the system and the surroundings so there cannot be any change in the entropy of the system since the change in entropy depends on Q. Also, some interviewed students explicitly tried to consider a process in which there is no energy exchange between the system and surroundings and the entropy of the system increases, but they often could not come up with any appropriate process. Those who did make a connection generally considered ΔS in the free expansion process in which the entire system is insulated from the surroundings. Those students usually concluded that since the entropy of the isolated system increases in the free expansion process, ΔS can be greater than zero for an isolated system.

5.8.4 Difficulties With ΔS in an Irreversible Isochoric Process

Item 29 is the only problem on the survey with the word "irreversible" in it. In this irreversible isochoric process, the most common difficulty was that the entropy of the gas increases but the entropy of the system and reservoir taken together does not change (see Table 18). Interviews suggest that some students with this type of reasoning thought that the entropy of the system and the reservoir taken together cannot change because the gas was in a closed container. The second most common difficulty was that the entropy is conserved both for the system alone and also for the system and the reservoir taken together. Interviews suggest that some of these students thought that the entropy of the gas cannot increase when it is in a closed container.

5.8.5 Difficulties With ΔS for a Free Expansion Process

Table 18 shows that on item 14 on the survey, the most common difficulty related to entropy is that it is conserved for the free expansion process. Approximately 20-30% of introductory students from calculus-based and algebra-based courses even after instruction and 10% of students from upper-level courses selected this option. Interviews suggest that a common reason for this difficulty is assuming that entropy cannot change for an isolated system.

5.8.6 Difficulties With ΔS in Spontaneous Mixing of Two Gases

Table 18 shows that on item 32 the most common difficulty was that the entropy of the combined system increases if two inert ideal gases in adjacent chambers are allowed to mix spontaneously but the entropy of each of the sub-systems does not change (the combined system is in an insulated case). Interviews suggest that at least for some of the students, this type of difficulty is due to combinatorics and students were unable to reason that there

Table 18: Prevalence of Difficulties with Entropy Change in Irreversible Processes

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Not all processes have $\mathbf{\Delta S_{Isolated} = 0}$ (correct)	***16***	**58**	**27**	**19**	**27**	**16**
$\Delta S_{Isolated} = 0$ for all processes	*16*	42	73	81	73	84
Irreversible Isochoric Process						
$\mathbf{\Delta S_U > 0, \Delta S_{System} > 0.}$ (correct)	***29***	**68**	**43**	**35**	**57**	**40**
$\Delta S_U = 0$, $\Delta S_{System} > 0$.	*29*	16	31	30	20	28
$\Delta S_U = 0$, $[\Delta S_{System} = 0]$	*29*	9	8	15	10	17
$\mathbf{\Delta S_{Isolated} > 0}$ for a spontaneous process (correct)	***12, 14, 24, 32***	**84**	**61**	**56**	**62**	**59**
$[\Delta S_{Isolated} = 0]$ for a spontaneous process	*12, 14, 24, 32*	16	39	44	38	41
Free Expansion Process						
$\mathbf{\Delta S_{Isolated} > 0}$ (correct)	***14***	**89**	**77**	**69**	**79**	**72**
$\Delta S_{Isolated} = 0$	*14*	11	23	31	21	28
Spontaneous Heat Transfer						
$\mathbf{\Delta S_{Isolated} > 0}$ and $\mathbf{\Delta S_{Cold} > 0}$ *(correct)*	***12, 24***	**63**	**31**	**24**	**32**	**30**
$[\Delta S_{Isolated} = 0]$ and $\mathbf{\Delta S_{Cold} > 0}$	*12, 24*	16	41	42	41	41
$\mathbf{\Delta S_{Isolated} > 0}$ and $[\Delta S_{Cold} = 0]$	*12, 24*	6	9	13	11	11
$\mathbf{\Delta S_{Isolated} > 0}$ and $\mathbf{\Delta S_{Cold} > 0}$ but also $\Delta S_{Hot} > 0$	*12, 24*	14	13	13	11	12
Spontaneous Mixing						
$\mathbf{\Delta S_{Isolated} > 0}$, and each $\mathbf{\Delta S_{subsys} > 0}$	***32***	**62**	**37**	**33**	**40**	**27**
$\Delta S_{Isolated} > 0$, but each $[\Delta S_{subsys} = 0]$	*32*	21	23	21	22	31
$[\Delta S_{Isolated} = 0]$, and each $[\Delta S_{subsys} = 0]$	*32*	8	21	23	18	16
$[\Delta S_{Isolated} = 0]$, but each $\Delta S_{subsys} > 0$	*32*	6	13	15	15	13
Second law *(correct)*	***13, 25, 33***	**88**	**63**	**58**	**72**	**53**
First law	*13, 25, 33*	25	53	60	40	54
Newton's 2^{nd} law	*13, 25, 33*	4	13	17	11	25

are more states available to the system if each gas is allowed to mix with the other gas in the adjacent chamber. It is reasonable to expect that at least advanced students have learned combinatorics to be able to reason about these types of issues. Table 18 also shows that the second most common difficulty on item 32 is that the entropy of the combined system and the entropy of each of the sub-systems does not change (each is conserved) if two inert ideal gases in adjacent chambers are allowed to mix spontaneously. Interviews suggest that at least for some of the students, this difficulty is due to the fact that the combined system is in an insulated case. Table 18 also shows that the third most common difficulty on item 32 is that the entropy of the combined system does not change (is conserved) but the entropy of each of the sub-systems increases. Interviews suggest that some students with this type of reasoning thought that since each gas is allowed to mix, each component will have an increased entropy but since the combined system was in an insulated case, there is no change in the entropy of the combined system.

5.8.7 Difficulties With ΔS in Spontaneous Heat Transfer Between Two Solids at Different Temperatures

Table 18 shows that on item 12, the most common difficulty was that for a closed system, the entropy of the sub-system at a lower temperature increases but that of the combined system does not change if there is spontaneous heat transfer from one part of the system to another when two solids at different temperatures are placed in contact. Interviews suggest that some students with this difficulty thought that the entropy of the combined system cannot change for an isolated system. Table 18 also shows that on item 12, the second most common difficulty was that for a closed system, the entropy of each sub-system and the combined system all increase if there is spontaneous heat transfer from one part of the system to another when two solids at different temperatures are placed in contact. Interviews suggest that at least some of these students thought that the entropy always increases for everything at least in an irreversible process. In addition, Table 18 shows that on item 12, the third most common difficulty was that the entropy of the combined system increases but the entropy of the cold sub-system does not change when two solids at different temperatures are in contact. Interviews suggest that at least some of these students thought that the second law of thermodynamics says that the entropy of the combined system increases but the entropy of each sub-system does not change in such processes.

5.8.8 Difficulties With ΔS in Spontaneous Heat Transfer Between Two Gases at Different Temperatures

Table 18 shows that on item 24, the most common difficulty was that for a closed system, the entropy of the sub-system at a lower temperature increases but that of the combined system does not change if there is spontaneous heat transfer from one part of the system to another. Table 18 also shows that on item 24, the second most common difficulty was that for a closed system, the entropy of each of the sub-system and the combined system all increase if there is spontaneous heat transfer from one part of the system to another. Table 18 shows that on item 24, the third most common difficulty was that for a closed system, the entropy of the combined system increases if there is spontaneous heat transfer from one part of the system to another but the entropy of each sub-system does not change. Interviews suggest that the common reasons for these difficulties were analogous to those for spontaneous heat transfer between two solids at different temperatures in contact that was discussed in the preceding section.

5.8.9 Difficulties With the Relation Between Entropy Change, Heat Transfer, and Temperature

Several prior investigations have focused on student difficulties with the relation between entropy, heat transfer and temperature [63, 65, 77, 79, 87]. Our findings are consistent with these studies. While some of the interviewed students did not recall the relationship between heat transfer to the system and the change in entropy of the system, even those interviewed students who remembered (often vaguely) that the change in the entropy of the system is somehow related to the net heat transfer between the system and the surroundings, generally appeared not to be able to exploit this information effectively to solve the survey problems. In fact, some of the students who remembered the relation between entropy and net heat transfer noted that if there is no heat transfer between the system and the surroundings, the entropy of the system must be conserved (including for irreversible processes). They argued that if $Q = 0$, then the entropy of the system cannot change. Also, some interviewed students did not know that the relation between the change in entropy, heat transfer and temperature applies only for a reversible process, but since entropy is a state variable, in order to find the change in entropy for an irreversible process, one can carry out the integral over a reversible process that takes the system from the same initial state to the final state.

5.8.10 Difficulty With the Law Governing the Direction of Spontaneous Processes

Prior research indicates that students do not always connect the second law of thermodynamics with spontaneous processes that occur in nature [63, 77, 87–89]. Items 13, 25, and 33 on the survey ask students to choose why a spontaneous process occurs in a certain direction and which law(s) will be violated if a spontaneous process took place in the reverse direction. On these questions, students had to identify that the second law of thermodynamics will be violated if a spontaneous process took place in the opposite direction. Items 13 and 25 pertaining to situations in which there is net heat transfer from a solid at a higher temperature to a solid at a lower temperature in contact or from a gas at a higher temperature to a gas at a lower temperature in adjacent chambers. Question 33 focused on mixing of two inert ideal gases in adjacent chambers. Apart from the second law of thermodynamics, students could select the first law and/or Newton's second law. These laws were chosen based upon our earlier interviews that suggested that some students thought that the first law of thermodynamics prevents a process from happening in reverse. Some students were even concerned whether Newton's second law would be violated in these types of reverse processes (especially in the mixing process). Table 18 shows the percentages of students who selected different options. Some interviewed students explicitly noted that since the first law relates to the conservation of energy, it will always be violated if a spontaneous process took place in a direction opposite to that observed.

In the interviews, some students claimed that the reason the first law (energy conservation) or Newton's second law is violated is because the momentum conservation is violated in these types of processes. In some cases, students reasoned about heat transfer as though it involved transfer of particles from one region to another. For example, in problems 13 and 25, some students stated that if the process took place in the opposite direction, the particles at lower temperature or lower energy will move towards particles with higher energy (at a higher temperature) and then bounce back off of them and this collision process imparts to the particles an overall higher momentum than was available originally which is not possible. In interviews, some students spent a significant amount of time trying to determine if these processes violate Newton's second law. On a positive note, introductory students taking the survey as a pretest often thought that Newton's second law is violated in such processes, but the fraction of students who thought this by the end of the semester is markedly lower.

5.9 DIFFICULTIES RELATED TO DISTINCTION BETWEEN SYSTEMS AND UNIVERSE

Most of our findings are consistent with prior research studies that have investigated the confusion between the system and the universe [63, 64, 87, 88]. In order to solve problems about changes in the entropy of the system or the universe using the second law of thermodynamics, one must understand the law as it pertains to the entropy change of the universe in a particular process. For example, students should understand that one version of the second law of thermodynamics states that the entropy of the universe increases in an irreversible process and remains constant in a reversible process, and it is possible for the entropy of the system alone to decrease or increase in a particular thermodynamic process regardless of whether the process is reversible or irreversible. Many students had difficulty with the survey problems related to entropy because they could not make a clear distinction between the system and the universe and how the second law formulated in terms of entropy applies to various systems, subsystems and the universe.

Some interviewed students also treated cyclic processes and reversible processes in a very similar manner in some problems (not necessarily all since student reasoning is context dependent) as though the word "cyclic" has the same meaning as the word "reversible". They did not realize that a cyclic process can be irreversible and in such a process the change in the entropy of the system will be zero but the change in the entropy of the environment will be greater than zero because it is the system that returns to the initial state after a complete cycle but the final state of the environment is different from the initial state after a complete cycle. Thus, the entropy of the universe increases in an irreversible cyclic process. The fact that some students are unclear about the distinction between the system, environment and the universe makes these concepts challenging for them.

5.10 DIFFICULTY MAKING USE OF THE PV DIAGRAM REPRESENTATION

There have been several prior investigations involving student understanding of PV diagrams [63, 67, 78, 83] There were eight problems out of 33 items on the survey on which PV diagrams were used in addition to stating the problems using a verbal description. The role of diverse representations, e.g., verbal, diagrammatic, mathematical, etc. in physics problem solving has been studied extensively [20, 25, 27, 36, 91–98]; studies focusing on expertise suggest that experts are facile in translating from one representation of knowledge to another. Experts often convert a problem from one representation to another to facilitate their problem solving process and make the problem solving task easier. On the other hand, students who are developing expertise not only have great difficulty in converting a problem from one representation to another, they also have difficulty in making sense of the information in various representations and exploiting it to solve problems.

When presented with a problem in different representations, students are likely to perform better in the representation with which they are more familiar or comfortable. The diagrammatic representation is often a favored representation for experts in the problem solving and reasoning process. One reasonable question to ask is whether students perform better on the survey items in which a PV diagram is provided in addition to the verbal description than on other problems with the same underlying physics principles but with only the verbal description of the problem. Our analysis suggests that students did not in general benefit from having the PV diagram. Interviews suggest that many students do not know how to take advantage of such diagrams to reason about the problems. However, one difference we found in interviews and individual discussions with students is that, while a majority of introductory students in both calculus-based and algebra-based courses had great difficulty reasoning using a PV diagram, many upper-level students attempted to reason with it. This observation does not mean that upper-level students were very successful in reasoning with the PV diagram. In fact, some upper-level students even tried to draw PV diagrams for processes for which a diagram was not provided (e.g., item 28), but they often had difficulty in actually representing the processes on the PV diagram except when graduate students worked in pairs and reasoned about the PV diagrams as discussed in the companion paper. The student pairs were more likely to be successful in using and interpreting the PV diagram correctly or drawing it for problems in which a diagram was not provided.

5.11 ARE WE PRIMING STUDENTS BY PLACING SURVEY QUESTIONS ON SIMILAR TOPICS CONSECUTIVELY?

On the survey, some consecutive questions are on similar topics and one may wonder whether students would be primed by having such questions one after the other. For example, items 9 and 10 both involve reasoning about the relation between the work done by the gas and the net heat transfer to the system. Similarly, items 26 and 27 both involve work done by the gas in an isothermal process. Although experts who provided feedback on the survey noticed that some questions that were placed one after the other could prime students to solve the second problem correctly, unfortunately, none of the students interviewed ever explicitly mentioned the similarity or relation between consecutive questions and whether placing them one after the other benefited them in any way.

5.12 INFORMATION THAT SIMPLIFIES EXPERT THINKING CAN DERAIL STUDENT REASONING

Certain information in a problem may give clarity to experts and simplify their reasoning process while solving a problem. For example, if the problem statement says that a thermodynamic process is reversible and the problem is about changes to the entropy of the system or the universe, such information made it easier for experts to reason that the entropy of the system can increase or decrease but the entropy of the universe will not change. On the other hand, many students who had alternative conceptions about what the word reversible meant, concluded that the entropy of the system will also not change. Similarly, on problems related to cyclic processes, experts immediately note that the state variables will be unchanged but other physical quantities are not zero in general after a complete cycle. However, the phrase "cyclic process" led many students to conclude that none of the thermodynamic quantities changed in such a process after a complete cycle. Thus, the information that the process is reversible or cyclic completely derailed their reasoning. The same kind of difficulty was observed on survey problems that used the word isothermal. As noted, in interviews, students often interpreted the word isothermal to mean that either both the temperature was constant and net heat transfer between the system and surroundings was zero or at least $Q = 0$ even if they did not explicitly mention anything about the change in temperature in an isothermal process. The rest of their reasoning was impacted by this initial interpretation of an isothermal process.

5.13 CONCLUDING REMARKS

We find that many students in all groups did not have the expertise to solve the conceptual survey problems in an expert-like manner and they did not invoke appropriate concepts in the order in which they should be invoked to solve a problem. Unlike experts, constraints on a process were often ignored or used incorrectly by students. Students' responses were also often context dependent in that they did not give consistent answers to various problems involving the same underlying concepts. In the framework of bounded rationality, they selected an answer to each problem that "satisfices" [107]. Since students' knowledge chunks are smaller, they may be overloaded by the number of potentially relevant concepts they must keep in working memory to reason about each survey problem methodically [107–109]. On the other hand, since experts have a good knowledge hierarchy and their knowledge chunks are larger, they can manage cognitive load while solving problems and systematically focus on various concepts involved in each problem during the problem solving process. For instance, issues related to whether the problem can be solved using the first law or second law (or both), whether a process is reversible or irreversible (and whether reversibility is relevant for solving the problem), whether the process described is isothermal, isobaric, isochoric, or adiabatic, and how to apply these constraints to solve the problem.

5.13.1 Recommendations for Instructors

In order to help students learn to make sense of the thermodynamic processes and correctly apply the first and second laws of thermodynamics, instructors must use physics education research based methods to help students learn content along with developing students' problem solving and reasoning skills. Most thermodynamics problems on the survey cannot be solved using a plug and chug approach with memorized algorithms. We find that students often used heuristics that were too simplistic such as using only two quantities in the first law instead of reasoning with all the three quantities or using a compensation heuristic to solve problems involving the relation between the internal energy of the system, work done by the system and net heat transfer to or from the system. They often claimed that when some quantity increases, another must decrease or if some quantity does not change then another should not change either without considering other aspects of the problem correctly. Interviews suggest that sometimes, the same student reasoned about two of the three quantities in the first law, choosing a set of two out of three quantities depending upon the cues of the problem statement without realizing that he/she may have skipped a quantity that should be included in the reasoning involving conservation of energy. Expertise research suggests that such a problem solving strategy is typical of students who do not have a robust knowledge structure in the domain [107–109].

In order to apply the fundamental laws of thermodynamics appropriately in diverse situations, students must discern the differences between different types of thermodynamic systems, quantities, processes, etc. The variety of situations one encounters in relation to the foundational topics in thermodynamics covered in the survey can provide a great opportunity for developing students' problem solving, reasoning and meta-cognitive skills. The probability of students performing well on a conceptual survey like the one discussed here using an algorithmic problem solving approach is very low. In order to solve these conceptual problems, students must learn how to systematically focus on different information in the problem statement rather than dealing with it haphazardly and only taking into account some of the information, which invariably leads to incorrect inference.

For example, on problem 1 about the entropy of the gas in a reversible adiabatic expansion, the first thing students must reason about is that in a reversible process, the entropy of the universe remains constant. Then, they must reason that since the given process is adiabatic, there is no net heat transfer to or from the system. Thus, entropy of the system remains constant for this reversible process because the process is adiabatic. Students should realize that the answer to this question would be the same regardless of whether it was a reversible adiabatic expansion or compression. They should also be able to reason that if the process was not adiabatic but it was a reversible process, the entropy of a system can increase or decrease but the change in the entropy of the surroundings should be such that the entropy of the universe does not change. They should also learn to argue that if a process was irreversible, the entropy of the universe would increase (which is due to the overall changes in the entropy of the system and the surroundings). In other words, even in a seemingly simple conceptual problem, there is a systematic reasoning chain required to solve the problem correctly. If students do not understand the relevance of reversibility in determining the change in the entropy of the universe, the difference between the system and the universe, or do not know how to take into account the fact that the process is adiabatic, they are unlikely to solve the problem correctly.

On many problems, students were unable to distinguish between the system, environment and the universe. Instruction should be designed carefully to help students learn, e.g., that in a cyclic process since the initial and final states of the system are the same in a complete cycle, there is no change in any state variable including the entropy for the system. On the other hand, all we can say about the environment is that there is no change in the internal energy of the environment in the complete cycle due to the conservation of energy, but the initial and final entropy of the environment will be the same only if the cyclic process is reversible. In particular, if the cyclic process is irreversible, the final entropy of the environment must be greater than its initial entropy so that the entropy of the universe (which is the sum of the entropy of the system and environment) increases, consistent with the second law of thermodynamics. Helping students learn to treat changes to the system, environment and the universe due to various thermodynamic processes instead of mainly focusing on discussions of the changes to the system will not only help students apply the second law of thermodynamics correctly in various situations but also reduce the confusion between, e.g., a cyclic and a reversible process.

We found that many students at all levels confused "closed container" with an "isolated system" in which there is no heat transfer between the system and the surroundings. Interviews suggest that some of the worst performance on the survey questions was on items in which a gas in a closed container was heated because students treated such a system as an isolated system. It is possible that this difficulty will decrease if instructors do not use the terms "isolated system" and "closed system" interchangeably because this type of language can give students the impression that "closed container" and "isolated system" are the same.

We find that many students do not realize or take advantage of the richness of the PV diagrams partly because they have not been provided sufficient opportunity and scaffolding support to learn to reason with the PV diagrams systematically. Helping students learn to make sense of PV diagrams for various types of processes and the relations among quantities may provide an excellent opportunity to help them learn the role of representation and its value in the problem solving process. Instruction should focus on helping students do sense making and learn to reason with PV diagrams for various processes. It would be helpful for students to reason how a PV diagram may even provide redundancy in the amount of information given so that one can reason about a particular outcome in a given situation in more than one way.

Finally, instructors should know the common difficulties that students have. For example, students commonly confuse the free energy and internal energy, *even if the instructor does not introduce free energy* in a physics course. An instructor can design instruction appropriately to help students avoid these types of pitfalls only if they are equipped with knowledge of these common difficulties.

6.0 CONCLUSION AND FUTURE DIRECTIONS

6.1 FUTURE EXTENSIONS OF QUANTUM MECHANICS RESULTS

This thesis explores issues related to common student difficulties with some topics in undergraduate quantum mechanics and thermodynamics courses. The quantum mechanics research focuses on using the research on student difficulties for the development and evaluation of a Quantum Interactive Learning Tutorial (QuILT) to help students learn about the time-dependence of expectation values using the context of Larmor precession of spin and evaluating the role of asking students to self-diagnose their mistakes on midterm examination on their performance on subsequent problem solving. The QuILT on Larmor precession of spin has both paper-pencil activities and a simulation component. It was satisfying to see advanced undergraduate physics students improve significantly after learning from the QuILT and more students adopted quantum mechanical reasoning. The simulation portion of the tutorial was improved to reduce cognitive load for the students, but still has a late 1990's look and feel. In think-aloud interviews, it is appears that the simulation portion of the tutorial could be improved by a redesign of the user interface. While it may seem obvious that an instructor should award points for behaviors that the instructor deems desirable, in practice incentivizing students appropriately has not been investigated in diverse situations. One of the goals of PER is to identify ways to enhance student learning and assessment can play a key role in providing incentives to students to improve. Incentives other than grades exist, but are not usually effective at motivating students. It would be useful to investigate how students with different backgrounds are impacted by the grade incentives and if such an incentive is more effective for students with certain characteristics. It would also be interesting to investigate how student performance in a laboratory course may be impacted by self-diagnosis of some activity connected to grade incentive.

6.2 FUTURE EXTENSION OF STPFASL ANALYSIS

The results from the development and validation of the thermodynamic survey, STPFaSL, show that, with a well-designed survey, a wealth of knowledge can be obtained about student difficulties. It was tremendously gratifying to contribute to what I hope will be a widely used tool to help measure and improve the teaching of thermodynamics. At the very least, we have shed some light on how conceptually difficult thermodynamics is for students, and that while students improve as they advance to higher level courses, instructors could do significantly better by adopting research-based learning tools.

However, the process of obtaining this information is an extremely intensive task that only an expert could perform. The entanglement between the various topics, the context-dependence of learning, and pedagogical knowledge required to help students in the context of thermodynamics make for a task only an artisan can perform. However, this raises some practical questions about future application of SPFaSL.

6.2.1 Question: Can an *Arbitrary* Student Population be Analyzed?

The entangled nature of thermodynamics makes it difficult to invert an arbitrary set of responses on a survey into a model of a group of students' understanding. For instance, in order to apply the first and second law to a single segment of a process, a decision matrix like Figure 9 is followed to reduce the problem to algebraic statements. A student can make a mistake anywhere along this line of reasoning, and end up with an erroneous conclusion. This is why I claim that going from their conclusion backwards to discover their line of reasoning requires both content and pedagogical expertise.

One question that I am interested in exploring is if a student's Performance can be Predicted. One answer is "yes", provided that the student is interviewed several times using a think-aloud interview, and a detailed profile of the student's difficulties is constructed.

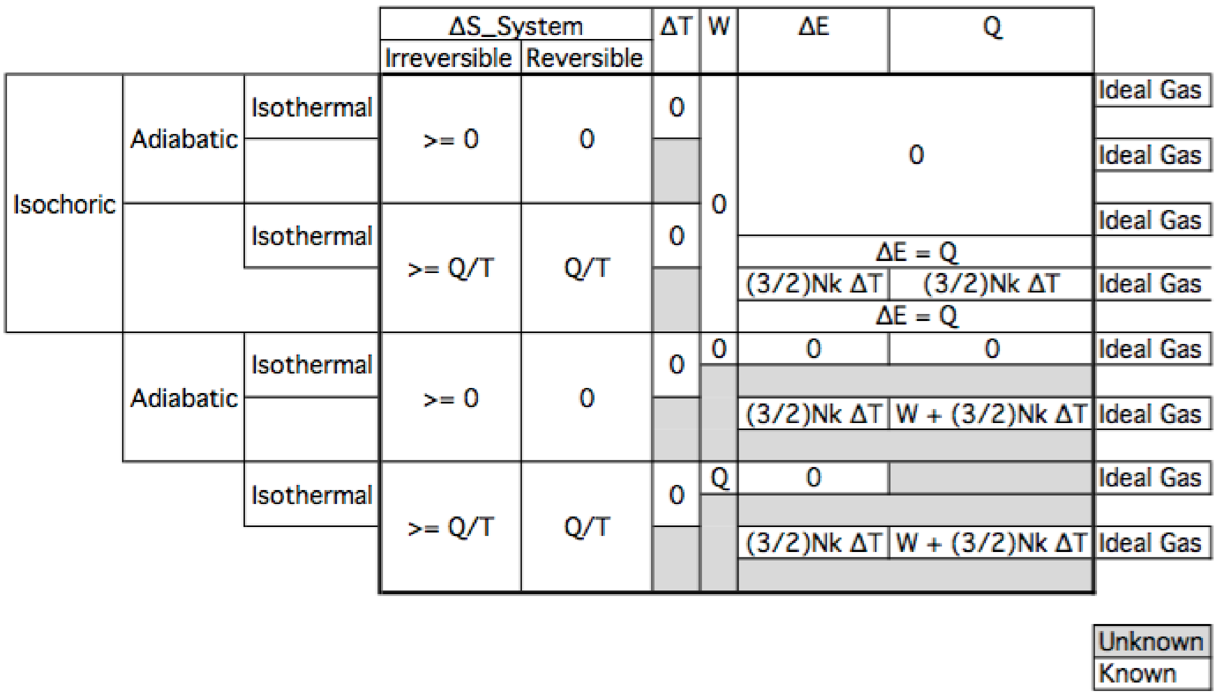


Figure 9: Mapping from a Description of a Single Thermodynamical Process to Algebraic Statements. The white boxes to the left and right label possible constraints or knowledge about a process that might be given in a typical problem. The information gained about each quantity $\Delta S_{system}, \Delta T, W, \Delta E$, and Q is found inside the box with the thick border.

But this approach is not practical or scalable to a large number of students. However, the purpose of a survey (e.g., STPFaSL), which can be administered to a large number of students, is to inform an instructor about the performance of a group of students to measure the effectiveness of instruction overall. Is there a way to extend the survey data to diagnosing students, or even a single student? Or, can a subset of the questions predict performance on another subset of non-identical questions? Item response theory (IRT) does not give a useful answer because IRT models student responses as a *linear* mapping from a (hidden) vector of student proficiencies to a single-parameter probability distribution of a correct response. Usually this is bounded above at 1, and bounded below at "guessing," or $1/n$, where n is the number of choices available for that item. What we see in student difficulties is that this model would be very poor: first, students are drawn by superficial features of problems towards distractor choices, they are not merely guessing. They are reliably attracted to effective distractors. However, this is a minor problem. The important problem with IRT is that the mapping from student knowledge of a subject, even if it is to a single parameter, is *certainly* not a linear one. It would be useful to test a *multiplicative* mapping of a hidden vector of student proficiencies to a parameter. For a knowledge-rich domain like physics, a student must master several concepts simultaneously to apply them to a problem. Missing one concept present in a problem is enough to dramatically reduce the reliability that a correct answer will be reached.

6.3 SUMMARY

Studying cognitive science and applying it to student understanding has been a fruitful exercise for researchers and students alike. When learning tools such as those discussed in this thesis are developed, and positive results (improvement in student performance) are observed, one can only hope that the influence of PER continues to grow. Viewing instructor evaluation through the lens of student performance can only yield benefits to a society. However, evidence-based teaching and learning can only become a widespread practice when institutional infrastructure supports these improvements.

APPENDIX A

TEXT-ONLY PORTION OF LARMOR PRECESSION TUTORIAL

The tutorial for time dependence of expectation values for Larmor precession of a spin-1/2 particle consists of a warm-up, the pre- and posttests, the main text portion and a simulation portion. Only the main text portion is shown here, as it was used during think-aloud interviews during the validation process.

An electron in a uniform magnetic field: Larmor precession of spin

For reference, if we choose the eigenstates of $\hat{S_z}$ as the basis vectors, the components of the spin angular momentum are given by:
$S_z = \hbar/2 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $S_x = \hbar/2 \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $S_y = \hbar/2 \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$.
The following identities may be useful:
$(e^{i\theta} + e^{-i\theta})/2 = \cos(\theta)$, $(e^{i\theta} - e^{-i\theta})/(2i) = \sin(\theta)$
$\cos^2(\theta) - \sin^2(\theta) = \cos(2\theta)$, $2\sin(\theta)\cos(\theta) = \sin(2\theta)$, $\cos^2(\theta) + \sin^2(\theta) = 1$

1. In quantum mechanics, the energy associated with the spin magnetic moment has a corresponding Hamiltonian operator $\hat{H}$ for the spin degrees of freedom. If the uniform magnetic field is $\vec{B} = B_0\hat{k}$, then $\hat{H} = -\gamma B_0 \hat{S}_z$. Which one of the following is the $\hat{H}$ in the matrix form in the chosen basis?
   (a) $-\gamma B_0 \hbar/2 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$
   (b) $-\gamma B_0 \hbar/2 \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$
   (c) $-\gamma B_0 \hbar/2 \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$
   (d) $-\gamma B_0 \hbar/2 \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$

2. If we measured the energy of the above system, which one of the following gives the allowed values of energy E_+ and E_-?
   (a) $\mp\gamma B_0 \hbar/2$
   (b) $\mp\hbar/2$
   (c) $\mp\gamma B_0 \hbar$
   (d) None of the above

3. Consider the following conversation between Andy and Caroline about the above Hamiltonian operator:
   • Andy: $\hat{H}$ is essentially $\hat{S_z}$ except for some multiplicative constants. Therefore, the eigenstates of $\hat{S_z}$ will also be the eigenstates of $\hat{H}$.
   • Caroline: No. The presence of magnetic field will make the eigenstates of $\hat{S_z}$ and $\hat{H}$

different. The eigenstates of $\hat{H}$ will change with time in a non-trivial manner.

- Andy: I disagree. If the magnetic field had a time dependence, e.g., $B = B_0\hat{k} + B_1\cos(\omega t)\hat{i}$, the eigenstates of $\hat{H}$ will change with time in a non-trivial manner but not for the present case where $\vec{B}$ is constant.

With whom do you agree?

(a) Andy

(b) Caroline

(c) Neither

4. If the eigenstates of $\hat{S_z}$ and $\hat{H}$, $|\uparrow\rangle_z$ and $|\downarrow\rangle_z$, are chosen as the basis vectors, which one of the following is their matrix representation?

(a) $\binom{1}{0}$ and $\binom{0}{1}$

(b) $\binom{1}{1}$ and $\binom{0}{0}$

(c) $1/\sqrt{2}\binom{1}{1}$ and $1/\sqrt{2}\binom{1}{-1}$

(d) $1/\sqrt{2}\binom{1}{i}$ and $1/\sqrt{2}\binom{1}{-i}$

5. If we choose $|\uparrow\rangle_z$ and $|\downarrow\rangle_z$ as the basis vectors for the two dimensional spin space, which one of the following is the correct expression for a general state $|\chi\rangle$?

(a) $|\chi\rangle = a|\uparrow\rangle_z + b|\downarrow\rangle_z$ where $|a| + |b| = 1$.

(b) $|\chi\rangle = a|\uparrow\rangle_z + b|\downarrow\rangle_z$ where $|a|^2 + |b|^2 = 1$.

(c) $|\chi\rangle = a|\uparrow\rangle_z + b|\downarrow\rangle_z$ where a and b can be any integers.

(d) $|\chi\rangle = a|\uparrow\rangle_z \times b|\downarrow\rangle_z$ where a and b can be any integers.

6. If the state of the system at an initial time $t = 0$ is given by $|\chi(t = 0)\rangle = a|\uparrow\rangle_z + b|\downarrow\rangle_z$, which one of the following is the correct matrix representation of this state in the chosen basis?

(a) $\binom{a}{b}$

(b) $\binom{a-b}{a+b}$

(c) $\binom{a+b}{a-b}$

(d) $1/\sqrt{2}\binom{a}{b}$

7. If the state of the system at an initial time $t = 0$ is given by $|\chi(t=0)\rangle = a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z$, which one of the following is the state, $|\chi(t)\rangle$, after a time t?
   (a) $e^{i\gamma B_0 t/2}(a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z)$
   (b) $e^{-i\gamma B_0 t/2}(a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z)$
   (c) $e^{i\gamma B_0 t/2}[(a+b)\,|\uparrow\rangle_z + (a-b)\,|\downarrow\rangle_z]$
   (d) $e^{i\gamma B_0 t/2} a\,|\uparrow\rangle_z + e^{-i\gamma B_0 t/2} b\,|\downarrow\rangle_z$

8. If the state of the system at an initial time $t = 0$ is given by $|\chi(t=0)\rangle = a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z$, which one of the following is the correct matrix representation of this state, $|\chi(t)\rangle$, after a time t?
   (a) $e^{i\gamma B_0 t/2}\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$
   (b) $e^{-i\gamma B_0 t/2}\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$
   (c) $e^{-i\gamma B_0 t/2}\left(\begin{smallmatrix} a+b \\ a-b \end{smallmatrix}\right)$
   (d) $\left(\begin{smallmatrix} a e^{i\gamma B_0 t/2} \\ b e^{-i\gamma B_0 t/2} \end{smallmatrix}\right)$

   Note: All of the following questions refer to the system for which the Hamiltonian operator is $\hat{H} = -\gamma B_0 \hat{S}_z$.

   In all of the questions below, $|\chi(t)\rangle$ is the state you calculated above and assume that the coefficients a and b in the state are non-zero unless mentioned specifically otherwise.

9. Which one of the following gives the correct outcomes of measuring S_z in the state $|\chi(t)\rangle$?
   (a) $\hbar/2$ with a probability $a e^{i\gamma B_0 t/2}$ and $-\hbar/2$ with a probability $b e^{-i\gamma B_0 t/2}$.
   (b) $\hbar/2$ with a probability $a^2 e^{i\gamma B_0 t}$ and $-\hbar/2$ with a probability $b^2 e^{-i\gamma B_0 t}$.
   (c) $\hbar/2$ with a probability $|a|^2$ and $-\hbar/2$ with a probability $|b|^2$.
   (d) $\hbar/2$ and $-\hbar/2$ with equal probability.

10. Consider the following conversation between Andy and Caroline about measuring $\hat{S_z}$ in the state $|\chi(t)\rangle$:
   • Andy: Since the probability of measuring $\hbar/2$ is $|a|^2$, $-\hbar/2$ is $|b|^2$ and $|a|^2 + |b|^2 = 1$, we can choose our a and b as $a = e^{i\phi_1}\cos(\alpha)$ and $b = e^{i\phi_2}\sin(\alpha)$ where α, ϕ_1 and ϕ_2 are real numbers.

• Caroline: I agree. Since $\cos^2(\alpha)+\sin^2(\alpha) = 1$ it gives the same relation as $|a|^2+|b|^2 = 1$ and there is no loss of generality.

Do you agree with Andy and Caroline?

(a) Yes.

(b) No.

11. Calculate the expectation value of $\hat{S_z}$ in the state $|\chi(t)\rangle$. Express your answer in terms of α defined earlier as $a = e^{i\phi_1}\cos(\alpha)$ and $b = e^{i\phi_2}\sin(\alpha)$.

12. Consider the following conversation between Pria and Mira about $\langle\hat{S}_Z\rangle$ in state $|\chi(t)\rangle$:

• Pria: Since the state of the system $|\chi(t)\rangle$ evolves in time, the expectation value $\langle S_Z\rangle$ will depend on time.

• Mira: I disagree with the second part of your statement. The time development of the expectation value of any operator $\hat{A}$ is given by $\frac{d\langle\hat{A}\rangle}{dt} = \frac{i}{\hbar}\langle[\hat{H},\hat{A}]\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle$. In our case, $[\hat{H},\hat{S}_Z] = 0$ and the operator $\hat{S_z}$ does not have any explicit time dependence so $\frac{\partial\hat{S}_Z}{\partial t} = 0$. Thus, $\frac{d\langle\hat{S}_Z\rangle}{dt} = 0$ and the expectation value will not change with time.

With whom do you agree?

(a) Pria

(b) Mira

(c) Neither

13. Which one of the following is the expectation value $\langle\hat{S}_Z\rangle = \langle\chi(t)|\hat{S}_Z|\chi(t)\rangle$?

(a) $\cos(2\alpha)\ \cos(\gamma B_0 t)\ \hbar/2$

(b) $\sin(2\alpha)\ \sin(\gamma B_0 t)\ \hbar/2$

(c) $\cos(2\alpha)\ \hbar/2$

(d) $\sin(2\alpha)\ \hbar/2$

14. Which one of the following is true about the expectation value $\langle\hat{S_x}\rangle$ in the state $|\chi(t)\rangle$?

(a) $\langle \hat{S_x} \rangle$ depends on time because $[\hat{H}, \hat{S_x}] \neq 0$

(b) $\langle \hat{S_x} \rangle$ depends on time because $\hat{S_x}$ has an explicit time-dependence.

(c) $\langle \hat{S_x} \rangle$ is time-independent because $[\hat{H}, \hat{S_x}] = 0$

(d) None of the above.

15. Calculate the expectation value of $\hat{S_x}$ in the state $|\chi(t)\rangle$. Express your answer in terms of α defined earlier as $a = e^{i\phi_1}\cos(\alpha)$ and $b = e^{i\phi_2}\sin(\alpha)$.

16. Which one of the following is the expectation value $\langle \hat{S_x} \rangle$ in the state $|\chi(t)\rangle$?

(a) $\cos(2\alpha)\,\cos(\gamma B_0 t + \phi_1 - \phi_2)\,\hbar/2$

(b) $\sin(2\alpha)\,\cos(\gamma B_0 t + \phi_1 - \phi_2)\,\hbar/2$

(c) $\cos(2\alpha)\,\hbar/2$

(d) $\sin(2\alpha)\,\hbar/2$

17. Choose all of the following statements that are true about the expectation value $\langle \hat{S_y} \rangle$ in the state $|\chi(t)\rangle$:

(I) $\langle \hat{S_y} \rangle$ depends on time because $[\hat{H}, \hat{S_y}] \neq 0$

(II) $\langle \hat{S_y} \rangle$ depends on time because $\hat{S_y}$ has an explicit time-dependence.

(III) $\langle \hat{S_y} \rangle$ is time-independent because $[\hat{H}, \hat{S_y}] = 0$

(a) (I) only

(b) (II) only

(c) (III) only

(d) (I) and (II) only

18. Based upon your responses for $\langle \hat{S_z} \rangle$ and $\langle \hat{S_x} \rangle$ above, which one of the following is a good guess for the expectation value $\langle \hat{S_y} \rangle$ in state $|\chi(t)\rangle$? Note: Please work this out explicitly as homework similar to your calculation for $\langle \hat{S_x} \rangle$ above.

(a) $\cos(2\alpha)\ \cos(\gamma B_0 t + \phi_1 - \phi_2)\ \hbar/2$

(b) $-\sin(2\alpha)\ \sin(\gamma B_0 t + \phi_1 - \phi_2)\ \hbar/2$

(c) $\cos(2\alpha)\ \hbar/2$

(d) $\sin(2\alpha)\ \hbar/2$

19. Explain why $\langle \hat{S}_z \rangle$ above does not depend on time where as $\langle \hat{S_x} \rangle$ and $\langle \hat{S_y} \rangle$ do.

20. Choose all of the following statements that are true about the expectation value $\langle \vec{S} \rangle$ in the state $|\chi(t)\rangle$:

(I) $\langle \vec{S} \rangle = \langle \hat{S_x} \rangle \hat{\imath} + \langle \hat{S_y} \rangle \hat{\jmath} + \langle \hat{S_z} \rangle \hat{k}$

(II) $\langle \vec{S} \rangle$ will depend on time because $[\hat{H}, \vec{S}] \neq 0$

(III) $\langle \vec{S} \rangle$ cannot depend on time because the expectation value of an observable is its time-averaged value.

(a) (I) only

(b) (II) only

(c) (III) only

(d) (I) and (II) only

21. Choose all of the following statements that are true about the expectation value $\langle \hat{S} \rangle$ in the state $|\chi(t)\rangle$:

(I) The z component of $\langle \hat{S} \rangle$, i.e., $\langle \hat{S_z} \rangle$, is time-independent.

(II) The x and y components of $\langle \hat{S} \rangle$ change with time and they are always "out of phase" with each other for all times.

(III) The magnitude of the maximum values of $\langle \hat{S_x} \rangle$ and $\langle \hat{S_y} \rangle$ are the same but when $\langle \hat{S_x} \rangle$ is a maximum $\langle \hat{S_y} \rangle$ is a minimum and vice versa.

(a) (I) and (II) only

(b) (I) and (III) only

(c) (II) and (III) only

(d) (I), (II) and (III)

22. Choose all of the following statements that are true about the vector $\langle \hat{S} \rangle$ in the spin space in the state $|\chi(t)\rangle$:

    (I) The vector $\langle \hat{S} \rangle$ can be thought to be precessing about the z axis at an angle 2α.

    (II) The vector $\langle \hat{S} \rangle$ can be thought to be precessing about the z axis with a frequency $\omega = \gamma B_0$.

    (III) All the three components of vector $\langle \hat{S} \rangle$ change as it precesses about the z axis.

(a) (I) and (II) only

(b) (I) and (III) only

(c) (II) and (III) only

(d) (I), (II) and (III)

23. Choose a three dimensional coordinate system in the spin space with the z axis in the vertical direction. Draw a sketch showing the precession of $\langle \hat{S} \rangle$ about the z axis when the state of the system starts out in $a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z$. Show the angle that $\langle \hat{S} \rangle$ makes with the z axis and the precession frequency explicitly.

    Show the projection of $\langle \hat{S} \rangle$ along the x, y and z axes at two separate times. Explain in words why the projection of $\langle \hat{S} \rangle$ along the z direction does not change with time but those along the x and y directions change with time.

Now consider a very specific initial state which is an eigenstate of $\hat{S_z}$, e.g., $|\uparrow\rangle_z$, in the following questions (as opposed to a very general initial state $a|\uparrow\rangle_z + b|\downarrow\rangle_z$ in the previous questions).

24. If the electron is initially in the state $|\chi(0)\rangle = |\uparrow\rangle_z$, write the state of the system $|\chi(t)\rangle$ after a time t. The Hamiltonian operator is $\hat{H} = -\gamma B_0 \hat{S}_z$.

25. Evaluate the expectation values of $\hat{S_x}$, $\hat{S_y}$ and $\hat{S_z}$ at time t in the above state. Which of these expectation values depend on time?

26. Explain why each expectation value you calculated in the previous question does or does not depend on time.

27. Calculate the expectation value $\langle\chi(t)|[\hat{H}, \hat{S_x}]|\chi(t)\rangle$ in the above state by writing $\hat{H} = -\gamma B_0 \hat{S}_z$ explicitly and acting with $\hat{S_z}$ on the state $|\chi(t)\rangle$.

28. Since $\frac{d\langle\hat{A}\rangle}{dt} = \frac{i}{\hbar}\langle[\hat{H}, \hat{A}]\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle$ what can you infer about the time-dependence of $\langle\hat{S_x}\rangle$ in the state $|\uparrow\rangle_z$ from your last response? What about $\langle\hat{S_y}\rangle$ or $\langle\hat{A}\rangle$ where the operator $\hat{A}$ does not have an explicit time-dependence?

29. Consider the following statements from Pria and Mira when the electron is initially in an eigenstate of $\hat{S_z}$. The Hamiltonian operator is $\hat{H} = -\gamma B_0 \hat{S}_z$.
   - Pria: The electron will NOT be in an eigenstate of $\hat{S_z}$ forever because the state will

evolve in time.

• Mira: I disagree. The eigenstates of $\hat{S_z}$ are also the eigenstates of $\hat{H}$. When the system is in an energy eigenstate or a stationary state, the time dependence is via an overall phase factor. The system stays in the stationary state. Since the system is in a stationary state, the expectation value of ANY operator (that does not have an explicit time dependence) will not depend on time as we saw due to $\langle\chi(t)|[\hat{H}, \hat{S_x}]\,|\chi(t)\rangle = 0$.

With whom do you agree? Explain why the other person is not correct.

(a) Pria

(b) Mira

30. Consider the following statements from Pria and Mira when the electron is initially in an eigenstate of $\hat{S_z}$. The Hamiltonian operator is $\hat{H} = -\gamma B_0 \hat{S}_z$.

• Pria: But shouldn't the expectation value $\langle\vec{S}\rangle = \langle S_x\rangle\hat{i} + \langle S_y\rangle\hat{j} + \langle S_z\rangle\hat{k}$ precess about the z axis whether the state is $|\uparrow\rangle_z$ or $a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z$?

• Mira: No. Since $|\uparrow\rangle_z$ is an eigenstate of the Hamiltonian, it is a stationary state which has a trivial time dependence. $\langle\hat{S_x}\rangle$ and $\langle\hat{S_y}\rangle$ do not depend on time since the state will remain an energy eigenstate and will evolve by a trivial overall time-dependent phase factor as $e^{-iE_+t/\hbar}\,|\uparrow\rangle_z$. There is no precession.

• Pria: How can that be?

• Mira: Another way to reason about it is by comparing $|\uparrow\rangle_z$ with $a\,|\uparrow\rangle_z + b\,|\downarrow\rangle_z$. If our state is $|\uparrow\rangle_z$ then $a = e^{i\phi_1}\cos(\alpha) = 1$ and $b = e^{i\phi_2}\sin(\alpha) = 0$ which implies that $\alpha = 0$ and $\langle\vec{S}\rangle$ is pointing along the z direction. Thus, there is no precession. If you calculate $\langle\hat{S_x}\rangle$ and $\langle\hat{S_y}\rangle$ in this state, they will both be always zero since the projection of $\langle\vec{S}\rangle$ along the x and y axis is zero.

With whom do you agree? Explain.

(a) Pria

(b) Mira

31. Calculate $\langle \hat{S_x} \rangle$ and $\langle \hat{S_y} \rangle$ in the state $e^{-iE_+t/\hbar} |\uparrow\rangle_z$ to verify your response above.

32. If the electron is initially in the state $|\downarrow\rangle_z$, what is $|\chi(t)\rangle$ at a time t. Find the angle that $\langle \vec{S} \rangle$ makes with the z axis. Is there any precession in this case? Explain.

33. Choose all of the following statements that are true if the system is initially in an eigenstate of $\hat{S_z}$, $|\uparrow\rangle_z$, and the Hamiltonian operator $\hat{H} = -\gamma B_0 \hat{S}_z$:

(I) $\langle \hat{S_x} \rangle$ depends on time because $[\hat{H}, \hat{S_x}] \neq 0$

(II) $\langle \hat{S_y} \rangle$ depends on time because $[\hat{H}, \hat{S_y}] \neq 0$

(III) The vector $\langle \hat{S} \rangle$ can be thought to be precessing about the z axis.

(a) (I) and (II) only

(b) (III) only

(c) (I), (II) and (III)

(d) None of the above

34. If $\hat{H} = -\gamma B_0 \hat{S}_z$ and the electron is initially in the state $|\uparrow\rangle_x$, what is the state after time t?

35. Based upon your answer to the preceding question, should the expectation value $\langle \hat{S_x} \rangle$, $\langle \hat{S_y} \rangle$ and $\langle \hat{S}_z \rangle$ depend on time if the initial state is an eigenstate of $\hat{S_x}$, i.e., $|\uparrow\rangle_x$? Work out each of these expectation values at time t to justify your answer for each case.

36. Consider the following conversation between three students about the preceding problem in which the Hamiltonian for the system $\hat{H} = -\gamma B_0 \hat{S}_z$:

Student A: If the electron is initially in an eigenstate of $\hat{S_x}$, expectation value of $\hat{S_x}$ is time-independent because the electron is stuck in the eigenstate.

Student B: I disagree. If the electron is initially in an eigenstate of $\hat{S_x}$, only the expectation value of $\hat{S_z}$ is time independent because $\hat{S_z}$ and the Hamiltonian H commute with each other. Since an eigenstate of $\hat{S_x}$ is not a stationary state, the time-dependence of the eigenstate of $\hat{S_x}$ is non-trivial. You can see this by writing the eigenstate of $\hat{S_x}$, e.g., $|\uparrow\rangle_x$, in terms of eigenstates of the Hamiltonian at time $t = 0$ and then writing down the state after time t explicitly.

Student C: S_z is a constant of motion since $\hat{S_z}$ and the Hamiltonian H commute with each other. Therefore, $\langle \hat{S}_z \rangle$ is time-independent in any state. With whom do you agree? Explain.

APPENDIX B

PROBLEMS AND RUBRIC FOR INCENTIVE STUDY

B.1 QUANTUM MECHANICS PROBLEMS

The following problems were given both in the midterm and final exams. The problem numbers refer to the numbering of the problems in the final exam. Students were given an additional sheet on which useful information was provided. For example, they were given the stationary state wave functions and energies for a one-dimensional infinite square well. For the one-dimensional Harmonic Oscillator, they were also given energies in terms of quantum number n, how the ladder operators relate to the position and momentum operators, the commutation relation between the raising and lowering operators, how a raising or lowering operator acting on the nth energy eigenstate of a one dimensional Harmonic Oscillator changes that state etc.

1) The eigenvalue equation for an operator $\hat{Q}$ is given by $\hat{Q}\,|\psi_i\rangle = \lambda_i\,|\psi_i\rangle$, with $i = 1...N$. Find an expression for $\langle\psi|\hat{Q}|\psi\rangle$, where $|\psi\rangle$ is a general state, in terms of $\langle\psi_i|\psi\rangle$.
2) For an electron in a one-dimensional infinite square well with well boundaries at $x = 0$ and $x = a$, measurement of position yields the value $x = a/2$. Write down the wave function immediately after the position measurement and without normalizing it show that if energy is measured immediately after the position measurement, it is equally probable to find the electron in any odd-energy stationary state.
3) Write an expression to show that the momentum operator $\hat{P}$ is the generator of translation in space. Then prove the relation. (Simply writing the expression is not sufficient... you need to prove it.)
4) Find the expectation value of potential energy in the n^{th} energy eigenstate of a one dimensional Harmonic Oscillator using the ladder operator method.

B.2 RUBRICS

A summary of the rubrics used to evaluate all four problems are shown in Tables 19, 20, 21, and 22. Each item receives an overall score, which is the average of the Invoking and Applying scores. Each of these general criteria scores is in turn the average of the specific criteria scores, which can take on the values 0, 1, 1/3, 1/2, 2/3, or N/A. N/A arises when certain criteria are not present to include in determining a score. Inter-rater reliability was tested for roughly 25% of the dataset and found to be better than 95%.

Table 19: Summary of the rubric used for problem 1. In this problem, students are asked to write the expectation value of an observable Q in terms of eigenstates and eigenvalues of the corresponding operator.

Problem 1	General Criteria	Specific Criteria
Overall Score	Invoking appropriate concepts	Spectral decomposition expressing identity operator in terms of a complete set of eigenstates $\|\psi_n\rangle$: $\hat{I} = \sum \|\psi_n\rangle \langle\psi_n\|$ Or expressing general state in terms of the eigenstates of $\hat{Q}$: $\|\psi\rangle = \sum c_n \|\psi_n\rangle$, where $c_n = \langle\psi_n\|\psi\rangle$
		Make use of $\hat{Q}\|\psi_n\rangle = \lambda_n \|\psi_n\rangle$
		$\langle\psi_n\|\psi\rangle^* = \langle\psi\|\psi_n\rangle$
		Using legitimate principles or concepts that are not appropriate in this problem.
		Using invalid principles or concepts (for instance, confusing a general state $\|\psi\rangle$ with an eigenstate $\|\psi_n\rangle$).
	Applying concepts	Inserting spectral decomposition into the expression for expectation value
		Eigenvalue evaluated and treated as number.
		Probability expressed in terms of $\langle\psi_n\|\psi\rangle^*$ and $\langle\psi\|\psi_n\rangle$

Table 20: Summary of the rubric used to solve problem 2. In this problem, a particle in a one dimensional infinite square well is first measured to be at the center of the well. Students are asked to show that a subsequent energy measurement can yield any odd-integer energy state with equal probability.

Problem 2	General Criteria	Specific Criteria
Overall Score	Invoking appropriate concepts	Measurement of position yields a Dirac delta function for the wave function immediately after the position measurement.
		Expand the wavefunction in terms of energy eigenfunctions: $\psi(x) = \sum c_n \psi_n(x)$
		Express probability amplitude for energy measurement as: $c_n = \int_{-\infty}^{\infty} \psi_n^*(x)\psi(x)\mathrm{d}x$
		Express the probability of measuring a given energy E_n as $\lvert c_n \rvert^2$
		Using legitimate principles or concepts that are not appropriate in this problem, e.g. invoking expectation values.
		Using invalid principles or concepts (for instance, confusing position eigenstates with energy eigenstates).
	Applying concepts	Applying delta function definition correctly to write the wavefunction after the position measurement: $\psi(x) = A\delta(x - \frac{a}{2})$
		Using provided stationary states for infinite square well: $\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) & \text{for } 0 \leq x \leq a \\ 0 & \text{otherwise} \end{cases}$
		Dirac delta function identity applied appropriately to calculate probability amplitude for energy measurement: $c_n = A\sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi x}{a}\right)\delta(x - \frac{a}{2})\mathrm{d}x$
		Find the probability for measuring energy E_n, $\lvert c_n \rvert^2 = \lvert A \rvert^2 \frac{2}{a} \sin^2(\frac{n\pi}{2}) = \lvert A \rvert^2 \frac{2}{a} \begin{cases} 1 & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}$

Table 21: Summary of the rubric used to solve problem 3. In this problem, students prove that momentum operator is the generator of translation in space.

Problem 3	General Criteria	Specific Criteria
Overall Score	Invoking appropriate concepts	Taylor expansion definition $f(x+x_0) = \sum\limits_{n=0}^{\infty} \frac{1}{n!} {x_0}^n \frac{\mathrm{d}}{\mathrm{d}x}^n f(x)$
		Momentum operator in position space in one dimension is $\hat{p} = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}$
		Expansion of exponential: $e^u = \sum\limits_{n=0}^{\infty} \frac{1}{n!} u^n$
		Using legitimate principles or concepts that are not appropriate in this problem.
		Using invalid principles or concepts (for instance, confusing position space with momentum space).
	Applying concepts	Partial derivative in terms of momentum operator: $\frac{\partial}{\partial x} = i\hat{p}/\hbar$
		$e^{i\hat{p}x_0/\hbar} = \sum\limits_n \frac{1}{n!} {x_0}^n (\frac{i\hat{p}}{\hbar})^n$
		Taylor expansion performed correctly to obtain: $f(x+x_0) = e^{i\hat{p}x_0/\hbar} f(x)$

Table 22: The summary of the rubric used to solve problem 4. In this problem, students are asked to find the expectation value of the potential energy for a one-dimensional harmonic oscillator when the system is in the n^{th} energy eigenstate.

Problem 4	General Criteria	Specific Criteria
Overall Score	Invoking appropriate concepts	$V = \frac{1}{2}m\omega^2 x^2 = \frac{1}{2}kx^2$ - or - $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$
		Express expectation value: $\langle f(x)\rangle = \int_{-\infty}^{\infty} \psi^*(x)f(x)\psi(x)\mathrm{d}x$
		Describe $\hat{x}$ (or $\hat{H}$) in terms of raising and lowering operators a_+ and a_- (as given in the formula sheet provided to students)
		Use orthogonality principle: $\int_{-\infty}^{\infty} \psi_m^*(x)\psi_n(x)\mathrm{d}x = \delta_{mn}$ Using legitimate principles or concepts that are not appropriate in this problem, e.g. ψ_n in terms of ψ_0.
		Using invalid principles or concepts (for instance, an incorrect definition of expectation value).
	Applying concepts	Proper expansion of $\langle x^2\rangle$ or $\langle H\rangle$, e.g. order of ladder operators in cross terms a_+a and a_-a_+ is correctly accounted for.
		Apply the operators correctly to obtain the correct states and coefficients. For instance, $a_+ \lvert\psi_n\rangle = \sqrt{n+1}\,\lvert\psi_{n+1}\rangle$
		Apply the orthogonality relation: $\int_{-\infty}^{\infty} \psi_n^*(x)\psi_{n\pm2}\,\mathrm{d}x = 0$, $\int_{-\infty}^{\infty} \psi_n^*(x)\psi_n\,\mathrm{d}x = 1$

B.3 EXAMPLE RESPONSES

Below, we provide typical sample student responses from incentivized group to show how students improved from pretest to posttest and from comparison groups to show how students deteriorated from the pretest to posttest on the same problem Each example includes both pretest and posttest (midterm and final exam) for a given student for a particular problem.

Incentivized Group **Problem 1**	**Student Work**
Pretest	$\langle\psi\|\hat{Q}\|\psi\rangle$ $= \sum_{i=1}^{N} \langle\psi\|\lambda_i\|\psi_i\rangle$ $= \sum_{i=1}^{N} \langle\psi\|(\langle\psi_i\|\psi\rangle)\|\psi_i\rangle$ $= \sum_{i=1}^{N} \|\langle\psi_i\|\psi\rangle\|^2$
Posttest	$\langle\psi\|\hat{Q}\|\psi\rangle = \sum_{i=1}^{N} \langle\psi\|\hat{Q}\|\psi_i\rangle\langle\psi_i\|\psi\rangle$ $= \sum_{i=1}^{N} \langle\psi\|\lambda_i\|\psi_i\rangle\langle\psi_i\|\psi\rangle$ $= \sum_{i=1}^{N} \lambda_i \langle\psi\|\psi_i\rangle\langle\psi_i\|\psi\rangle$ $= \sum_{i=1}^{N} \lambda_i \|\langle\psi_i\|\psi\rangle\|^2$

Figure 10: An example of a pretest and posttest solution pair for a student from the incentivized group which demonstrates improvement in student understanding. While the pretest solution shows difficulty with the expansion of a general state in terms of the eigenstates of the operator $\hat{Q}$ and confusion between the eigenvalue of $\hat{Q}$ and the probability amplitude for measuring Q, the posttest shows excellent use of Dirac notation to solve the problem.

Comparison Group

Problem 1	Student Work
Pretest	$\langle\psi\|\hat{Q}\|\psi\rangle =$
	$= \left\langle\psi\middle\|\hat{Q}\sum_{i=1}^{N}\psi_i\right\rangle\langle\psi_i\|\psi\rangle$
	$= \sum_{i=1}^{N}\langle\psi\|\hat{Q}\psi_i\rangle\langle\psi_i\|\psi\rangle$
	$= \sum_{i=1}^{N}\langle\psi\|q_i\psi_i\rangle\langle\psi_i\|\psi\rangle$
	$= \sum_{i=1}^{N}q_i\langle\psi\|\psi_i\rangle\langle\psi_i\|\psi\rangle$
	$\therefore \langle\psi\|\hat{Q}\|\psi\rangle = \sum_{i=1}^{N}q_i\langle\psi_i\|\psi\rangle^{*}\langle\psi_i\|\psi\rangle$
	in terms of $\langle\psi_i\|\psi\rangle$
Posttest	$\hat{Q}\|\psi_i\rangle = \lambda_i\|\psi_i\rangle$
	$\langle\psi\|\hat{Q}\|\psi\rangle = \int_{-\infty}^{\infty}\mathrm{dx}\langle\psi\|\hat{Q}\underbrace{\|\psi_i\rangle\langle\psi_i\|}_{1}\|\psi\rangle$
	$= \int_{-\infty}^{\infty}\mathrm{dx}\underbrace{\langle\psi\|\psi_i\rangle}_{\psi_i^{*}(x)}\hat{Q}\underbrace{\langle\psi_i\|\psi\rangle}_{\psi_i(x)}$

Figure 11: An example of a pretest and posttest solution pair for a student from the comparison group which demonstrates a typical deterioration. While the pretest solution is almost perfect, the posttest solution shows student difficulties including confusion between the discrete spectrum of the eigenstates of $\hat{Q}$ and the continuous spectrum of position eigenstates and difficulty with the treatment of the operator $\hat{Q}$.

Incentivized Group **Problem 2**	**Student Work**
Pretest	$\psi_n = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right)$ $\psi_n = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{2}\right)$ $= \sqrt{\frac{2}{a}}$ for $n = 1,3,5,\ldots$ otherwise $\psi_{\text{new}} = 0$
Posttest	$x_0 = \frac{a}{2}$ Immediately upon measuring the position, the wave function becomes a delta function: $\psi(t=0) = \delta(x - x_0)$ where $x_0 = \frac{a}{2}$ $c_n = \int \psi_n^*(x)\psi(t=0)\mathrm{d}x$ $= \sqrt{\frac{2}{a}}\int \sin\frac{n\pi x}{a}\delta(x-x_0)\mathrm{d}x = \sqrt{\frac{2}{a}}\sin\frac{n\pi x_0}{a} = \sqrt{\frac{2}{a}}\sin\frac{n\pi}{2} = c_n$ $\left\{\begin{array}{l}\text{Prob Energy}\\ = E_n\end{array}\right\} = \lvert c_n\rvert^2 = \frac{2}{a}\sin^2\frac{n\pi}{2} = \begin{cases}0 & n \text{ even}\\ \frac{2}{a} & n \text{ odd}\end{cases}$ $\left.\begin{array}{l}\text{Probability}\\ \text{Energy=E}_n\\ \text{where } n = \text{odd}\end{array}\right\} = \frac{2}{a}$ (odd energy states)

Figure 12: An example of a pretest and posttest solution pair for a student from the incentivized group which demonstrates improvement in student understanding. The pretest solution shows that the student incorrectly inserts $x = a/2$ in the expression for the energy eigenfunction instead of calculating the probability of measuring energy. The posttest solution is essentially perfect.

Problem 2	Student Work

Pretest

$$\psi(x,0)=\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$

$\langle x\rangle=\frac{a}{2}$ after a pos. measurement

$$\langle x\rangle=\int_0^a\psi_n^* x\psi_n=\frac{a}{2}=\frac{2}{a}\left(\frac{a^2}{4}\right)$$

$$\Psi(x,t)=\sum_{n=1}^{\infty}c_n\psi_n(x)e^{\frac{-iE_nt}{\hbar}}=A\left(c_1\psi_1e^{\frac{-iE_1t}{\hbar}}+\ldots+c_n\psi_ne^{\frac{-iE_nt}{\hbar}}\right)$$

$$\psi_n=A\sin\left(\frac{n\pi}{a}x\right)$$

$$\langle H\rangle=\int_0^a A^2\sin\left(\frac{n\pi}{a}x\right)\left(-\frac{\hbar^2}{2m}\partial x^2+0\right)\sin\left(\frac{n\pi}{a}x\right)\mathrm{dx}$$

$$=A^2\int_0^a\sin\left(\frac{n\pi}{a}x\right)\left[+\frac{\hbar^2}{2m}\left(\frac{n\pi}{a}\right)^2\sin\left(\frac{n\pi}{a}x\right)\right]\mathrm{dx}$$

$$=A^2\int_0^a\sin^2\left(\frac{n\pi}{a}x\right)\left(\frac{\hbar^2}{2m}\frac{n^2\pi^2}{a^2}\right)\mathrm{dx}$$

$$=A^2\left(\frac{\hbar^2n^2\pi^2}{2ma^2}\right)\int_0^a\sin^2\left(\frac{n\pi}{a}x\right)\mathrm{dx}$$

$$=A^2\left(\frac{\hbar^2n^2\pi^2}{2ma^2}\right)\left(\frac{1}{2}\right)$$

$$=A^2E_n(0.5)$$

50% probability of energy being in odd energy eigenstate.
50% probability of energy being in even energy eigenstate.

Posttest

$$\psi(x)=\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right)\rightarrow \begin{array}{l}\psi=A\sin kx+\cancel{B}^{\text{for } x=0}\cos kx\\ \text{for } x=\frac{a}{2}\quad \frac{ka}{2}=n\pi \quad \text{so } k=\frac{2n\pi}{a}\end{array}$$

$$\frac{-\hbar^2}{2m}\partial x^2\psi'+0\psi'=E\psi'$$

$$\partial x^2\psi'=\frac{-2mE'}{\hbar^2}\psi' \qquad \delta x^2\psi'=-k'^2\psi' \quad \text{where } k'=\frac{\sqrt{+2mE}}{\hbar}$$

$$\psi'(x)=A\sin\left(\frac{2n\pi}{a}x\right)$$

$$\hat{H}\psi'=\frac{-\hbar^2}{2m}\partial x^2\psi'(x)=A\left(-\frac{\hbar^2}{2m}\partial x\left(\frac{2n\pi}{a}\cos(\frac{2n\pi}{a}x)\right)\right)$$

$$=A\left(+\frac{\hbar^2}{2m}\left(\frac{\psi n^2\pi^2}{a^2}\right)\sin(2\frac{n\pi}{a}x)\right)=\frac{2\hbar^2n^2\pi^2}{ma^2}\psi(x)$$

So $E_n'=\frac{2\hbar^2n^2\pi^2}{ma^2}$ for $\psi'(x)$.

Figure 13: An example of a pretest and posttest solution pair for a student from the comparison group. The student struggles with this problem in the pretest. In the posttest, no improvement is demonstrated.

Problem 3	Student Work
Pretest	$\langle\psi\vert\hat{P}\vert\psi\rangle = \int_{\text{All Space}} \mathrm{dx}\langle\psi\vert x\rangle\langle x\vert\hat{P}\vert\psi\rangle$ $= \int_{-\infty}^{\infty}\mathrm{dx}\,\psi^*(x)\hat{P}\psi(x)$ $= \int_{-\infty}^{\infty}\psi^*(x)(-i\hbar)\frac{\partial\psi(x)}{\partial \mathrm{x}}\mathrm{dx}$ $\langle\psi\vert\hat{P}\vert\psi\rangle = \langle x\vert -i\hbar\frac{\partial}{\partial \mathrm{x}}\vert\psi\rangle$ $= -i\hbar\frac{\partial}{\partial \mathrm{x}}\langle x\vert\psi\rangle$ $= -i\hbar\frac{\partial\psi(x)}{\partial \mathrm{x}} = \hat{P}\psi(x)$ $\langle x\vert\hat{P}\vert p_0\rangle = \langle x\vert p_0\vert p_0\rangle$ $= p_0\langle x\vert p_0\rangle$ $= p_0 A e^{ip_0 x/\hbar}$
Posttest	$\psi(x+x_0) = \sum_n^{\infty}\frac{1}{n!}x^n\left(\frac{\partial}{\partial \mathrm{x}}\right)^n\psi(x)$ $= \sum_n^{\infty}\frac{1}{n!}x^n\left(\frac{i}{\hbar}\hat{P}\right)^n\psi(x)$ $= \sum_n^{\infty}\frac{1}{n!}\left(\frac{ix\hat{P}}{\hbar}\right)^n\psi(x) = e^{\frac{-ix\hat{p}}{\hbar}}\psi(x)$ Since $-i\hbar\frac{\partial}{\partial \mathrm{x}} = \hat{P}$ $\frac{\partial}{\partial \mathrm{x}} = \frac{i}{\hbar}P$

Figure 14: An example of a pretest and posttest solution pair for a student from the incentivized group which demonstrates improvement in student understanding. The pretest shows exploratory derivations focused on expectation value of momentum that has nothing to do with the correct solution. The posttest solution is a succinct derivation of the desired result.

Problem 3	**Student Work**
Pretest	$f_p(x+x_0) = \sum_n^{\infty} c_n (x_0)^n \left(\frac{\partial}{\partial \mathrm{x}}\right)^n f(x)$ But $\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial \mathrm{x}} \quad \frac{i\hat{p}}{\hbar} = \frac{\partial}{\partial \mathrm{x}}$ $f_p(x+x_0) = \sum_n^{\infty} c_n (x_0)^n \left(\frac{i\hat{p}}{\hbar}\right)^n f(x)$ $= e^{\frac{i\hat{p}x_0}{\hbar}} f(x)$
Posttest	$\underline{\hat{P}\lvert\beta\rangle = \lvert\alpha\rangle\langle\alpha\lvert\beta\rangle}$ $p = \frac{\hbar}{i}\frac{\partial}{\partial \mathrm{x}}$

Figure 15: An example of a pretest and posttest solution pair for a student from the comparison group which demonstrates a typical deterioration. The pretest solution is nearly perfect, but in the posttest, no relevant knowledge is displayed.

Problem 4 **Student Work**

Pretest

Expectation value of potential energy

$$\langle V\rangle = ?$$

$$\langle E\rangle = \int \psi_n^*\left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\right)\psi_n(x)\,\mathrm{dx}$$

$$\psi_n = \frac{1}{\sqrt{n!}}(a_+)^n\psi_0(x)$$

$$\psi_n^* = \frac{1}{\sqrt{n!}}(a_+^*)^n\psi_0^*(x)$$

$$\therefore \frac{\partial^2}{\partial x^2}\psi_n = \frac{\partial^2}{\partial x^2}\frac{1}{\sqrt{n!}}(a_+)^n\psi_0(x)$$

$$= \frac{1}{\sqrt{n!}}(a_+^*)^n\frac{\partial^2}{\partial x^2}\psi_0^*(x)$$

$$= \frac{1}{\sqrt{n!}}(a_+^*)^n\frac{\partial^2}{\partial x^2}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-(m\omega/2\hbar)x^2}$$

$$= \frac{1}{\sqrt{n!}}(a_+^*)^n\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\frac{\partial}{\partial x}\left(\left(\frac{-m\omega}{2\hbar}\right)e^{-m\omega/2\hbar\, x^2}\,2x\right)$$

$$= \frac{1}{\sqrt{n!}}(a_+^*)^n\left(-\left(\frac{m\omega}{\pi\hbar}\right)^{5/4}\right)\left(e^{-\frac{m\omega}{\pi\hbar}x^2} + xe^{-(m\omega/2\hbar)x^2}\left(\frac{-m\omega}{2\hbar}\right)2x\right)$$

$$= \frac{1}{\sqrt{n!}}(a_+^*)^n - \left(\frac{m\omega}{\pi\hbar}\right)^{5/4}\left(e^{-\frac{m\omega}{\pi\hbar}x^2}\right)\left(1 - x^2\left(\frac{m\omega}{\hbar}\right)\right)$$

$$\langle E\rangle = \int \frac{1}{n!}$$

Posttest

Expectation value of potential energy in n^{th} state

$$V = \frac{1}{2}m\omega^2x^2$$

$$\langle V\rangle = \frac{1}{2}m\omega^2\langle x^2\rangle = \frac{1}{2}m\omega^2\int\psi_n^*x^2\psi_n\,\mathrm{dx}$$

we know $x = \frac{\hbar}{2m\omega}(a_+ + a_-)$

$$\Rightarrow x^2 = \left(\sqrt{\frac{\hbar}{2m\omega}}\right)^2(a_+ + a_-)^2 = \frac{\hbar}{2m\omega}\left(a_+^2 + a_-^2 + a_+a_- + a_-a_+\right)$$

$$\Rightarrow x^2 = \frac{\hbar}{2m\omega}\left(a_+^2 + a_-^2 + a_+a_- + a_-a_+\right)$$

$$\therefore \langle x^2\rangle = \int_0^\infty \psi_n^*x^2\psi_n\,\mathrm{dx} = \int\psi_n^*\frac{\hbar}{2m\omega}\left(a_+^2 + a_-^2 + a_+a_- + a_-a_+\right)\psi_n\,\mathrm{dx}$$

$$a_+^2\psi_n = a_+(a_+\psi_n) = a_+\left(\sqrt{n+1}\,\psi_{n+1}\right) = \sqrt{n+1}\sqrt{n+2}\,\psi_{n+2}$$

$$a_-^2\psi_n = a_-(a_-\psi_n) = a_-\left(\sqrt{n}\,\psi_n\right) = \sqrt{n}\sqrt{n-1}\,\psi_{n-1}$$

$$a_+a_-\psi_n = a_+\left(\sqrt{n}\,\psi_{n-1}\right) = \sqrt{n}\left(a_+\psi_{n-1}\right) = \sqrt{n}\sqrt{n}\,\psi_n = n\psi_n$$

$$a_-a_+\psi_n = a_-\left(\sqrt{n+1}\,\psi_{n+1}\right) = \sqrt{n+1}\left(a_-\psi_{n+1}\right) = (n+1)\psi_n$$

$$\therefore \int\psi^*x^2\psi_n\,\mathrm{dx} = \frac{\hbar}{2m\omega}\Big[\sqrt{n+1}\sqrt{n+2}\int\psi_n^*\psi_{n+2}\,\mathrm{dx} + \sqrt{n}\sqrt{n-1}\int\psi_n^*\psi_{n-1}\,\mathrm{dx}$$
$$+ n\int|\psi_n|^2\,\mathrm{dx} + (n+1)\int|\psi_n|^2\,\mathrm{dx}\Big]$$

$$\therefore \langle x^2\rangle = \frac{\hbar}{2m\omega}[(2n+1)]$$

$$\therefore \langle V\rangle = \frac{1}{2}m\omega^2\langle x^2\rangle = \frac{1}{2}\cancel{m}\cancel{\omega}^2\frac{\hbar}{2\cancel{m}\cancel{\omega}}(2n+1)$$

$$= \frac{\hbar\omega}{2}\left(n + \frac{1}{2}\right)$$

Hence $\langle V\rangle = \frac{1}{2}\hbar\omega\left(n + \frac{1}{2}\right)$ Half of total energy E_n

Figure 16: An example of a pretest(left) and posttest(right) solution pair for a student from the incentivized group which demonstrates improvement in student understanding. The pretest shows that the student struggles and eventually abandons the problem, but in the posttest, the student demonstrates proficient use of the ladder operators to solve the problem.

Problem 4	Student Work
Pretest	$\langle V\rangle$ for $\psi_n = \frac{a_+^n}{\sqrt{n!}}\psi_0$ $\left\langle \frac{1}{2}m\omega^2x^2\right\rangle = \int_{-\infty}^{\infty}\psi_n^*\left(\frac{1}{2}m\omega^2x^2\right)\psi_n\,\mathrm{dx}$ $x = \sqrt{\frac{\hbar}{2m\omega}}\left(a_+ + a_-\right)$ $x^2 = \frac{\hbar}{2m\omega}\left(a_+^2 + a_+a_- + a_-a_+ + a_-^2\right)$ $\langle V\rangle = \frac{1}{2}\cancel{m}\omega^2\left(\frac{\hbar}{2\cancel{m}\omega}\right)\int_{-\infty}^{\infty}\psi_n^*\left(a_+^2 + a_+a_- + a_-a_+ + a_-^2\right)\psi_n\,\mathrm{dx}$ $= \frac{\hbar\omega}{4}\int_{-\infty}^{\infty}\psi_n^*a_+^2\psi_n + \psi_n^*a_+a_-\psi_n + \psi_n^*a_-a_+\psi_n + \psi_n^*a_-^2\psi_n\,\mathrm{dx}$ $= \frac{\hbar\omega}{4}\int_{-\infty}^{\infty}\psi_n^*a_+a_-\psi_n + \psi_n^*a_-a_+\psi_n\,\mathrm{dx}$ $a_+a_-\psi_n = a_+\left(\sqrt{n}\psi_{n-1}\right)$ $= \sqrt{n+1}\sqrt{n}\psi_{n-1+1}$ $= \sqrt{n^2+n}\psi_n$ $a_-a_+\psi_n = a_-\left(\sqrt{n+1}\psi_n\right)$ $= \sqrt{n^2+n}\psi_n$ $\int_{-\infty}^{\infty}\psi_n^*\sqrt{n^2+n}\psi_n\,\mathrm{dx}$ $= \sqrt{n^2+n}$ if ψ_n is normalized $= \frac{\hbar\omega}{4}\psi\sqrt{n^2+n}$ $= \frac{\hbar\omega}{2}\sqrt{n^2+n}$
Posttest	$\langle V\rangle = \int\psi_n^*V\psi_n\,\mathrm{dx}$ $= \left\langle E - \frac{p^2}{2m}\right\rangle \qquad p^2 = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial \mathrm{x}^2}$ $= \int\frac{a_+^n}{n!}\psi_0\left(E - \frac{p^2}{2m}\right)a_+^n\psi_0\,\mathrm{dx}$ $= \frac{E}{n!}\int\psi_n\psi_n\,\mathrm{dx} + \frac{\hbar^2}{2m}\int\psi_n\frac{\partial^2}{\partial \mathrm{x}^2}\psi_n\,\mathrm{dx}$

Figure 17: An example of a pretest and posttest solution pair for a student from the comparison group which demonstrates a typical deterioration. The student's pretest performance is good, except for some minor errors with the coefficients when dealing with the ladder operators but the posttest solution demonstrates that this proficiency has deteriorated.

APPENDIX C

FULL TEXT OF STPFASL

The design and validation of the survey are discussed in Chapter 4. Quantitative analysis of the types of difficulties is presented in chapter 5. Detailed analysis results can be found in the appendix following, appendices D. In this section, we show the survey in its entirety.

Survey of Thermodynamic Processes and First and Second Laws

- Please select only one of the five choices, (a)-(e) for each of the 33 questions.
- All temperatures T are **absolute** temperatures.
- All experiments involving a gas as the system are performed with a fixed amount of gas.
- The following equations may be useful for an ideal monatomic gas system where the symbols have the usual meaning: the internal energy $E_{int} = (3/2)NkT$ and $PV = NkT$.
- Thermal reservoirs are significantly larger than the system so that heat transfer between the system and the reservoir does not change the temperature of the reservoir.
- An adiabatic process is one in which there is no heat transfer between a system and its surroundings.

The following abbreviations are used throughout the survey:

- W = work done by the system (it is positive when the gas expands).
- Q = net heat transfer to the system.

Also, Q_1 , Q_2 in a particular problem will refer to the net heat transfer to the system in process 1 and process 2, respectively, etc.

You perform an experiment with a gas such that it undergoes a reversible adiabatic expansion. Answer questions (1) and (2) below about this experiment.

(1) Which one of the following statements is true about the change in entropy of the gas?

a. Entropy of the gas increases because the gas expands.

b. Entropy of the gas decreases because the gas expands.

c. Entropy of the gas remains constant for this reversible process because the process is adiabatic.

d. Entropy of the gas remains constant because the entropy of the gas does not change for a reversible process.

e. None of the above.

(2) Which one of the following statements must be true for the gas that undergoes a reversible adiabatic expansion process?

a. The internal energy must decrease, and the work done by the gas must be positive.

b. The internal energy must decrease, and the work done by the gas must be negative.

c. The internal energy must increase, and the work done by the gas must be positive.

d. The internal energy must increase, and the work done by the gas must be negative.

e. None of the above.

You perform an experiment with an ideal monatomic gas such that it undergoes a reversible isothermal compression. Answer questions (3) and (4) below about this experiment:

(3) Which one of the following is true about the net heat transfer during the reversible isothermal compression?

a. There is net heat transfer to the gas.

b. There is net heat transfer away from the gas.

c. There is no net heat transfer to or from the gas.

d. Net heat transfer could be to or from the gas depending on the initial values of P and V.

e. None of the above.

(4) Which one of the following is true about the entropy of the gas for the reversible isothermal compression?

a. It increases.

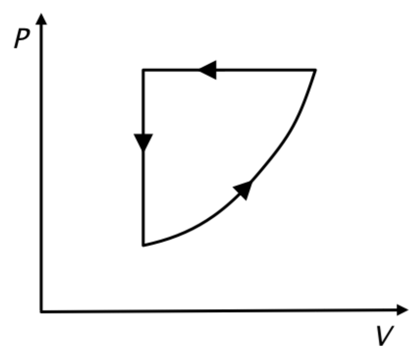
b. It decreases.

c. It remains the same.

d. It could increase or decrease depending on the initial values of P and V.

e. None of the above.

You perform an experiment with a gas in a container involving a piston and obtain the cycle shown on the PV diagram below in which the cycle is traversed counterclockwise. Use this PV diagram to answer questions (5), (6), (7), and (8).



In each of the questions (5), (6), (7) and (8) below, choose the statement that is true about the system for one complete cycle:

(5) For one complete cycle shown, the final internal energy of the gas

a. is the same as the initial internal energy.

b. is greater than the initial internal energy.

c. is less than the initial internal energy.

d. could be greater or less than the initial internal energy depending on the details of the cycle only shown schematically here.

e. is the same as the initial internal energy if the process is reversible, but otherwise it would be less.

(6) Which one of the following statements is true regarding the net work done by the gas for one complete cycle shown?

a. The net work done by the gas is positive.

b. The net work done by the gas is negative.

c. The net work done by the gas is zero.

d. The net work done by the gas could be positive or negative depending on the details of the cycle only shown schematically here.

e. The net work done by the gas is zero only if the process is reversible, but otherwise it would be negative.

(7) For one complete cycle shown, the final entropy of the gas

a. is greater than the initial entropy.

b. is less than the initial entropy.

c. is the same as the initial entropy.

d. could be greater or less than the initial entropy depending on the details of the cycle only shown schematically here.

e. is the same as the initial entropy if the process is reversible, but otherwise it would be greater.

(8) Which one of the following statements is true regarding the net heat transfer in one complete cycle shown?

a. There is no net heat transfer to or from the gas.

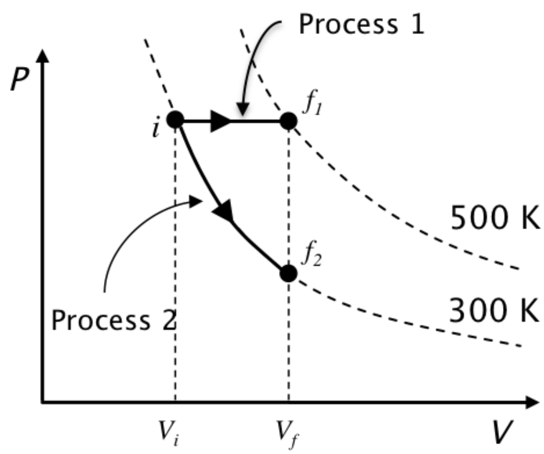
b. There is net heat transfer to the gas.

c. There is net heat transfer away from the gas.

d. Net heat transfer could be to *or* from the gas depending on the details of the cycle only shown schematically here.

e. There is no net heat transfer to or from the gas if the process is reversible, but otherwise net heat transfer is away from the gas.

(9) You carry out two experiments each with one mole of an ideal monatomic gas starting at the same point i shown on the PV diagram below. The two curved dashed lines on the PV diagram show two different isotherms at 300K and 500K. Process 1 is a constant pressure process starting at 300K and ending at point f_1 at 500K whereas process 2 (which starts at the same point as process 1) proceeds along the 300K isotherm to point f_2 as shown below. Both processes end with the same final volume V_f. Which one of the following is true about the relations between the work done by the gas and net heat transfer to the gas for the two processes?



a. $W_1 < W_2$ and $Q_1 < Q_2$.

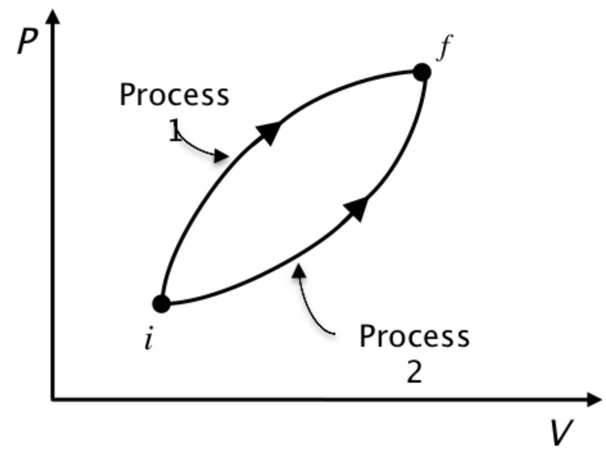
b. $W_1 > W_2$ and $Q_1 < Q_2$.

c. $W_1 < W_2$ and $Q_1 > Q_2$.

d. $W_1 > W_2$ and $Q_1 = Q_2$.

e. $W_1 > W_2$ and $Q_1 > Q_2$.

(10) You carry out two different processes each with one mole of a gas that start in the same thermodynamic state, i, and end in the same state, f, shown on the PV diagram below. Which one of the following statements correctly relates the net heat transfer to the gas in process 1, Q_1, with the net heat transfer to the gas in process 2, Q_2?



a. $Q_1 = Q_2$, because the same amount of work is done in both processes.

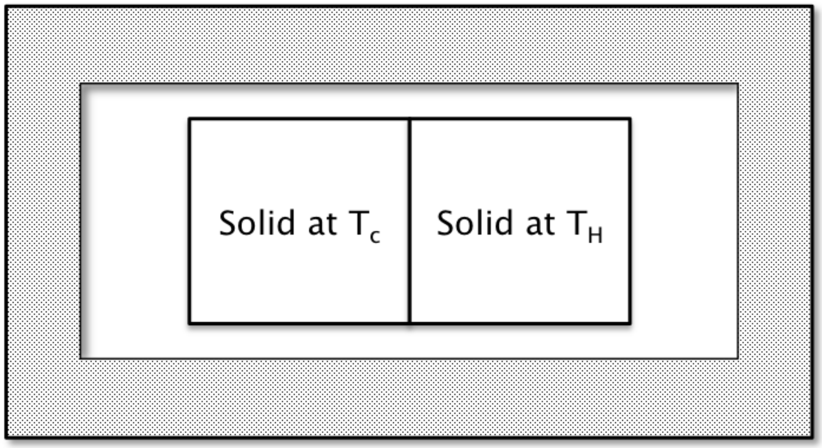
b. $Q_1 = Q_2$, because the two processes share the same initial and final states.

c. $Q_1 < Q_2$, because more work is done in process 2.

d. $Q_1 < Q_2$, because more work is done in process 1.

e. $Q_1 > Q_2$, because more work is done in process 1.

You perform an experiment involving two solids which are initially at temperatures T_C and T_H, respectively (where $T_C < T_H$). You place them in contact with each other inside of an insulated case to isolate them from everything else. Neglect thermal expansion, stresses, etc. Answer questions (11) through (13) related to this system.



(11) Choose all of the following statements that must be true for the process ending when thermal equilibrium between the two solids has been established:

I. The internal energy of the combined system of two solids increases.
II. The internal energy of the solid initially at temperature T_H increases.

a. I only.

b. II only.

c. Neither I nor II.

d. Both I and II.

e. Not enough information.

(12) Choose *all* of the following statements that must be true after the establishment of thermal equilibrium between the two solids:

I. The entropy of the solid initially at temperature T_C has increased.
II. The entropy of the solid initially at temperature T_H has increased.
III. The entropy of the combined system of two solids has increased.

a. I only.

b. II only.

c. III only.

d. I and III only.

e. I, II, and III.

(13) When the solids are placed in contact, no spontaneous net heat transfer will take place from the solid at the lower temperature T_C to the solid at the higher temperature T_H. Choose all of the following statements that explain why this process does not occur:

I. This process will violate Newton's second law.
II. This process will violate the first law of thermodynamics.
III. This process will violate the second law of thermodynamics.

a. I only.

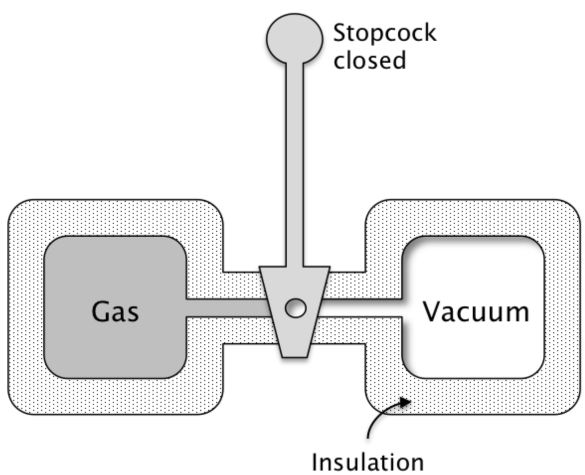
b. II only.

c. III only.

d. I and II only.

e. II and III only.

(14) The figure below shows the initial stage of a free-expansion process you carry out in the laboratory in which a gas is initially in thermal equilibrium and confined via a stopcock to one half of an insulated container with two identical halves. Then, you open the stopcock, and the gas expands freely to fill both halves of the container and eventually reaches an equilibrium state. Which one of the following statements is true about the system for this free expansion process?



a. Entropy remains the same; no work is done; internal energy remains the same.

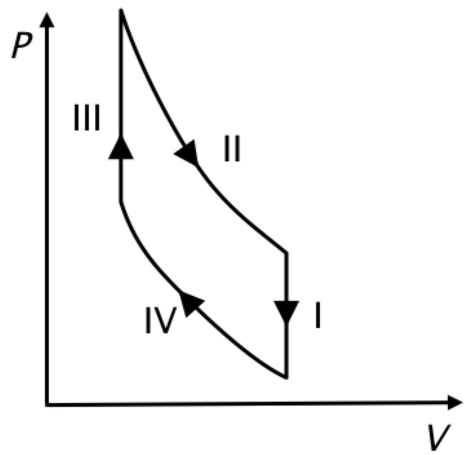
b. Entropy remains the same; positive work is done; internal energy decreases.

c. Entropy increases; positive work is done; internal energy remains the same.

d. Entropy increases; positive work is done; internal energy decreases.

e. Entropy increases; no work is done; internal energy remains the same.

(15) You perform an experiment involving a reversible cyclic thermodynamic process with one mole of a gas shown below on the PV diagram. Which one of the following statements is true about the gas <u>for one cycle</u> (clockwise) of the process?



I, III: constant volume processes
II, IV: adiabatic processes

a. The final entropy is same as the initial entropy; the final internal energy is the same as the initial internal energy; the net heat transfer is zero.

b. The final entropy is same as the initial entropy; the final internal energy is greater than the initial internal energy; the net heat transfer is zero.

c. The final entropy is greater than the initial entropy; the final internal energy is greater than the initial internal energy; the net heat transfer is zero.

d. The final entropy is greater than the initial entropy; the final internal energy is the same as the initial internal energy; there is net heat transfer to the gas.

e. The final entropy is same as the initial entropy; the final internal energy is the same as the initial internal energy; there is net heat transfer to the gas.

(16) Choose all of the following statements that must be true:

I. The entropy of any system that cannot exchange energy with its surrounding environment remains unchanged.
II. No heat engine can be more efficient than a reversible heat engine between given high temperature and low temperature reservoirs.
III. The net heat transfer to the system from a hot reservoir cannot be completely converted to mechanical work in a cyclic process.

a. I only.

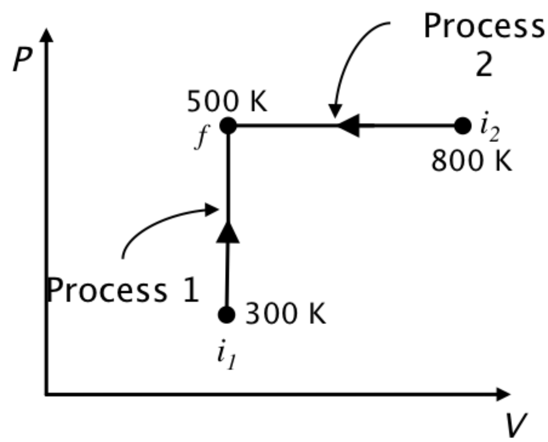
b. I and II only.

c. I and III only.

d. II and III only.

e. I, II, and III.

(17) You carry out two experiments each with one mole of an ideal monatomic gas such that both processes end in the same state shown on the PV diagram below. Process 1 is a constant volume process starting at 300K at point i_1 and ending at point f at 500K whereas process 2 is a constant pressure process starting at 800K at point i_2 and ending at point f at 500K. Which one of the following statements is true for the change in the internal energy of the gas for the two processes?



a. Internal energy does not change for process 1, and internal energy decreases for process 2.

b. Internal energy does not change for process 1, and internal energy increases for process 2.

c. Internal energy increases for both processes.

d. Internal energy does not change for either process.

e. Internal energy increases for process 1, and decreases for process 2.

(18) Choose all of the following thermodynamic processes in which there is <u>NO</u> net heat transfer between the system and the surroundings:

I. Any cyclical process.
II. Any adiabatic process.
III. Any isothermal process.

a. I only.

b. II only.

c. III only.

d. I and II only.

e. II and III only.

(19) Which one of the following statements is true for a Carnot engine? A heat engine operates using an appropriate working substance (system) that expands and compresses during each cycle. $|Q_H|$ = magnitude of net heat transfer to the system from the hot reservoir, $|Q_C|$ = magnitude of net heat transfer between the system and the cold reservoir, T_H = temperature of the hot reservoir, T_C = temperature of the cold reservoir.

a. Since a Carnot engine is a reversible engine, it is 100% efficient.

b. $|Q_H/T_H| > |Q_C/T_C|$

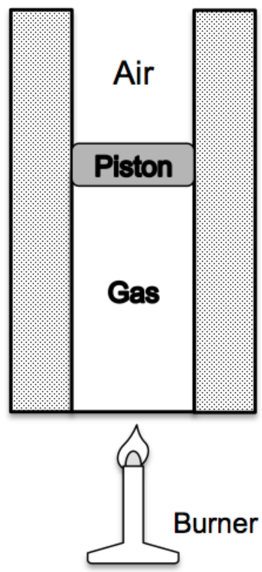
c. $|Q_H/T_H| < |Q_C/T_C|$

d. In one complete cycle, the non-zero change in entropy of the working substance is compensated by an opposite change in the net entropy of the hot and cold reservoirs.

e. In one complete cycle, the increase in entropy of the cold reservoir is compensated by a corresponding decrease in entropy of the hot reservoir.

(20) Chris heats an ideal gas in a closed container with a piston which can move freely, without friction. As the gas is heated with a burner, it expands and the piston moves slowly. There is no other thermal contact between the gas and the environment. Choose all of the following thermodynamic processes that are reasonable approximate descriptions for the given scenario:

I. A constant pressure process (isobaric).
II. An isothermal process.
III. An adiabatic process.



a. I only.

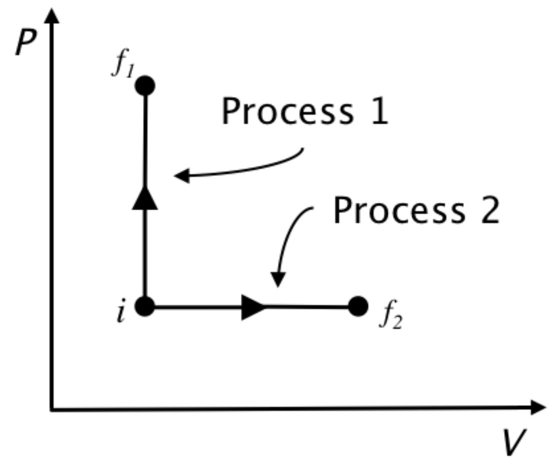
b. II only.

c. III only.

d. I and II only.

e. II and III only.

(21) You carry out two experiments each with one mole of an ideal monatomic gas such that both processes start in the same state i as shown on the PV diagram below. Process 1 is a constant volume (isochoric) process and process 2 is a constant pressure (isobaric) process. Choose all of the following statements that are correct about the two processes:



I. Between the gas and surroundings, there is no net heat transfer in process 1, but there is net heat transfer in process 2.
II. There is work done by the gas in process 2.
III. There is no change in the internal energy of the gas in process 1 but there is a change in the internal energy in process 2.

a. I only.

b. II only.

c. III only.

d. I and II only.

e. II and III only.

(22) For a given thermodynamic system, choose all of the following quantities whose values are determined by the state of the system and not by the process that led to that state.

I. Work done by the system.
II. Internal energy of the system.
III. Entropy of the system.

a. I only.

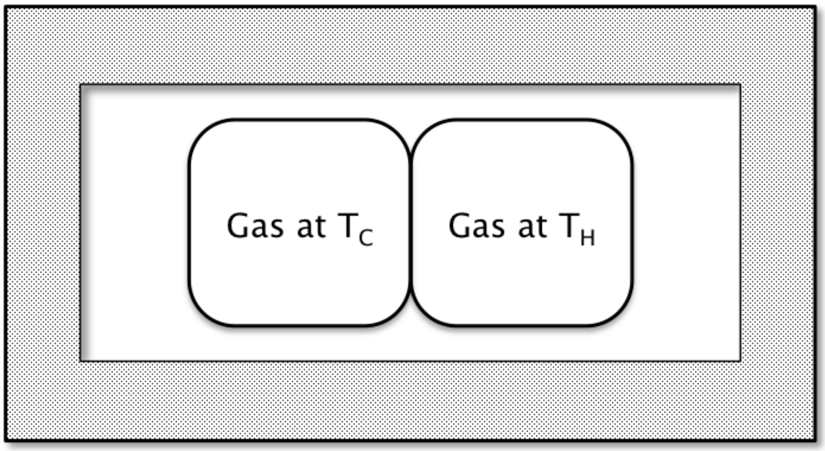
b. II only.

c. I and II only.

d. II and III only.

e. I, II, and III.

You perform an experiment in which two adjacent identical chambers, each with one mole of the same type of a gas, are initially at temperatures T_C and T_H, respectively ($T_C < T_H$). You place the two chambers in an insulated case to isolate them from everything else but in contact with each other such that while the gas molecules are confined to their original chamber, heat transfer can take place freely between the two chambers. The total volume of each chamber remains fixed throughout the process. Answer questions (23) through (25) related to this system.



(23) Choose all of the following statements that must be true for the process that ends when thermal equilibrium has been established:

I. The work done by the combined system of gases in the two chambers is zero.
II. The internal energy of the combined system has increased.

a. I only.
b. II only.
c. Neither I nor II.
d. Both I and II.
e. Not enough information.

(24) Choose *all* of the following statements that must be true after the establishment of thermal equilibrium between the gases in the two chambers:

I. The entropy of the gas in the chamber initially at temperature T_C has increased.
II. The entropy of the gas in the chamber initially at temperature T_H has increased.
III. The entropy of the combined system of two gases has increased.

a. I only.

b. II only.

c. III only.

d. I and III only.

e. I, II and III.

(25) There is no spontaneous net heat transfer from the chamber at the lower temperature T_C to the chamber at the higher temperature T_H when the two chambers at different temperatures are initially placed in contact. Choose all of the following statements that explain why this process does not occur:

I. This process will violate Newton's second law.
II. This process will violate the first law of thermodynamics.
III. This process will violate the second law of thermodynamics.

a. I only.

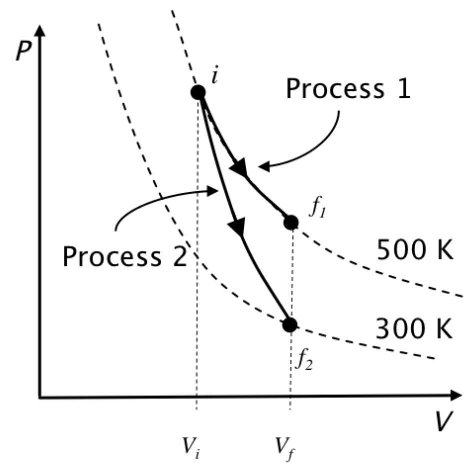
b. II only.

c. III only.

d. I and II only.

e. II and III only.

(26) You carry out two experiments each with one mole of a gas shown on the PV diagram below. Processes 1 and 2 (which are isothermal and adiabatic, respectively) both start at the same point i at 500K and the gas has the same final volume V_f at the end of both processes. Which one of the following statements is true about the two processes? W is the work done by the gas.



a. No work is done in the isothermal process since the temperature does not change, but W is positive in the adiabatic expansion.

b. W is larger for the adiabatic expansion because the change in pressure between the initial and final points is larger in that process.

c. W is equal for both processes because the final and initial volumes are the same.

d. W is larger in the adiabatic expansion because additional work must be done to change the temperature.

e. W is larger in the isothermal expansion because, for each small change in gas volume between the initial and final points, the pressure is higher.

(27) You expand an ideal monatomic gas isothermally. Choose all of the following statements that are true about this process:

I. There is no net heat transfer between the gas and its environment.
II. The internal energy of the gas does not change.
III. The work done by the gas is positive.

a. I only

b. I and II only.

c. I and III only.

d. II and III only.

e. I, II and III.

(28) You perform an experiment with an ideal monatomic gas in a closed container with a movable piston in contact with a thermal reservoir such that it undergoes an isothermal *reversible* expansion. There is net heat transfer to the gas during the process. Choose *all* of the following statements that must be true during the reversible isothermal expansion:

I. The entropy of the gas increases.
II. The entropy of the gas and the thermal reservoir taken together does not change.
III. The internal energy of the gas does not change.

a. I only.

b. I and II only.

c. I and III only.

d. II and III only.

e. I, II, and III.

(29) You perform an experiment with a gas in a closed container in contact with a thermal reservoir and it undergoes a constant volume (isochoric) *irreversible* thermodynamic process. There is net heat transfer to the gas during the process. Choose *all* of the following statements that must be true during the irreversible isochoric process:

I. The entropy of the gas increases.
II. The entropy of the gas and the thermal reservoir taken together does not change.
III. The internal energy of the gas does not change.

a. I only.

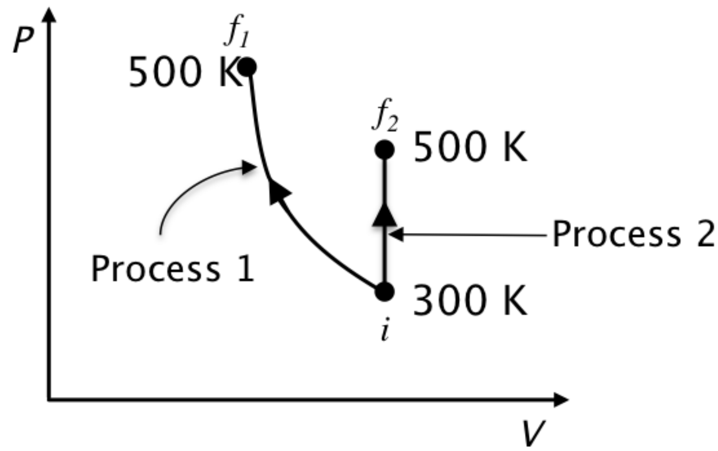
b. I and II only.

c. I and III only.

d. II and III only.

e. I, II, and III.

(30) You carry out two experiments each with one mole of an ideal monatomic gas. Both processes start at 300K at the same point i and both end at 500K but at different points f_1 and f_2 as shown on the PV diagram below. Processes 1 and 2 are adiabatic and isochoric (constant volume), respectively. Choose all of the following statements that are true about these processes:



I. The change in internal energy of the gas is equal for processes 1 and 2.
II. The magnitude of the work done on the gas in process 1 is equal to the net heat transfer to the gas in process 2.

a. I only.

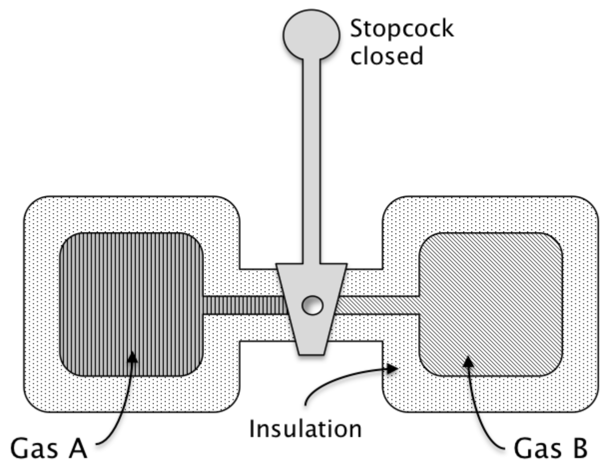
b. II only.

c. Neither I nor II.

d. Both I and II.

e. Not enough information.

You perform an experiment with one mole of an ideal gas A in one chamber and one mole of a different ideal gas B in a separate identical chamber. The two chambers are at the same temperature and pressure. You place the two chambers in an insulated case to isolate them from everything else but in contact with each other such that the chemically inert gas molecules can move freely between the two chambers. The total volume of the system remains fixed throughout the process. Answer questions (31) through (33) related to this system with the following initial and final states: initial state = one mole of gas A in the left chamber and one mole of gas B in the right chamber; final state = gas A and gas B mixed uniformly throughout the entire volume.



(31) Select the statement(s) that must be true for the process that ends when equilibrium has been established:

I. The work done by the combined system of two gases is zero.
II. The internal energy of the combined system has increased.

a. I only.

b. II only.

c. Neither I nor II.

d. Both I and II.

e. Not enough information.

(32) Choose *all* of the following statements that must be true after the establishment of equilibrium:

I. The entropy of gas A has increased compared to its initial value.
II. The entropy of gas B has increased compared to its initial value.
III. The entropy of the combined system of two gases has increased.

a. III only.

b. I and II only.

c. I, II, and III.

d. None of the above.

e. Not enough information.

(33) The reverse process to that described in questions (31) and (32) does not occur spontaneously, i.e., no spontaneous process will begin with gas A and gas B mixed uniformly throughout the entire volume and end with gas A in one chamber and gas B in the other. Choose all of the following statements that explain why this process does not occur:

I. This process will violate Newton's second law.
II. This process will violate the first law of thermodynamics.
III. This process will violate the second law of thermodynamics.

a. I only.

b. II only.

c. III only.

d. I and II only.

e. II and III only

APPENDIX D

SUPPLEMENTAL DATA AND ANALYSIS FROM STPFASL

Table 23: Average percentage scores for each of the five choices for each item on the survey for each group. Pre or pretest refers to the data before instruction in a particular course in which the survey topics in thermodynamics were covered (we note that students in a course may have learned these topics in other courses). Post or posttest refers to data after instruction in relevant concepts in that particular course. Abbreviations for various student groups: Upper (students in junior/senior level thermodynamics and physics graduate students in their first semester of a graduate program), calc (students in introductory calculus-based courses), Algebra (students in introductory algebra-based courses). The first column is the percentage of students who answer the item correctly, and the corresponding response. The four remaining columns are the percentage of incorrect answer (and choice), ranked by frequency.

Problem #	Correct (%)	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
1	**24 (C)**	39 (D)	28 (A)	5 (E)	4 (B)	*Upper Post*
	28 (C)	32 (A)	29 (D)	8 (B)	3 (E)	*Calc Post*
	19 (C)	51 (A)	16 (D)	12 (B)	2 (E)	*Calc Pre*
	29 (C)	36 (D)	29 (A)	4 (B)	2 (E)	*Algebra Post*
	24 (C)	46 (A)	19 (D)	10 (B)	1 (E)	*Algebra Pre*
2	**57 (A)**	14 (C)	10 (D)	9 (B)	9 (E)	*Upper Post*
	35 (A)	24 (C)	16 (D)	13 (E)	11 (B)	*Calc Post*
	29 (A)	33 (C)	14 (B)	14 (D)	9 (E)	*Calc Pre*
	39 (A)	20 (C)	15 (D)	14 (E)	13 (B)	*Algebra Post*
	27 (A)	38 (C)	13 (D)	12 (E)	10 (B)	*Algebra Pre*

continued on next page

Problem #	**Correct (%)**	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
3	**51 (B)**	23 (A)	18 (C)	8 (D)	0 (E)	*Upper Post*
	28 (B)	38 (C)	22 (A)	11 (D)	1 (E)	*Calc Post*
	28 (B)	29 (A)	24 (C)	19 (D)	1 (E)	*Calc Pre*
	29 (B)	33 (C)	19 (D)	18 (A)	1 (E)	*Algebra Post*
	23 (B)	29 (D)	26 (A)	21 (C)	1 (E)	*Algebra Pre*
4	**31 (B)**	42 (C)	23 (A)	4 (D)	0 (E)	*Upper Post*
	28 (B)	36 (C)	26 (A)	9 (D)	1 (E)	*Calc Post*
	37 (B)	25 (C)	23 (A)	13 (D)	1 (E)	*Calc Pre*
	32 (B)	43 (C)	12 (A)	11 (D)	1 (E)	*Algebra Post*
	44 (B)	20 (C)	19 (A)	17 (D)	0 (E)	*Algebra Pre*
5	**74 (A)**	9 (E)	7 (D)	6 (C)	4 (B)	*Upper Post*
	68 (A)	18 (E)	6 (C)	4 (D)	4 (B)	*Calc Post*
	61 (A)	15 (E)	9 (D)	9 (B)	7 (C)	*Calc Pre*
	50 (A)	27 (E)	9 (D)	8 (C)	5 (B)	*Algebra Post*
	63 (A)	16 (E)	11 (D)	6 (B)	4 (C)	*Algebra Pre*
6	**60 (B)**	23 (A)	9 (C)	5 (E)	3 (D)	*Upper Post*
	27 (B)	37 (C)	21 (A)	9 (D)	5 (E)	*Calc Post*
	10 (B)	56 (C)	18 (A)	8 (D)	8 (E)	*Calc Pre*
	9 (B)	47 (C)	28 (A)	10 (D)	6 (E)	*Algebra Post*
	8 (B)	53 (C)	20 (A)	12 (D)	8 (E)	*Algebra Pre*
7	**41 (C)**	36 (A)	9 (E)	7 (B)	7 (D)	*Upper Post*
	53 (C)	22 (A)	13 (E)	6 (B)	6 (D)	*Calc Post*
	55 (C)	17 (A)	10 (E)	9 (B)	8 (D)	*Calc Pre*
	43 (C)	22 (E)	21 (A)	7 (B)	7 (D)	*Algebra Post*
	62 (C)	12 (A)	11 (D)	10 (B)	5 (E)	*Algebra Pre*
8	**41 (C)**	36 (B)	11 (A)	10 (D)	3 (E)	*Upper Post*
	21 (C)	33 (A)	27 (B)	12 (D)	7 (E)	*Calc Post*
	12 (C)	40 (A)	25 (B)	13 (D)	9 (E)	*Calc Pre*
	10 (C)	37 (A)	24 (B)	19 (D)	10 (E)	*Algebra Post*
	12 (C)	38 (A)	21 (B)	18 (D)	12 (E)	*Algebra Pre*
9	**56 (E)**	18 (B)	16 (C)	7 (A)	3 (D)	*Upper Post*
	40 (E)	23 (C)	17 (B)	11 (D)	9 (A)	*Calc Post*
	26 (E)	24 (C)	19 (D)	19 (B)	12 (A)	*Calc Pre*
	38 (E)	21 (C)	17 (B)	15 (D)	10 (A)	*Algebra Post*
	32 (E)	21 (C)	19 (D)	15 (B)	13 (A)	*Algebra Pre*

continued on next page

Problem #	**Correct (%)**	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
10	**53 (E)**	19 (D)	11 (B)	11 (A)	6 (C)	*Upper Post*
	23 (E)	30 (B)	23 (A)	14 (D)	10 (C)	*Calc Post*
	15 (E)	40 (B)	27 (A)	9 (C)	9 (D)	*Calc Pre*
	23 (E)	36 (B)	27 (A)	8 (D)	6 (C)	*Algebra Post*
	17 (E)	41 (B)	28 (A)	11 (C)	3 (D)	*Algebra Pre*
11	**80 (C)**	9 (A)	4 (E)	4 (D)	3 (B)	*Upper Post*
	69 (C)	14 (A)	8 (B)	6 (D)	2 (E)	*Calc Post*
	57 (C)	17 (A)	11 (B)	11 (D)	4 (E)	*Calc Pre*
	65 (C)	19 (A)	10 (B)	3 (D)	2 (E)	*Algebra Post*
	50 (C)	26 (A)	11 (D)	9 (B)	4 (E)	*Algebra Pre*
12	**62 (D)**	17 (A)	13 (E)	6 (C)	2 (B)	*Upper Post*
	31 (D)	40 (A)	15 (E)	10 (C)	5 (B)	*Calc Post*
	24 (D)	44 (A)	13 (E)	13 (C)	6 (B)	*Calc Pre*
	32 (D)	41 (A)	12 (C)	12 (E)	3 (B)	*Algebra Post*
	29 (D)	39 (A)	17 (E)	11 (C)	4 (B)	*Algebra Pre*
13	**74 (C)**	14 (E)	9 (B)	3 (A)	0 (D)	*Upper Post*
	43 (C)	24 (E)	24 (B)	5 (D)	4 (A)	*Calc Post*
	36 (C)	30 (E)	23 (B)	7 (D)	5 (A)	*Calc Pre*
	61 (C)	18 (B)	14 (E)	4 (A)	3 (D)	*Algebra Post*
	43 (C)	24 (B)	19 (E)	10 (D)	4 (A)	*Algebra Pre*
14	**55 (E)**	18 (C)	16 (D)	8 (A)	3 (B)	*Upper Post*
	30 (E)	32 (C)	17 (A)	15 (D)	6 (B)	*Calc Post*
	21 (E)	28 (C)	20 (D)	17 (A)	14 (B)	*Calc Pre*
	28 (E)	31 (C)	19 (D)	19 (A)	3 (B)	*Algebra Post*
	20 (E)	28 (D)	24 (C)	16 (A)	12 (B)	*Algebra Pre*
15	**40 (E)**	29 (D)	22 (A)	6 (C)	4 (B)	*Upper Post*
	22 (E)	43 (A)	17 (D)	11 (C)	8 (B)	*Calc Post*
	11 (E)	43 (A)	18 (C)	14 (D)	13 (B)	*Calc Pre*
	11 (E)	50 (A)	14 (D)	13 (B)	12 (C)	*Algebra Post*
	9 (E)	53 (A)	15 (D)	12 (C)	11 (B)	*Algebra Pre*

continued on next page

Problem #	Correct (%)	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
16	**58 (D)**	22 (E)	15 (C)	4 (B)	1 (A)	*Upper Post*
	27 (D)	27 (C)	27 (E)	14 (B)	5 (A)	*Calc Post*
	19 (D)	33 (C)	21 (E)	17 (B)	10 (A)	*Calc Pre*
	27 (D)	40 (E)	19 (C)	9 (B)	5 (A)	*Algebra Post*
	16 (D)	36 (C)	24 (E)	12 (B)	11 (A)	*Algebra Pre*
17	**72 (E)**	14 (C)	7 (B)	4 (A)	3 (D)	*Upper Post*
	58 (E)	19 (C)	10 (A)	8 (B)	5 (D)	*Calc Post*
	50 (E)	19 (C)	12 (B)	11 (A)	9 (D)	*Calc Pre*
	48 (E)	21 (C)	14 (A)	10 (B)	7 (D)	*Algebra Post*
	55 (E)	20 (C)	9 (A)	9 (B)	7 (D)	*Algebra Pre*
18	**75 (B)**	9 (C)	6 (D)	6 (E)	3 (A)	*Upper Post*
	37 (B)	28 (C)	16 (E)	11 (D)	8 (A)	*Calc Post*
	20 (B)	34 (C)	16 (E)	15 (A)	15 (D)	*Calc Pre*
	46 (B)	21 (E)	18 (C)	9 (D)	6 (A)	*Algebra Post*
	37 (B)	23 (C)	15 (D)	14 (E)	12 (A)	*Algebra Pre*
19	**28 (E)**	34 (B)	19 (C)	16 (D)	4 (A)	*Upper Post*
	25 (E)	24 (B)	19 (C)	17 (D)	15 (A)	*Calc Post*
	18 (E)	26 (B)	20 (D)	19 (C)	17 (A)	*Calc Pre*
	37 (E)	25 (D)	18 (B)	10 (A)	10 (C)	*Algebra Post*
	21 (E)	27 (D)	22 (B)	18 (C)	11 (A)	*Algebra Pre*
20	**51 (A)**	20 (C)	16 (D)	10 (B)	3 (E)	*Upper Post*
	31 (A)	30 (C)	16 (D)	14 (B)	9 (E)	*Calc Post*
	23 (A)	26 (C)	22 (D)	17 (B)	12 (E)	*Calc Pre*
	44 (A)	22 (C)	14 (B)	11 (D)	9 (E)	*Algebra Post*
	23 (A)	24 (C)	20 (B)	18 (D)	16 (E)	*Algebra Pre*
21	**70 (B)**	12 (E)	10 (D)	5 (C)	2 (A)	*Upper Post*
	49 (B)	16 (D)	16 (E)	9 (A)	9 (C)	*Calc Post*
	35 (B)	20 (C)	19 (D)	13 (A)	13 (E)	*Calc Pre*
	44 (B)	21 (E)	15 (D)	12 (C)	9 (A)	*Algebra Post*
	37 (B)	19 (D)	18 (C)	16 (E)	10 (A)	*Algebra Pre*
22	**78 (D)**	16 (B)	4 (C)	3 (A)	0 (E)	*Upper Post*
	52 (D)	16 (B)	12 (E)	12 (C)	7 (A)	*Calc Post*
	44 (D)	19 (E)	17 (B)	12 (C)	8 (A)	*Calc Pre*
	49 (D)	16 (B)	13 (C)	11 (A)	11 (E)	*Algebra Post*
	51 (D)	18 (E)	16 (B)	10 (C)	6 (A)	*Algebra Pre*

continued on next page

Problem #	**Correct (%)**	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
23	**62 (A)**	22 (C)	9 (D)	5 (B)	2 (E)	*Upper Post*
	52 (A)	20 (C)	14 (D)	11 (B)	2 (E)	*Calc Post*
	43 (A)	21 (C)	18 (B)	12 (D)	5 (E)	*Calc Pre*
	46 (A)	20 (C)	18 (D)	14 (B)	1 (E)	*Algebra Post*
	40 (A)	25 (B)	20 (C)	11 (D)	3 (E)	*Algebra Pre*
24	**64 (D)**	15 (A)	14 (E)	5 (C)	2 (B)	*Upper Post*
	32 (D)	42 (A)	12 (E)	9 (C)	6 (B)	*Calc Post*
	24 (D)	39 (A)	13 (C)	12 (E)	11 (B)	*Calc Pre*
	31 (D)	41 (A)	11 (E)	9 (C)	8 (B)	*Algebra Post*
	30 (D)	43 (A)	11 (C)	9 (B)	7 (E)	*Algebra Pre*
25	**69 (C)**	20 (E)	9 (B)	2 (D)	0 (A)	*Upper Post*
	38 (C)	26 (B)	23 (E)	7 (D)	6 (A)	*Calc Post*
	29 (C)	28 (B)	25 (E)	12 (D)	5 (A)	*Calc Pre*
	57 (C)	18 (E)	17 (B)	5 (D)	3 (A)	*Algebra Post*
	35 (C)	21 (B)	21 (E)	15 (D)	8 (A)	*Algebra Pre*
26	**65 (E)**	12 (D)	11 (B)	10 (C)	2 (A)	*Upper Post*
	41 (E)	19 (B)	18 (C)	15 (D)	7 (A)	*Calc Post*
	17 (E)	25 (B)	23 (D)	22 (C)	13 (A)	*Calc Pre*
	31 (E)	26 (B)	20 (D)	14 (C)	9 (A)	*Algebra Post*
	17 (E)	27 (D)	27 (B)	18 (C)	11 (A)	*Algebra Pre*
27	**74 (D)**	10 (C)	6 (E)	6 (B)	4 (A)	*Upper Post*
	34 (D)	31 (C)	13 (E)	11 (B)	10 (A)	*Calc Post*
	25 (D)	27 (C)	21 (B)	15 (E)	11 (A)	*Calc Pre*
	35 (D)	29 (C)	16 (B)	11 (E)	9 (A)	*Algebra Post*
	28 (D)	25 (C)	25 (B)	14 (A)	8 (E)	*Algebra Pre*
28	**14 (E)**	26 (C)	26 (D)	20 (B)	14 (A)	*Upper Post*
	14 (E)	27 (D)	27 (B)	17 (C)	16 (A)	*Calc Post*
	11 (E)	25 (B)	23 (D)	22 (A)	18 (C)	*Calc Pre*
	14 (E)	30 (D)	28 (B)	14 (A)	14 (C)	*Algebra Post*
	9 (E)	33 (D)	23 (A)	18 (B)	18 (C)	*Algebra Pre*

continued on next page

Problem #	Correct (%)	1^{st}	2^{nd}	3^{rd}	4^{th}	Level
29	**68 (A)**	15 (B)	9 (D)	7 (C)	1 (E)	*Upper Post*
	43 (A)	24 (B)	17 (C)	8 (D)	7 (E)	*Calc Post*
	35 (A)	23 (B)	20 (C)	15 (D)	7 (E)	*Calc Pre*
	57 (A)	16 (B)	13 (C)	10 (D)	4 (E)	*Algebra Post*
	40 (A)	21 (B)	17 (D)	15 (C)	7 (E)	*Algebra Pre*
30	**37 (D)**	24 (A)	22 (C)	11 (B)	5 (E)	*Upper Post*
	19 (D)	36 (C)	23 (B)	19 (A)	4 (E)	*Calc Post*
	17 (D)	33 (C)	25 (B)	18 (A)	8 (E)	*Calc Pre*
	14 (D)	47 (C)	22 (B)	9 (A)	8 (E)	*Algebra Post*
	11 (D)	42 (C)	21 (A)	18 (B)	8 (E)	*Algebra Pre*
31	**57 (A)**	24 (C)	11 (B)	6 (D)	2 (E)	*Upper Post*
	42 (A)	20 (B)	18 (C)	16 (D)	4 (E)	*Calc Post*
	36 (A)	21 (C)	20 (B)	17 (D)	5 (E)	*Calc Pre*
	46 (A)	21 (D)	16 (C)	14 (B)	2 (E)	*Algebra Post*
	35 (A)	25 (B)	21 (D)	16 (C)	3 (E)	*Algebra Pre*
32	**62 (C)**	21 (A)	8 (D)	6 (B)	3 (E)	*Upper Post*
	37 (C)	23 (A)	21 (D)	13 (B)	7 (E)	*Calc Post*
	33 (C)	23 (D)	21 (A)	15 (B)	8 (E)	*Calc Pre*
	40 (C)	22 (A)	18 (D)	15 (B)	5 (E)	*Algebra Post*
	27 (C)	31 (A)	16 (D)	13 (B)	12 (E)	*Algebra Pre*
33	**75 (C)**	12 (E)	6 (B)	4 (D)	2 (A)	*Upper Post*
	43 (C)	22 (B)	19 (E)	9 (D)	7 (A)	*Calc Post*
	32 (C)	24 (B)	20 (E)	13 (A)	10 (D)	*Calc Pre*
	49 (C)	18 (E)	15 (B)	12 (D)	6 (A)	*Algebra Post*
	25 (C)	23 (A)	19 (B)	18 (E)	16 (D)	*Algebra Pre*

Table 24: Average percentage scores for Graduate students (left) vs graduate student pairs (right) for each of the five options in each item.

#	(%)	Incorrect %				(%)	Incorrect %			
1	**42 (C)**	38 (D)	13 (A)	4 (B)	2 (E)	**81 (C)**	10 (D)	10 (A)	0 (E)	0 (B)
2	**62 (A)**	18 (B)	13 (E)	7 (D)	0 (C)	**81 (A)**	10 (B)	5 (E)	5 (C)	0 (D)
3	**64 (B)**	18 (A)	13 (C)	4 (D)	0 (E)	**81 (B)**	14 (A)	5 (C)	0 (E)	0 (D)
4	**40 (B)**	38 (C)	22 (A)	0 (E)	0 (D)	**71 (B)**	14 (C)	14 (A)	0 (E)	0 (D)
5	**62 (A)**	13 (D)	11 (E)	11 (C)	2 (B)	**71 (A)**	14 (D)	10 (B)	5 (E)	0 (C)
6	**56 (B)**	24 (A)	18 (C)	2 (E)	0 (D)	**76 (B)**	10 (C)	10 (A)	5 (D)	0 (E)
7	**52 (C)**	18 (A)	11 (E)	11 (B)	7 (D)	**52 (C)**	33 (A)	5 (E)	5 (D)	5 (B)
8	**33 (C)**	40 (B)	16 (D)	11 (A)	0 (E)	**48 (C)**	38 (B)	14 (D)	0 (E)	0 (A)
9	**62 (E)**	18 (C)	9 (B)	7 (D)	4 (A)	**86 (E)**	10 (B)	5 (D)	0 (C)	0 (A)
10	**58 (E)**	18 (D)	13 (B)	9 (C)	2 (A)	**81 (E)**	19 (D)	0 (C)	0 (B)	0 (A)
11	**82 (C)**	7 (D)	4 (E)	4 (A)	2 (B)	**81 (C)**	10 (E)	10 (A)	0 (D)	0 (B)
12	**51 (D)**	24 (A)	16 (E)	7 (C)	2 (B)	**86 (D)**	14 (A)	0 (E)	0 (C)	0 (B)
13	**69 (C)**	13 (E)	13 (B)	4 (A)	0 (D)	**86 (C)**	5 (E)	5 (B)	5 (A)	0 (D)
14	**53 (E)**	16 (D)	16 (C)	16 (A)	0 (B)	**81 (E)**	10 (C)	5 (D)	5 (A)	0 (B)
15	**43 (E)**	20 (D)	20 (A)	11 (C)	5 (B)	**57 (E)**	33 (D)	5 (B)	5 (A)	0 (C)
16	**49 (D)**	38 (E)	9 (C)	4 (B)	0 (A)	**67 (D)**	29 (E)	5 (C)	0 (B)	0 (A)
17	**64 (E)**	11 (C)	9 (B)	9 (A)	7 (D)	**90 (E)**	5 (C)	5 (A)	0 (D)	0 (B)
18	**64 (B)**	16 (E)	9 (D)	7 (C)	4 (A)	**90 (B)**	5 (E)	5 (D)	0 (C)	0 (A)
19	**36 (E)**	27 (D)	23 (B)	14 (C)	0 (A)	**52 (E)**	24 (D)	14 (B)	10 (C)	0 (A)
20	**62 (A)**	16 (D)	13 (C)	7 (B)	2 (E)	**76 (A)**	19 (D)	5 (E)	0 (C)	0 (B)
21	**67 (B)**	16 (D)	11 (C)	7 (E)	0 (A)	**90 (B)**	10 (E)	0 (D)	0 (C)	0 (A)
22	**62 (D)**	31 (B)	4 (C)	2 (A)	0 (E)	**76 (D)**	19 (B)	5 (E)	0 (C)	0 (A)
23	**57 (A)**	16 (D)	16 (C)	9 (B)	2 (E)	**65 (A)**	20 (C)	10 (D)	5 (B)	0 (E)
24	**58 (D)**	16 (A)	13 (E)	7 (C)	7 (B)	**85 (D)**	10 (A)	5 (E)	0 (C)	0 (B)
25	**64 (C)**	18 (E)	16 (B)	2 (D)	0 (A)	**85 (C)**	10 (B)	5 (D)	0 (E)	0 (A)
26	**60 (E)**	16 (B)	13 (D)	9 (C)	2 (A)	**86 (E)**	10 (D)	5 (A)	0 (C)	0 (B)
27	**64 (D)**	16 (B)	7 (E)	7 (C)	7 (A)	**100 (D)**	0 (E)	0 (C)	0 (B)	0 (A)
28	**9 (E)**	36 (C)	25 (B)	16 (A)	14 (D)	**29 (E)**	48 (C)	14 (D)	5 (B)	5 (A)
29	**64 (A)**	20 (B)	9 (C)	7 (D)	0 (E)	**86 (A)**	14 (B)	0 (E)	0 (D)	0 (C)
30	**22 (D)**	31 (A)	22 (C)	20 (B)	4 (E)	**48 (D)**	48 (A)	5 (E)	0 (B)	0 (C)
31	**49 (A)**	24 (C)	16 (B)	7 (D)	4 (E)	**67 (A)**	19 (B)	14 (C)	0 (E)	0 (D)
32	**62 (C)**	13 (A)	11 (D)	9 (B)	4 (E)	**76 (C)**	14 (D)	5 (B)	5 (A)	0 (E)
33	**68 (C)**	11 (E)	9 (D)	9 (B)	2 (A)	**76 (C)**	19 (B)	5 (E)	0 (D)	0 (A)

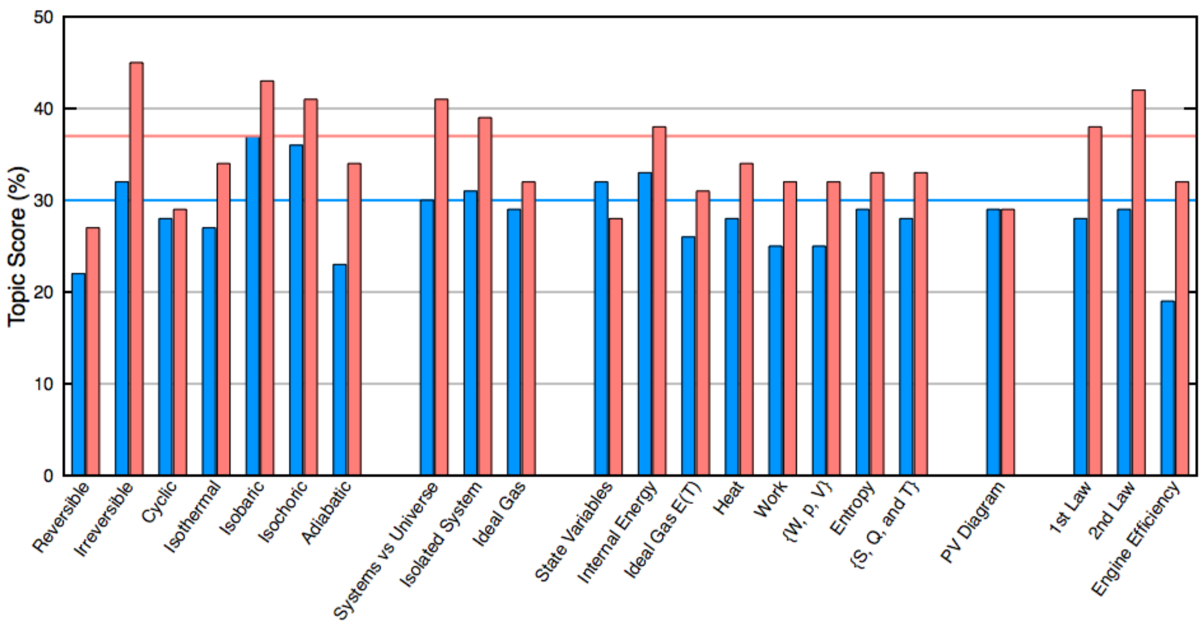


Figure 18: Average percentage scores for students in algebra-based introductory physics courses on STPFaSL by topic before and after instruction (for the pretest, the number of students N=218, blue, and for the posttest, N=382, red). It is worth noting that students in the algebra-based introductory courses improved more than those in calculus-based introductory courses. In fact, students in algebra-based courses perform better on posttest than students in the calculus-based introductory courses on posttest on this conceptual survey.

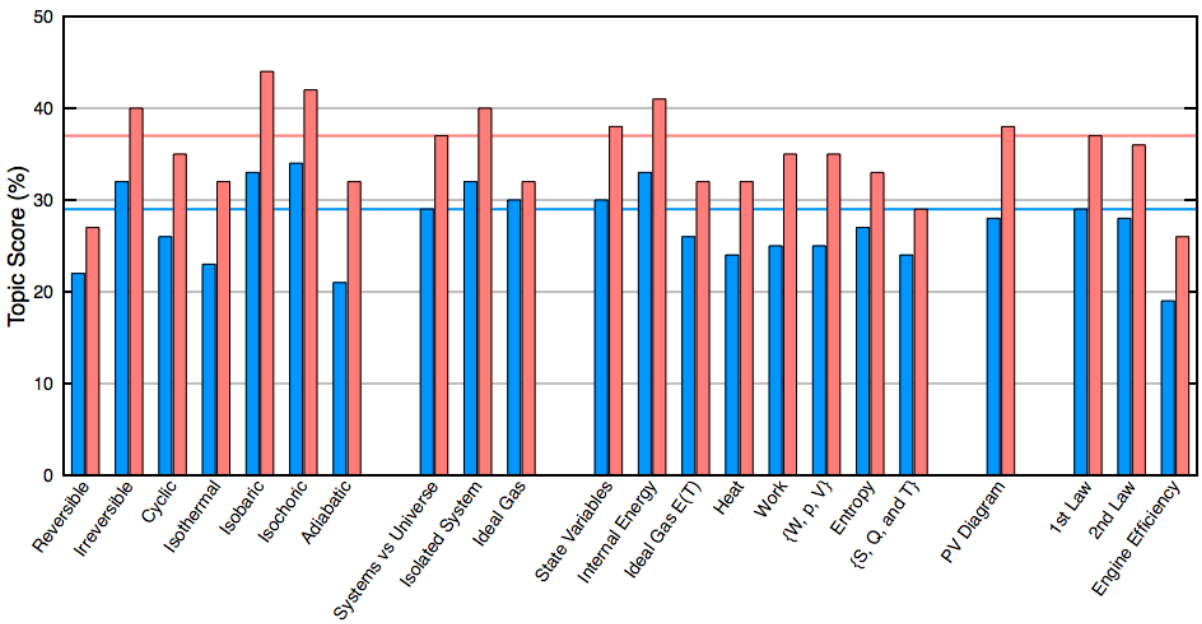


Figure 19: Average percentage scores for students in calculus-based introductory physics courses on STPFaSL by topic before and after instruction (for the pretest, the number of students N=704, blue, and for the posttest, N=505, red).

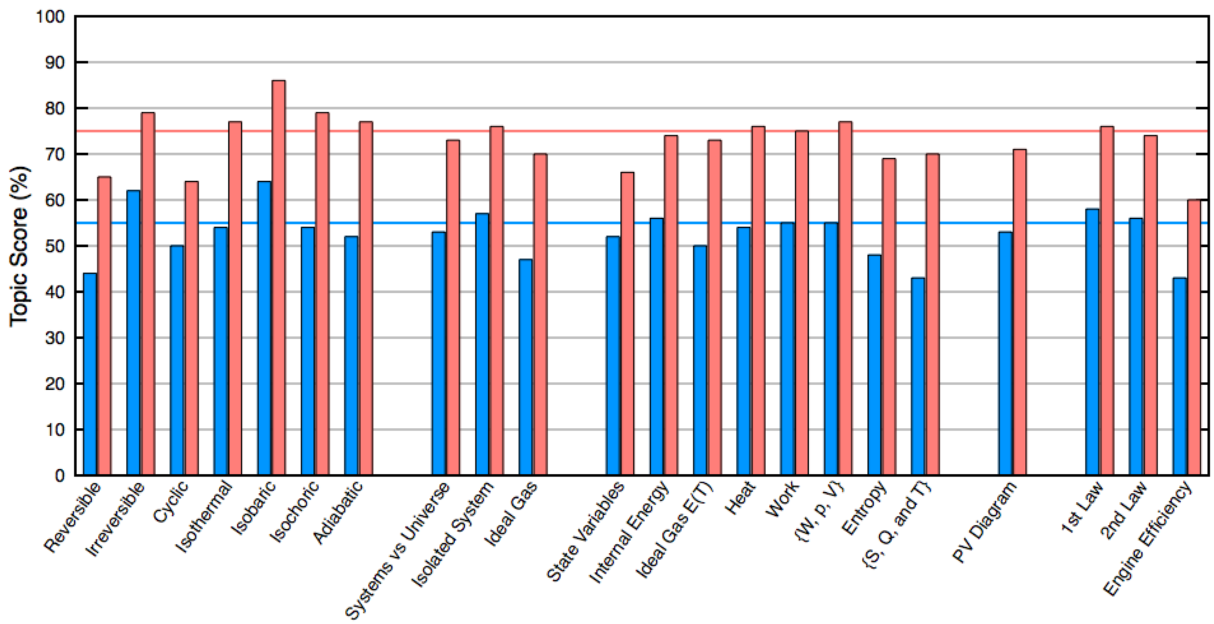


Figure 20: Average percentage scores for physics graduate students individually and then in pairs on STPFaSL by topic (for individual graduate students, the number of students N=45, blue, and for the paired test,N=21, red).

APPENDIX E

EXTENDED DIFFICULTIES DATA FOR STPFASL

Table 25: Prevalence of Difficulties Related to Cyclic Processes for each item

		Prevalence (%) by Level				
		Upper	Calc	Calc	Algebra	Algebra
Correct Answer, Difficulties	Item #	Post	Post	Pre	Post	Pre
$\mathbf{\Delta E = 0}$ **after a cycle. (correct)**	***5***	**74**	**68**	**61**	**50**	**63**
$\Delta E \neq 0$ after a cycle.	*5*	26	32	39	50	37
$\Delta E = 0$ only if a cycle is reversible.	*5*	9	18	15	27	16
When explicitly given that a cycle is reversible, $\Delta E \neq 0$	*15*	10	19	31	25	23
$\mathbf{W \neq 0}$ **for a cycle. (correct)**	***6***	**83**	**49**	**28**	**37**	**28**
$W = 0$ for a cycle.	*6*	14	43	64	53	61
$W = 0$ only if a cycle is reversible.	*6*	5	5	8	6	8
$\mathbf{Q \neq 0}$ **for a cycle. (correct)**	***8***	**76**	**48**	**38**	**34**	**33**
$Q = 0$ for a cycle.	*8*	14	40	49	47	50
	18	9	19	29	15	26
$Q = 0$ only if a cycle is reversible.	*8*	3	7	9	10	12
When explicitly given that a cycle is reversible, $Q = 0$	*15*	32	62	75	75	76
$\mathbf{\Delta S = 0}$ **after a cycle. (correct)**	***7***	**41**	**53**	**55**	**43**	**62**
$\Delta S \neq 0$ after a cycle. (additional question 19*)	*7*	59	47	45	57	38
$\Delta S = 0$ only if a cycle is reversible.	*7*	9	13	10	22	5
When explicitly given that a cycle is reversible, $\Delta S \neq 0$	*15*	35	28	32	26	27

Table 26: Prevalence of Difficulties Related to State Variables for each item

Difficulties	Item #	Prevalence (%) by Level				
		Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Not treating E as a state variable.	*5*	26	32	39	50	37
	15	10	19	31	25	23
	22	3	7	8	11	6
Treating W as though it is a state variable.	*6*	9	37	56	47	53
	10	11	23	27	27	28
	22	7	32	39	35	33
Treating Q as though it is a state variable.	*8*	11	33	40	37	38
	10	11	30	40	36	41
	15	32	62	75	75	76
	18	9	19	29	15	26
Not treating S as a state variable.	*7*	59	47	45	57	38
	15	35	28	32	26	27
	22	22	36	37	41	31
Treating all three quantities $\{E, Q, S\}$ or $\{E, W, S\}$ as state variables.	*15*	22	43	43	50	53
	22	0	12	19	11	18

Table 27: Prevalence of Difficulties with Internal Energy (E) for each item

Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
Not recognizing that E is proportional to T for an ideal gas. (additional questions 9*, 28*)	*17*	28	42	50	52	45
	27	14	42	38	38	39
	30	38	63	65	77	69
$\Delta E \neq 0$ for an ideal gas that undergoes an isothermal process.	*27*	14	42	38	38	39
	28	34	42	48	42	40
$\Delta E \neq 0$ for an isolated system.	*11*	13	20	28	23	37
	14	19	21	33	22	40
	23	14	26	31	33	36
	31	17	36	37	36	46
$\Delta E = 0$ in an isochoric process	*17*	14	23	32	31	25
	21	17	26	33	33	34
	29	17	32	41	27	39
Incorrect sign of ΔE	*2*	27	47	52	40	58
Not treating E as a state variable.	*5*	26	32	39	50	37
	15	10	19	31	25	23
	22	3	7	8	11	6
$\Delta E \neq 0$ after a cycle.	*5*	26	32	39	50	37

Table 28: Prevalence of Difficulties with work done by a system (W) for each item. Bold text indicates a correct response.

		Prevalence (%) by Level				
		Upper	Calc	Calc	Algebra	Algebra
Correct Answer, Difficulties	Item #	Post	Post	Pre	Post	Pre
$W \neq 0$ for various spontaneous processes.	*14*	37	53	62	53	64
	23	27	32	39	35	45
	31	36	38	41	31	42
$W \neq 0$ for a spontaneous process with only particle transfer	*14*	37	53	62	53	64
	31	36	38	41	31	42
$W \neq 0$ for a spontaneous process with only heat transfer	*23*	27	32	39	35	45
$W \leq 0$ in an isothermal expansion	*27*	10	22	33	25	39
$W = 0$ in an isobaric process	*21*	8	18	33	21	28
Not recognizing W as area under the curve in a PV diagram (additional questions 6*, 10*, 15*, 21*)	*9*	24	32	36	30	34
	26	35	59	83	69	83
When comparing an isothermal expansion to an adiabatic expansion on a PV diagram where both start in the same state and both have the same ΔV:						
$\mathbf{W_{isothm.} > W_{adi}}$ **(correct)**	***26***	**65**	**41**	**17**	**31**	**17**
$\lvert\Delta P_{adi}\rvert > \lvert\Delta P_{isothm.}\rvert \Rightarrow W_{adi} > W_{isothm.}$	*26*	11	19	25	26	27
$\lvert\Delta T_{adi}\rvert > \lvert\Delta T_{isothm.}\rvert \Rightarrow W_{adi} > W_{isothm.}$	*26*	12	15	23	20	27
$\Delta V_{adi} = \Delta V_{isothm.} \Rightarrow W_{adi} = W_{isothm.}$	*26*	10	18	22	14	18
Incorrect sign of W (additional question 27*)	*2*	21	32	31	32	26
	6	27	44	64	76	72
Treating W as though it is a state variable.	*6*	9	37	56	47	53
	10	11	23	27	27	28
	22	7	32	39	35	33
$W = 0$ for a cycle.	*6*	14	43	64	53	61

Table 29: Prevalence of difficulties with heat transfer to a system (Q) for each item. Bold text indicates a correct response.

Correct Answer, Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
$\mathbf{Q = 0}$ in an adiabatic process *(correct)*	***18***	**88**	**64**	**51**	**76**	**65**
$Q = 0$ for a gas heated in a closed container (additional question 28*)	*20*	23	39	38	31	40
$Q = 0$ in an isothermal process (additional questions 4*, 28*)	*3*	18	38	24	33	21
	18	16	44	50	39	37
	27	26	66	75	65	72
$Q = 0$ in a *reversible isothermal process (additional questions* 4*, 28*)*	*3*	18	38	24	33	21
$Q = 0$ in a reversible cycle (additional question 8*)	*15*	32	62	75	75	76
$Q = 0$ in an isochoric process	*21*	12	26	32	24	29
Incorrect sign of Q (additional questions 9*, 10*)	*3*	31	44	51	37	53
	8	47	56	67	71	63
Treating Q as though it is a state variable.	*8*	11	33	40	37	38
	10	11	30	40	36	41
	15	32	62	75	75	76
	18	9	19	29	15	26
$Q_1 = Q_2$ if two processes share initial and final states on a PV diagram.	*10*	22	53	67	63	70
$Q = 0$ for a cycle.	*8*	14	40	49	47	50
	18	9	19	29	15	26

Table 30: Prevalence of incorrect use of a compensation heuristic in first law problems for each item.

Difficulties	Item #	Prevalence (%) by Level Upper Post	Calc Post	Calc Pre	Algebra Post	Algebra Pre
An increase in one quantity implies a decrease in another quantity. (additional question 14*)	*2*	10	16	14	15	13
	9	34	40	42	37	36
	10	19	14	9	8	3
Two of the three quantities in the first law are zero in a process. (additional question 29*)	*17*	11	18	23	24	18
	21	30	51	65	56	63
	27	6	13	15	11	8

Table 31: Prevalence of difficulties with entropy change in reversible processes for each item. Bold text indicates a correct statement, which was given as part of a response. S_U is the entropy of the universe and S_{system} is the entropy of the system under consideration.

		Prevalence (%) by Level				
		Upper	Calc	Calc	Algebra	Algebra
Correct Answer, Difficulties	Item #	Post	Post	Pre	Post	Pre
Reversible Adiabatic Process $\Rightarrow \mathbf{\Delta S_{system} = 0}$ ***(correct)***	***1***	**24**	**28**	**19**	**29**	**24**
Reversibility (alone) $\Rightarrow \Delta S_{system} = 0$	*1*	39	29	16	36	19
$\Delta S_{system} > 0$	*1*	28	32	51	29	46
Reversible Isothermal Compression $\mathbf{\Delta S_{system} < 0}$ ***(correct)***	***4***	**31**	**28**	**37**	**32**	**44**
$\Delta S_{system} = 0$	*4*	42	36	25	43	20
$\Delta S_{system} > 0$	*4*	23	26	23	12	19
Reversible Isothermal Expansion $\mathbf{\Delta S_{System} > 0}$ ***and*** $\mathbf{\Delta S_U = 0}$ ***(correct)***	***28***	**34**	**40**	**36**	**42**	**27**
$[\Delta S_{System} = 0]$ and $\Delta S_U = 0$	*28*	26	27	23	30	33
$\Delta S_{system} > 0$ and $\Delta S_U > 0$	*28*	40	33	41	28	40

Table 32: Prevalence of difficulties with entropy change in irreversible processes for each item. Bold text indicates a correct statement, which was given as part of a response. Statements in brackets are the interpretation of student response informed by interview. For instance, "[$\Delta S_{Isolated} = 0$]" represents a case where students do not select $\Delta S_{Isolated} > 0$.

		Prevalence (%) by Level				
		Upper	Calc	Calc	Algebra	Algebra
Correct Answer, Difficulties	Item #	Post	Post	Pre	Post	Pre
Not all processes have $\mathbf{\Delta S_{Isolated} = 0}$ (correct)	***16***	**58**	**27**	**19**	**27**	**16**
$\Delta S_{Isolated} = 0$ for all processes	*16*	42	73	81	73	84
Irreversible Isochoric Process **$\mathbf{\Delta S_{Isolated} > 0}$, $\mathbf{\Delta S_{System} > 0}$. (correct)**	***29***	**68**	**43**	**35**	**57**	**40**
$\Delta S_{Isolated} = 0$, $\Delta S_{System} > 0$.	*29*	16	31	30	20	28
$\Delta S_{Isolated} = 0$, [$\Delta S_{System} = 0$]	*29*	9	8	15	10	17
$\mathbf{\Delta S_{Isolated} > 0}$ for a spontaneous process (correct)	***12***	**81**	**55**	**50**	**57**	**57**
	14	**89**	**77**	**69**	**79**	**72**
	24	**83**	**53**	**50**	**51**	**48**
	32	**83**	**60**	**54**	**63**	**59**
[$\Delta S_{Isolated} = 0$] for a spontaneous process	*12*	19	45	50	43	43
	14	11	23	31	21	28
	24	17	47	50	49	52
	32	17	40	46	37	41
Free Expansion Process **$\mathbf{\Delta S_{Isolated} > 0}$ (correct)**	***14***	**89**	**77**	**69**	**79**	**72**
$\Delta S_{Isolated} = 0$	*14*	11	23	31	21	28
Spontaneous Heat Transfer **$\mathbf{\Delta S_{Isolated} > 0}$ and $\mathbf{\Delta S_{Cold} > 0}$ *(correct)***	***12***	**62**	**31**	**24**	**32**	**29**
	24	**64**	**32**	**24**	**31**	**30**
[$\Delta S_{Isolated} = 0$] and **$\mathbf{\Delta S_{Cold} > 0}$**	*12*	17	40	44	41	39
	24	15	42	39	41	43
$\mathbf{\Delta S_{Isolated} > 0}$ and [$\Delta S_{Cold} = 0$]	*12*	6	10	13	12	11
	24	5	9	13	9	11
$\mathbf{\Delta S_{Isolated} > 0}$ and **$\mathbf{\Delta S_{Cold} > 0}$** but also $\Delta S_{Hot} > 0$	*12*	13	15	13	12	17
	24	14	12	12	11	7
Spontaneous Mixing **$\mathbf{\Delta S_{Isolated} > 0}$ and each $\mathbf{\Delta S_{subsys} > 0}$**	***32***	**62**	**37**	**33**	**40**	**27**
$\Delta S_{Isolated} > 0$ but each [$\Delta S_{subsys} = 0$]	*32*	21	23	21	22	31
[$\Delta S_{Isolated} = 0$] and each [$\Delta S_{subsys} = 0$]	*32*	8	21	23	18	16
[$\Delta S_{Isolated} = 0$] but each $\Delta S_{subsys} > 0$	*32*	6	13	15	15	13
Second law *(correct)*	***13***	**88**	**67**	**66**	**75**	**62**
	25	**89**	**61**	**55**	**75**	**56**
	33	**88**	**62**	**52**	**67**	**42**
First law	*13*	23	53	59	35	53
	25	31	56	65	40	56
	33	23	50	54	45	53
Newton's 2^{nd} law	*13*	3	9	11	7	14
	25	2	13	17	8	23
	33	6	17	23	18	39

APPENDIX F

CODE TO ANALYZE STPFASL

Here is the code that was used to analyze the prevalence of difficulties for STPFaSL. It produced all the results, and typeset most of the tables, as well.

```
#!/usr/bin/env python
import sys
from numpy import *
from numpy.lib import recfunctions
from numpy.core.defchararray import isalpha
from scipy.stats.mstats import pearsonr
import re

PossibleResponses = ['A', 'B', 'C', 'D', 'E']
itemsList = [str(i) for i in range(1,33+1)]

catalog = load('consolidated3.npy')
STPFaSLkey = catalog[itemsList][0]

itemsArr = array(itemsList)

latexdelim = '&'
latexEOL = 2*'\\'
nonlevelswidth =9
itemswidth = 1.5
diffwidth = nonlevelswidth - itemswidth - 0.1

class Topic:
  def __init__(self, name="", items=itemsArr):
    self.name=name
    self.items=items
  def Sample(self, sample):
```

```
    return Sample(" ".join([sample.name, self.name]), sample.cat,
        self.items, key=catalog[self.items][0])


class Difficulty:
  def __init__(self, name, items=[], includeStatistics=True):
    self.name = name
    self.items = items

  def printSummaryRow(self, samples, displayWide=False):
    isCorrect = all([i.includesCorrect for i in self.items])

    #Start with the difficulties name:
    rowname = self.name
    #if not isCorrect: rowname = "{\\noindent\\hspace{+1ex}"+
        rowname + "}"
    row_strings = [rowname]

    #The items
    if not displayWide: row_strings.append(", ".join([re.sub("
        ([^0-9\*]*)", "", i.num) for i in self.items]))

    #The percentages
    for s in samples:
      prevalence=[]
      for i in self.items:
        if not "*" in i.num: prevalence.append(
            proportionresponding(s, i))
      row_strings.append(per_str(array(prevalence).mean()))

    #if we're a correct answer, bold everything in the row.
    if isCorrect:
      row_strings = ["\\bfseries{"+s+"}" for s in row_strings]

    print latexdelim.join(row_strings) + latexEOL

  def printDetailedRows(self, samples):
    regular_items = [i for i in self.items if not "*" in i.num]
    starred_items = [i for i in self.items if "*" in i.num]

    numItems = len(regular_items)
    column1 = self.name
    if (len(starred_items)>0):
      column1 += " (additional question"
      if (len(starred_items)>1): column1 += "s"
```

```
    column1 += " " + ", ".join([s.num for s in starred_items]) +
        ")"

  if numItems > 1:
    column1 = "\\multirow{" + str(numItems) + "}{"+str(diffwidth
      )+"cm}{" + column1 + "}"
  for i in regular_items:
    row_strings = [column1, i.num]
    for s in samples:
      row_strings.append(per_str(proportionresponding(s, i)))
    if i.includesCorrect: row_strings = ["\\bfseries{"+s+"}" for
        s in row_strings]
    print latexdelim.join(row_strings) + latexEOL
    column1 = ""
  print "\\hline"

def proportionresponding(sample, sItem):
  responseslist = sample.cat[sItem.safenum]#level[item.safenum]
  matchingresponses = zeros_like(1*(responseslist==''))
  validresponses = zeros_like(1*isalpha(responseslist))
  for r in sItem.vsresponse:
    validresponses += 1*(responseslist==r)

  for r in sItem.response:
    matchingresponses += 1*(responseslist==r)

  return float(matchingresponses.sum())/float(validresponses.sum()
    )

class surveyItem:
  def __init__(self, itemNum, response, vsresponse=
      PossibleResponses):
    self.num=itemNum #For printing
    self.safenum=re.sub("([^0-9]*)", "", itemNum)#Digits Only,
        String Representation

    self.response=response
    self.vsresponse=vsresponse
    ###Reimplement next line for Sample?
    self.includesCorrect = any([r in STPFaSLkey[self.safenum] for
        r in response])
    self.isStar = "*" in self.num

  def responseRate(self, sample):
    return proportionresponding(sample, self)
```

```
class Sample:
  def __init__(self, name = "", cat=None, items=itemsList, key=
     STPFaSLkey):
    s=self
    s.name  = name
    s.items = items#Strings for slicing Slices
    s.cat= cat
    s.catArr = array([[response for response in survey] for survey
        in s.cat[s.items]], dtype='S5')
    (s.numStudents, s.numItems) = s.catArr.shape
    s.key   = array([k for k in key], dtype='S5')
    s.scores_raw = 1.0 * (s.catArr==s.key)
    s.valid=isalpha(s.catArr)
    s.scores = ma.masked_where(logical_not(s.valid), s.scores_raw)
    axes={'oStudents':0, 'oItems':1}

    #Dimension #0 is over students, dimension #1 is over items
    s.itemScores = s.scores.mean(0) #Called item difficulty
    s.studentScores = s.scores.mean(1) #Student's Score (Fraction)
    s.score = s.studentScores.mean()
    s.median = median(s.studentScores)
    slicehigh = s.studentScores>s.median
    slicelow = s.studentScores<s.median
    # s.numDiscStudents = ((1*slicehigh)+(1*slicelow)).sum()
    s.studentCompleted = (1*(s.scores!=2.0)).sum(1)
    s.meanItemsCompleted = s.studentCompleted.mean()

    s.studentPnts = s.studentScores*s.meanItemsCompleted

    s.pbc = array([pearsonr(s.scores[:, i], s.studentScores)[0]
       for i in range(0,s.numItems)])
    # s.disc     = array([((s.scores[slicehigh][i]==1).sum()-(s.
       scores[slicelow][i]==1).sum())/(s.numDiscStudents/2.0) for
       i in range(0, s.numItems)])
    # s.antidisc = array([((s.scores[slicelow][i]==0).sum()-(s.
       scores[slicehigh][i]==0).sum())/(s.numDiscStudents/2.0) for
        i in range(0, s.numItems)])

    s.pbc[isnan(s.pbc)]=1.0 #Edge case when an entire sample has
       the same score for an item.
    (s.kr8, s.kr20) = s.kr()
```

```
    ##NOte that indexing s.cat (the array) as:
    ##up.cat[:, [2,3,4]], would select numbers 3,4,5 (=[2,3,4 ]+
        1))
    #correlation coefficients does not honor masking:corrcoef(wat.
        scores[1:, :].T, wat.scores[1:].mean(axis=1).T, rowvar=1)
  def kr(s):
    k = s.meanItemsCompleted
    sig_x = k * s.studentScores.std() #units are points
    sig_i = s.scores.std(axis = 0)
    var_x = sig_x*sig_x #Elementwise, of course only 1 element
    var_i = sig_i*sig_i
    #for kr8
    minusb = (1.0 - var_i.sum()/var_x)/2.0
    rad = sqrt((s.pbc*s.pbc * var_i).sum()/var_x + minusb*minusb)
    return (minusb+rad, (k/(k-1.0)) * (1.0-(var_i.sum()/var_x)))

  def __str__(self):
    oldprintops = get_printoptions()
    #set_printoptions(precision=3, suppress = True)
    set_printoptions(formatter={'float': '{: 3.2f}'.format})
    thestring = "\n".join([
      "****************************",
      "Name : " +self.name,
      "Catalog of " + str(self.numItems) + " questions, and " +
         str(self.numStudents)+" students.",
      "With Answer Key: " + ", ".join(self.key),
      "Mean: " + per_str(self.scores.mean()) + "%",
      "Stdv: " + nicefloat(self.scores.std()),
      "NormStd: " + nicefloat(self.studentScores.std()*self.
         meanItemsCompleted),
      "Or, in Points: ",
      "Mean: " + nicefloat(self.scores.sum(axis=1).mean()),
      "Stdv: " + nicefloat(self.scores.sum(axis=1).std()),
      "KR20: " + nicefloat(self.kr20),
      "KR-8: " + nicefloat(self.kr8),
      "PBC's:",
      str(array(self.pbc)),
      "Item Performance:",
      str(array(self.itemScores)),
      "****************************",
      ])
    #set_printoptions(oldprintops)
    set_printoptions(edgeitems=3,infstr='inf', linewidth=75,
       nanstr='nan', precision=8,suppress=False, threshold=1000,
       formatter=None)
```

```
    return thestring

def nicefloat(x):
  return "{: 3.2f}".format(x)

def printTexDoc(tables=None, tabletype="Summary"):
  print "\documentclass[pdftex, final]{pittetd}"
  print "\usepackage{units}"
  print "\usewithpatch{graphicx}"
  print "\usepackage{amsmath,amsthm, amssymb}"
  print "\usepackage{braket}"
  print "\usepackage[normalem]{ulem}"
  print "\usepackage{booktabs}"
  print "\usepackage{tabularx}"
  print "\usepackage{tabulary}"
  print "\usepackage{multirow}"
  print "\usepackage{bigstrut}"
  print "\usepackage{longtable}"

  print "\usepackage{pdflscape}"
  print "\usepackage{array}"
  print "\usepackage{subfiles}"
  print "\usepackage[numbers,sort&compress]{natbib}"
  print "\usepackage[table]{xcolor}"
  print "\usepackage[super]{nth}"
  print "\usepackage{hhline}"
  print ""
  print "\usepackage{etoolbox}"

  print "\makeatletter"
  print "\patchcmd{\@classz}"
  print "  {\CT@row@color}"
  print "  {\oldCT@column@color}"
  print "  {}"
  print "  {}"
  print "\patchcmd{\@classz}"
  print "  {\CT@column@color}"
  print "  {\CT@row@color}"
  print "  {}"
  print "  {}"
  print "\patchcmd{\@classz}"
  print "  {\oldCT@column@color}"
  print "  {\CT@column@color}"
  print "  {}"
  print "  {}"
```

```
print "\makeatother"

print "\patch{amsmatch}"
print "\patch{amsthm}"
print "\\let\\Tabular\\tabular"
print "\\def\\tabular{\\global\\rownum=0\\relax\\Tabular}"
print "\\newcommand{\inlinecomment}[1]{}"
print "\\newcommand{\op}[1]{\ensuremath{\hat{#1}}}"
print "\\newcommand{\commutator}[2]{\ensuremath{[#1, #2]}}"
print "\\newcommand{\up}{\ket{\uparrow}}"
print "\\newcommand{\down}{\ket{\downarrow}}"
print "\\newcommand{\\x}{\\displaystyle x}"
print "\\newcommand{\\y}{\\displaystyle y}"
print "\\newcommand{\\z}{\\displaystyle z}"
print "\\newcolumntype{Y}{>{\centering\\arraybackslash}X}"
#print "\\newcolumntype{M}[1]{m{#1}}"
print "\\newcolumntype{M}[1]{>{\\raggedright}m{#1}}"
print "\\newcolumntype{B}[1]{>{\\noindent\\raggedright}b{#1}}"
print "\\newcolumntype{P}[1]{>{\\raggedright}p{#1}}"
print "\\newcommand{\\abs}[1]{\\lvert #1 \\rvert}"
print "\\newcommand{\\adiabat}{adi}"
print "\\newcommand{\\isotherm}{isothm.}"

print "\\renewcommand{\\arraystretch}{1.05}"
print "\\setlength{\\bigstrutjot}{6pt}"
print "\\begin{document}"

#if tabletype in "blocktables":
  # print "\\begin{center}"
  # print "\LTcapwidth=\\textwidth"
  # print "\\begin{longtable}{r|r|rrrr|l}"
  # print "      \cline{2-6}"
  # print "       {Problem \#} & {\\textbf{Correct} (\%)} & {$1^{st}$}& {$2^{nd}$} & {$3^{rd}$} & {$4^{th}$} & {Level}\\\\*"
  # print "       \cline{2-6}"
  # print "       \endfirsthead"
  # print "     \cline{2-6}"
  # print "  {Problem \#} & {\\textbf{Correct}} & {$1^{st}$}& {$2^{nd}$} & {$3^{rd}$} & {$4^{th}$}& {Level}\\\\*"
  # print "     \cline{2-6}"
  # print "     \endhead"
  # print "    \multicolumn{6}{r}{\small\sl continued on next page}\\\\*"
  # print "     \endfoot"
  # print " \endlastfoot"
```

```
    # for t in tables:
    #   t.printBlockTable()
    # print "\end{longtable}"
    # print "\end{center}"
  #else:
  for t in tables:
    t.printTable(tabletype)

  print "\end{document}"


class lTable:
  def __init__(self, name, difficulties=None, displayWide=False,
      itemswidth = itemswidth):
    self.name = name
    self.displayWide = displayWide
    self.itemswidth=itemswidth
    if(difficulties):
      self.difficulties = difficulties # list of Difficultys
    else:
      self.difficulties = []
    self.longCaption=name
    self.labelString = "".join(self.name.split()[-3:]).translate(
       None, ".")

  def setCorrect(self):
    realdiffs = [d for d in self.difficulties if not isinstance(d,
        basestring)]
    self.includesCorrect = any([si.includesCorrect for d in
       realdiffs for si in d.items ])

  def printLevelsHeader(self, samples, displayWidearg=None):
    if displayWidearg==None:
      displayWide = self.displayWide
    else:
      displayWide = displayWidearg

    #hdr_strings = ["{\\bfseries Correct Answer}"]
    #if not displayWide: hdr_strings.append("Item ")
    #hdr_strings += ["{\\footnotesize " + s.name.split(' ', 1)[0]
       + "}" for s in samples]
    # for s in samples:
    #   hdr_strings.append("{\\footnotesize " + s.name.split(' ',
       1)[0] + "}")
    #hdr_strings += [latexEOL, "\n"]
```

```
    #hdr_strings +=["Difficulties"]

    #print "\\hline"
    column1title = "Difficulties"
    if self.includesCorrect: column1title = "{\\bfseries Correct Answer}, " + column1title
    print latexdelim.join(["\\multicolumn{1}{l}{}", "\\multicolumn{1}{r|}{}"] + ["{\\footnotesize " + s.name.split(' ', 1)[0] + "}" for s in samples]) + latexEOL
    print latexdelim.join(["\\multicolumn{1}{l}{"+column1title+"}", "\multicolumn{1}{r|}{Item \#}"] + ["{\\footnotesize " + s.name.split(' ', 1)[-1] + "}" for s in samples]) + latexEOL

    #print latexdelim.join(["", "Item"] + ["{\\footnotesize " + s.name.split(' ', 1)[0] + "}" for s in samples]) + latexEOL
    #print latexdelim.join(["{\\bfseries Correct Answer}, Difficulties", "\#"] + ["{\\footnotesize " + s.name.split(' ', 1)[-1] + "}" for s in samples]) + latexEOL
    print "\\hline"

    #print latexdelim.join(hdr_strings) + latexEOL
    #print "\\hline"

  def caption(self, longCaption=None):
    if longCaption: self.longCaption = longCaption
    return self.longCaption

  def label(self, labelString=None):
    if(labelString): self.labelString=labelString
    return self.labelString

  def tabularHead(self, tabularHeadString=None):
    """Sets and returns a head defining the columns"""
    if(tabularHeadString):
      self.tabularHeadString=tabularHeadString
    elif self.displayWide:
      self.tabularHeadString="\\begin{tabulary}{\\linewidth}{|B{" + str(nonlevelswidth) + "cm}|CCCCC|}"
    else:
      self.tabularHeadString="\\begin{tabulary}{\\linewidth}{ | >{\\noindent}B{" + str(nonlevelswidth-self.itemswidth-0.1) + "cm} | >{\\itshape\\footnotesize}M{"+str(self.itemswidth)+"cm}|CCCCC|}"
    return self.tabularHeadString
```

```
def printSummaryTabulary(self):
  print self.tabularHead()
  if self.displayWide:
    print "\\cline{2-6}"
    print "\\multicolumn{1}{l␣|}{}␣&␣\\multicolumn{5}{␣c␣|}{
       Prevalence␣(\\%)␣by␣Level}\\\\"
  else:
    print "\\cline{3-7}"
    print "\\multicolumn{2}{l␣|}{}␣&␣\\multicolumn{5}{␣c␣|}{
       Prevalence␣(\\%)␣by␣Level}\\\\"
  self.printLevelsHeader(samples)
  for d in self.difficulties:
    if(isinstance(d, basestring)):
      print d
    else:
      d.printSummaryRow(samples, self.displayWide)
  print "\\hline"
  print "\\end{tabulary}"

def printDetailedTabulary(self):
  print self.tabularHead("\\begin{tabulary}{\\linewidth}{|␣M{" +
      str(diffwidth) + "cm}␣|␣>{\itshape}C␣|␣C␣C␣C␣C␣C␣|}")
  print "\\cline{3-7}"
  print "\\multicolumn{2}{l␣|}{}␣&␣\\multicolumn{5}{␣c␣|}{
     Prevalence␣(\\%)␣by␣Level}\\\\"
  self.printLevelsHeader(samples, displayWidearg = False)
  for d in self.difficulties:
    if(isinstance(d, basestring)):
      print re.sub("multicolumn\{6\}", "multicolumn{7}", d)
    else:
      d.printDetailedRows(samples)
    #print "\\hline"
  print "\\end{tabulary}"

def printTable(self, tabletype="Summary"):
  print 80*"%"
  print "\\begin{table}"
  print "\\setlength{\\tabcolsep}{3.5pt}"
  if tabletype in "Appendix":
    self.labelString +="byItem"
    self.name += "␣for␣each␣item"
    self.caption(self.caption() + "␣for␣each␣item")
  print "\\label{"+self.label()+"}"
  print "\\caption["+self.name+"]{"+ self.caption() +"}"
```

```
    print "\small"
    print "\\noindent"

    if tabletype in "Summary":
      print "\\rowcolors{5}{}{gray!20}"

    if(tabletype in "Summary"):
      self.printSummaryTabulary()
    elif(tabletype in "Appendix"):
      self.printDetailedTabulary()
    else:
      print "%" + "No␣Table␣Data"

    print "\\end{table}"
    print 80*"%"

  def printSummaryTable(self):
    self.printTable("Summary")
  def printDetailedTabled(self):
    self.printTable("Detailed")

tables=[]

"""##################################################################"""



cyclicTable = lTable("Prevalence␣of␣Difficulties␣Related␣to␣Cyclic
    ␣Processes", itemswidth=1.2)
cyclicTable.difficulties.extend([
  Difficulty("$\\mathbf{\Delta␣E␣=␣0}$␣after␣a␣cycle.␣(correct)",
    (
      surveyItem('5', 'A'),
    )
  ),
  Difficulty("$\Delta␣E␣\\neq␣0$␣after␣a␣cycle.",
    (
      surveyItem('5', ('B', 'C', 'D', 'E')),
    )
  ),
  Difficulty("$\Delta␣E␣=␣0$␣only␣if␣a␣cycle␣is␣reversible.",
    (
      surveyItem('5', 'E'),
    )
  ),
```

```
Difficulty("When␣explicitly␣given␣that␣a␣cycle␣is␣reversible,␣$
    \\Delta␣E␣\\neq␣0$",
  (
    surveyItem('15', ('B', 'C')),
  )
),
"\\hline",
Difficulty("$\\mathbf{W␣\\neq␣0}$␣for␣a␣cycle.␣(correct)",
  (
    surveyItem('6', ('A', 'B')),
  )
),
Difficulty("$W␣=␣0$␣for␣a␣cycle.",
  (
    surveyItem('6', ('C', 'E')),
  )
),
Difficulty("$W␣=␣0$␣only␣if␣a␣cycle␣is␣reversible.",
  (
    surveyItem('6', 'E'),
  )
),
"\\hline",
Difficulty("$\\mathbf{Q␣\\neq␣0}$␣for␣a␣cycle.␣(correct)",
  (
    surveyItem('8', ('B', 'C')),
  )
),

Difficulty("$Q␣=␣0$␣for␣a␣cycle.",
  (
    surveyItem('8', ('A', 'E')),
    surveyItem('18', ('A', 'D')),
  )
),
Difficulty("$Q␣=␣0$␣only␣if␣a␣cycle␣is␣reversible.",
  (
    surveyItem('8', 'E'),
  )
),
Difficulty("When␣explicitly␣given␣that␣a␣cycle␣is␣reversible,␣$Q
    ␣=␣0$",
  (
    surveyItem('15', ('A', 'B', 'C')),
  )
```

```
  ),
  "\\hline",
  Difficulty("$\\mathbf{\Delta S = 0}$ after a cycle. (correct)",
    (
      surveyItem('7', 'C'),
    )
  ),
  Difficulty("$\Delta S \\neq 0$ after a cycle.",
    (
      surveyItem('7', ('A', 'B', 'D', 'E')),
      surveyItem('$19^{*}$', ''),
    )
  ),
  Difficulty("$\Delta S = 0$ only if a cycle is reversible.",
    (
      surveyItem('7', 'E'),
    )
  ),
  Difficulty("When explicitly given that a cycle is reversible, $\
     Delta S \\neq 0$",
    (
      surveyItem('15', ('C', 'D')),
    )
  ),

])
tables.append(cyclicTable)

"""##################################################################"""


stateReversibleTable = lTable("Prevalence of Difficulties Related 
   to State Variables")
stateReversibleTable.difficulties.extend([

  Difficulty("Not treating $E$ as a state variable.",
    (
      surveyItem('5', ('B', 'C', 'D', 'E')),
      surveyItem('15', ('B', 'C')),
      surveyItem('22', 'A'),
    )
  ),
  Difficulty("Treating $W$ as though it is a state variable.",
    (
      surveyItem('6', 'C'),
```

```
      surveyItem('10', 'A'),
      surveyItem('22', ('A','C','E')),
    )
  ),
  Difficulty("Treating $Q$ as though it is a state variable.",
    (
      surveyItem('8', 'A'),
      surveyItem('10', 'B'),
      surveyItem('15', ('A', 'B', 'C')),
      surveyItem('18', ('A', 'D')) #Also cyclic
    )
  ),
  Difficulty("Not treating $S$ as a state variable.",
    (
      surveyItem('7', ('A', 'B', 'D', 'E')),
      surveyItem('15', ('C', 'D')),
      surveyItem('22', ('A', 'B', 'C')),
    )
  ),
  Difficulty("Treating all three quantities \{$E$, $Q$, $S$\} or 
     \{$E$, $W$, $S$\} as state variables.",
    (
      surveyItem('15', 'A'),
      surveyItem('22', 'E'),
    )
  ),
  ])
tables.append(stateReversibleTable)

"""##################################################################"""

ETable = lTable("Prevalence of Difficulties with Internal Energy (
   $E$)", itemswidth=1.6)
ETable.difficulties.extend([
  Difficulty("Not recognizing that $E$ is proportional to $T$ for 
     an ideal gas.",
    (
      surveyItem('17', ('A', 'B', 'C', 'D')),
      surveyItem('27', ('A', 'C')),
      surveyItem('30', ('B', 'C', 'E')),
      surveyItem('$9^{*}$', ''),
      surveyItem('$28^{*}$', ''),
    )
  ),
```

```
  Difficulty("$\Delta E \\neq 0$ for an ideal gas that undergoes 
     an isothermal process.",
    (
      surveyItem('27', ('A', 'C')),
      surveyItem('28', ('A', 'B')),
    )
  ),
  Difficulty("$\Delta E \\neq 0$ for an isolated system.",
    (
      surveyItem('11', ('A', 'D')),
      surveyItem('14', ('B', 'D')),
      surveyItem('23', ('B', 'D')),
      surveyItem('31', ('B', 'D')),
    )
  ),
  Difficulty("$\Delta E = 0$ in an isochoric process",
    (
      surveyItem('17', ('A', 'B', 'D')),
      surveyItem('21', ('C', 'E')),
      surveyItem('29', ('C', 'D', 'E')),
    )
  ),
  "\\hline",
  Difficulty("Incorrect sign of $\Delta E$",
    (
      surveyItem('2', ('C', 'D'), ('A', 'B', 'C', 'D')),
    )
  ),
  "\\hline",
  Difficulty("Not treating $E$ as a state variable.",
    (
      surveyItem('5', ('B', 'C', 'D', 'E')),
      surveyItem('15', ('B', 'C')),
      surveyItem('22', 'A'),
    )
  ),
  Difficulty("$\Delta E \\neq 0$ after a cycle.",
    (
      surveyItem('5', ('B', 'C', 'D', 'E')),
    )
  ),
])
tables.append(ETable)
```

```
"""###########################################################"""

workTable = lTable("Prevalence of Difficulties with Work done by a
    system ($W$)")
workTable.difficulties.extend([
  Difficulty("$W \\ne 0$ for various spontaneous processes.",
    (
      surveyItem('14', ('B', 'C', 'D')),
      surveyItem('23', ('B', 'C')),
      surveyItem('31', ('B', 'C'))
    )
  ),
  Difficulty("$W \\ne 0$ for a spontaneous process with only
      particle transfer",#\multicolumn{1}{| r |}
    (
      surveyItem('14', ('B', 'C', 'D')),
      surveyItem('31', ('B', 'C')),
    )
  ),
  Difficulty("$W \\ne 0$ for a spontaneous process with only heat
      transfer",#\multicolumn{1}{| r |}
    (
      surveyItem('23', ('B', 'C')),
    )
  ),
  "\\hline",
  Difficulty("$W \\leq 0$ in an isothermal expansion",
    (
      surveyItem('27', ('A', 'B')),
    )
  ),
  "\\hline",
  Difficulty("$W = 0$ in an isobaric process",
    (
      surveyItem('21', ('A', 'C')),
    )
  ),
  "\\hline",
  "\\hline",
  Difficulty("Not recognizing $W$ as area under the curve in a PV
      diagram",
    (#"\\setlength{\\bigstrutjot}{em}  "
      surveyItem('9\\bigstrut', ('A', 'C')),
      surveyItem('26\\bigstrut', ('A', 'B', 'C', 'D')),
```

```
      surveyItem('$6^{*}$', 'C'), #Clear
      surveyItem('$10^{*}$', ('A', 'C')), #Clear
      surveyItem('$15^{*}$', ('A', 'B', 'C')),#Clear?
      surveyItem('$21^{*}$', ('A', 'C', 'E')),#Clear.
    )
  ),
  #"\\hiderowcolors ",
  "\\hline",
  "\multicolumn{7}{|p{0.9\linewidth}|}{\itshape When comparing an 
     isothermal expansion to an adiabatic expansion on a PV 
     diagram where both start in the same state and both have the 
     same $\Delta V$:}\\\\",
  "\\hline",
  #"\\showrowcolors\\",
  Difficulty("$\mathbf{W_{\\isotherm} > W_{\\adiabat}}$ (correct)"
     ,
    (
      surveyItem('26', 'E'),
    )
  ),
  Difficulty("$\\abs{\Delta P_{\\adiabat}} > \\abs{\Delta P_{\\
     isotherm}} \\Rightarrow W_{\\adiabat} > W_{\\isotherm}$",
    (
      surveyItem('26', 'B'),
    )
  ),
  Difficulty("$\\abs{\Delta T_{\\adiabat}} > \\abs{\\Delta T_{\\
     isotherm}} \\Rightarrow W_{\\adiabat} > W_{\\isotherm}$",
    (
      surveyItem('26', 'D'),
    )
  ),
  Difficulty("$\\Delta V_{\\adiabat} = \\Delta V_{\\isotherm} \\
     Rightarrow W_{\\adiabat} = W_{\\isotherm}$",
    (
      surveyItem('26', 'C'),
    )
  ),

  # Difficulty("$W_{\\isotherm} > W_{\\adiabat}$ because $\Delta
     V_{\\adiabat} = \Delta V_{\\isotherm}$.",
  #   (
  #     surveyItem('26', 'A'),
  #   )
  # ),
```

```
  "\\hline",
  "\\hline",
  Difficulty("Incorrect␣sign␣of␣$W$",
    (
      surveyItem('2', ('B', 'D'), ('A', 'B', 'C', 'D')),
      surveyItem('6', 'A', ('A', 'B')),
      surveyItem('$27^{*}$', ''),
    )
  ),
  "\\hline",
  Difficulty("Treating␣$W$␣as␣though␣it␣is␣a␣state␣variable.",
    (
      surveyItem('6', 'C'),
      surveyItem('10', 'A'),
      surveyItem('22', ('A','C','E')),
    )
  ),
  Difficulty("$W␣=␣0$␣for␣a␣cycle.",
    (
      surveyItem('6', ('C', 'E')),
    )
  ),

])

tables.append(workTable)
"""##########################################################"""

heatTable = lTable("Prevalence␣of␣Difficulties␣with␣Heat␣transfer␣
   to␣a␣system␣($Q$)", itemswidth=1.2)
number18adiabatic = surveyItem('18', ('B', 'D', 'E'))
number18adiabatic.includesCorrect = True
heatTable.difficulties.extend([
  Difficulty("$\mathbf{Q␣=␣0}$␣in␣an␣adiabatic␣process␣{\itshape␣(
     correct)}",
    (number18adiabatic, )
  ),
  Difficulty("$Q=0$␣in␣an␣isothermal␣processes",
    (
      surveyItem('3', 'C'),
      surveyItem('$4^{*}$', 'C'),
      surveyItem('18', ('C', 'E')),
      surveyItem('27', ('A', 'B', 'C', 'E')),
    )
  ),
```

```
Difficulty("$Q=0$ in an isochoric process",
  (
    surveyItem('21', ('A', 'D')),
  )
),
Difficulty("$Q=0$ for a gas which is heated in a closed
   container",
  (
    surveyItem('20', ('C', 'E')),
    surveyItem('$28^{*}$', 'D'),
  )
),
"\\hline",
"\\hline",
Difficulty("Incorrect sign of $Q$",
  (
    surveyItem('3', 'A', ('A', 'B')),
    surveyItem('8', 'B', ('B', 'C')),
    surveyItem('$9^{*}$', ''),
    surveyItem('$10^{*}$', ''),
  )
),
"\\hline",
"\\hline",
Difficulty("Treating $Q$ as though it is a state variable.",
  (
    surveyItem('8', 'A'),
    surveyItem('10', 'B'),
    surveyItem('15', ('A', 'B', 'C')),
    surveyItem('18', ('A', 'D')) #Also cyclic
  )
),
Difficulty('$Q_1 = Q_2$ if two process share initial and final
   states on a PV diagram.',
  (
    surveyItem('10', ('A', 'B')),
  )
),
Difficulty("$Q=0$ for a cycle.",
  (
    surveyItem('8', ('A', 'E')),
    surveyItem('18', ('A', 'D')),
  )
),
```

```
])

tables.append(heatTable)

"""############################################################"""


compensationTable = lTable("Prevalence of Incorrect Use of 
    Compensation Heuristic in First Law Problems")
compensationTable.difficulties.append(
    Difficulty("An increase in one quantity implies a decrease in 
        another quantity.",
        (
            surveyItem('2', 'D'),
            surveyItem('9', ('B', 'C')),
            surveyItem('10', 'D'),
            surveyItem('$14^{*}$', ''),
        )
    ))
compensationTable.difficulties.append(
    Difficulty("Two of the three quantities in the first law are 
        zero in a process.",
        (
            surveyItem('17', ('A', 'B')),
            surveyItem('21', ('A', 'C', 'D', 'E')),
            surveyItem('27', 'E'),
            surveyItem('$29^{*}$', ''),
        )
    ))
tables.append(compensationTable)

"""############################################################"""


# fns = "{\\footnotesize"
# entropyTable = lTable("Prevalence of Difficulties with Entropy
    of a System ($S$)")
# entropyTable.difficulties.extend([
#     Difficulty("$[\Delta S_{Isolated} = 0]$ or $\Delta S_{Isolated
    } = 0$ for an irreversible process.",
#         (
#             surveyItem(fns+'12(S)'+"}", ('A', 'B')),
#             surveyItem(fns+'14(S)'+"}", ('A', 'B')),
#             surveyItem(fns+'16(G)'+"}", ('A', 'B', 'C', 'E')),
#             surveyItem(fns+'24(S)'+"}", ('A', 'B')),
```

```
#           surveyItem(fns+'32(S)'+"}", ('B', 'D', 'E')),
#       )
#   ),
#   Difficulty("Not treating $S$ as a state variable.",
#       (
#           surveyItem(fns+'7(GC)'+"}", ('A', 'B', 'D', 'E')),
#           surveyItem(fns+'15(RC)'+"}", ('C', 'D')),
#           surveyItem(fns+'22(D)'+"}", ('A', 'B', 'C')),
#       )
#   ),
#   Difficulty("$S$ is a state variable, but only for reversible
    processes.",
#       (
#           surveyItem(fns+'7(GC)'+"}", 'E'),
#       )
#   ),
# ])
# tables.append(entropyTable)


"""##########################################################"""

entropyReversibleTable = lTable("Prevalence of Difficulties with 
    Entropy change in Reversible Processes", itemswidth=1.2)
entropyReversibleTable.difficulties.extend([
  #"\\rowcolor{gray!20}",
  #"\\hline",
  #"\\rowcolor{white}",
  "\\itshape{Reversible Adiabatic Process}\\\\",
  Difficulty("$\Rightarrow \mathbf{\Delta S_{system}=0}$ (correct)
      ",
    (
      surveyItem('1', 'C'),
    )
  ),
  Difficulty("Reversibility (alone) $\Rightarrow \Delta S_{system
      }=0$",
    (
      surveyItem('1', 'D'),
    )
  ),
  #"\\rowcolor{white}",
  Difficulty("$\Delta S_{system}>0$",
    (
      surveyItem('1', 'A'),
```

```
    )
),
"\\hline",
"\\hline",
#"\\rowcolor{gray!20}",
#"\\multicolumn{6}{| l |}{Item 4: Reversible Isothermal
    Compression}\\\\",
"\\itshape{Reversible Isothermal Compression}\\\\",
Difficulty("$\mathbf{\Delta S_{system} < 0}$ (correct)",
    (
        surveyItem('4', 'B'),
    )
),
#"\\hline",
#"\\rowcolor{white}",
Difficulty("$\Delta S_{system}=0$",
    (
        surveyItem('4', 'C'),
    )
),
#"\\rowcolor{white}",
Difficulty("$\Delta S_{system}>0$",
    (
        surveyItem('4', 'A'),
    )
),
"\\hline",
"\\hline",
#"\\rowcolor{gray!20}",
#"\\multicolumn{6}{| l |}{Item 28, $\mathbf{\Delta S_{system} >
    0}$ in a Reversible Isothermal Expansion}\\\\",
#"\\rowcolor{white}",
"\\itshape{Reversible Isothermal Expansion}\\\\",
Difficulty("$\mathbf{\Delta S_{System} > 0}$ and $\mathbf{\Delta
    S_{U}=0}$ (correct)",
    (
        surveyItem('28', ('B', 'E')),
    )
),
Difficulty("$[\Delta S_{System}=0]$ and $\Delta S_{U}=0$",
    (
        surveyItem('28', 'D'),
    )
),
#"\\rowcolor{white}",
```

```python
    Difficulty("$\Delta S_{system}>0$ and $\Delta S_{U}>0$",
        (
            surveyItem('28', ('A', 'C')),
        )
    ),
])

tables.append(entropyReversibleTable)
"""###########################################################"""

fns = "{\\footnotesize"
entropyIrTable=lTable("Prevalence of Difficulties with Entropy in Irreversible Processes", displayWide=False)
entropyIrTable.difficulties.extend([
    #"\\multicolumn{6}{| l |}{\\itshape Irreversible Isochoric Process (Item 29)}\\\\",
    Difficulty("Not all processes have $\\mathbf{\\Delta S_{Isolated} = 0}$ (correct)",
        (
            surveyItem(fns+'16(G)'+"}", 'D'),
        )
    ),
    Difficulty("$\Delta S_{Isolated} = 0$ for all processes",
        (
            surveyItem(fns+'16(G)'+"}", ('A', 'B', 'C', 'E')),
        )
    ),
    "\\hline",
    "\\hline",
    "{\\itshape Irreversible Isochoric Process}\\\\",
    Difficulty("$\mathbf{\Delta S_{Isolated} > 0}$, $\mathbf{\Delta S_{System}>0}$. (correct)",
        (
            surveyItem('29', 'A'),
        )
    ),
    Difficulty("$\Delta S_{Isolated} = 0$, $\Delta S_{System}>0$.",
        (
            surveyItem('29', ('B', 'E')),
        )
    ),
    Difficulty("$\Delta S_{Isolated} = 0$, [$\Delta S_{System} = 0$]",
        (
            surveyItem('29', ('D')),
```

```
    )
  ),
  "\\hline",
  "\\hline",
  Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0}$␣for␣a␣spontaneous
     ␣process␣(correct)",
    (
      surveyItem('12', ('C', 'D', 'E')),
      surveyItem('14', ('C', 'D', 'E')),
      surveyItem('24', ('C', 'D', 'E')),
      surveyItem('32', ('A', 'C')),
    )
  ),
  Difficulty("$[\Delta␣S_{Isolated}␣=␣0]$␣for␣a␣spontaneous␣
     process",
    (
      surveyItem('12', ('A', 'B')),
      surveyItem('14', ('A', 'B')),
      surveyItem('24', ('A', 'B')),
      surveyItem('32', ('B', 'D', 'E')),
    )
  ),
  "\\hline",

  #"\\multicolumn{6}{| l |}{\\itshape Free Expansion Process (Item
      14)}\\\\",
  "{\\itshape␣Free␣Expansion␣Process}\\\\",
  Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0}$␣(correct)",
    (
      surveyItem('14', ('C', 'D', 'E')),
    )
  ),
  Difficulty("$\Delta␣S_{Isolated}␣=␣0$",
    (
      surveyItem('14', ('A', 'B')),
    )
  ),

    # "\\hiderowcolors",
    # "\\hhline{*{7}-}",
    # "\\showrowcolors",
    "\\hline",
  #"\\multicolumn{6}{| l |}{\\itshape Spontaneous (Items 12, 24)
     }\\\\",
  "{\\itshape␣Spontaneous␣Heat␣Transfer}\\\\",
```

```
Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0}$␣and␣$\mathbf{\
   Delta␣S_{Cold}␣>␣0}$␣\itshape{(correct)}",
  (
    surveyItem('12', 'D'),
    surveyItem('24', 'D'),
  )
),
Difficulty("$[\Delta␣S_{Isolated}␣=␣0]$␣and␣$\mathbf{\Delta␣S_{
   Cold}␣>␣0}$",
  (
    surveyItem('12', ('A')),
    surveyItem('24', ('A')),
  )
),
Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0}$␣and␣$[\Delta␣S_{
   Cold}=0]$",
  (
    surveyItem('12', 'C'),
    surveyItem('24', 'C'),
  )
),
Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0}$␣and␣$\mathbf{\
   Delta␣S_{Cold}␣>␣0}$␣\linebreak[1]␣but␣also␣$\Delta␣S_{Hot}␣>
   ␣0$",
  (
    surveyItem('12', ('E')),#&B
    surveyItem('24', ('E')),
  )
),
# Difficulty("{\itshape Only} $\Delta S_{Hot} > 0$",
#   (
#     surveyItem('12', ('B')),#&B
#     surveyItem('24', ('B')),
#   )
# ),
"\\hline",
#"\\multicolumn{6}{| l |}",
"{\\itshape␣Spontaneous␣Mixing}\\\\",
Difficulty("$\mathbf{\Delta␣S_{Isolated}␣>␣0␣}$,␣and␣each␣$\
   mathbf{\Delta␣S_{subsys}␣>␣0}$",
  (
    surveyItem('32', 'C'),
  )
),
```

```
    Difficulty("$\Delta S_{Isolated} > 0 $, but each $[\Delta S_{
        subsys} = 0]$",
      (
        surveyItem('32', 'A'),
      )
    ),
    Difficulty("$[\Delta S_{Isolated} = 0]$, and each [$\Delta S_{
        subsys} = 0$]",
      (
        surveyItem('32', 'D'),
      )
    ),
    Difficulty("$[\Delta S_{Isolated} = 0]$, but each $\Delta S_{
        subsys}>0$",
      (
        surveyItem('32', 'B'),
      )
    ),
    "\\hline",
    Difficulty("Second law {\itshape (correct)}",
      (
        surveyItem('13', ('C', 'E')),
        surveyItem('25', ('C', 'E')),
        surveyItem('33', ('C', 'E')),
      )
    ),
    Difficulty("First law ",
      (
        surveyItem('13', ('B', 'D', 'E')),
        surveyItem('25', ('B', 'D', 'E')),
        surveyItem('33', ('B', 'D', 'E')),
      )
    ),
    Difficulty("Newton's $2^{nd}$ law",
      (
        surveyItem('13', ('A', 'D')),
        surveyItem('25', ('A', 'D')),
        surveyItem('33', ('A', 'D')),
      )
    ),
])
tables.append(entropyIrTable)

"""#########################################################"""
```

```python
for t in tables: t.setCorrect()
"""Clean up tables"""
uppers = catalog[catalog['upper']==1]

alg = catalog[catalog['level']=='Algebra']

algPre = alg[alg['preorpost']=='Pre']
#algPre = catalog[logical_and(catalog['level']=='Algebra', catalog
    ['preorpost']=='Pre')]
algPost = alg[alg['preorpost']=='Post']


calc = catalog[catalog['level']=='Calculus']
calcPre = calc[calc['preorpost']=='Pre']
calcPost = calc[calc['preorpost']=='Post']

#calcPost = catalog[ catalog['level']=='Calculus' && catalog['
    preorpost']=='Post']


allind = catalog[catalog['ind']==1]

c = catalog
slicepre  = c['preorpost']=="Pre"
slicepost = c['preorpost']=="Post"
slicecalc = c['level']=="Calculus"
slicealg = c['level']=="Algebra"
sliceintro = slicecalc + slicealg
slicePitt = c['institution']=='Pitt'
sampleIntroPre = Sample("Intro Pre", c[sliceintro*slicepre])
sampleIntroPost = Sample("Intro Post", c[sliceintro*slicepost])

samples = [
  Sample("Graduate Pairs", c[c['ind']==2]),
  Sample("Graduates", c[c['level']=='TA']),
  Sample("Upper Post", c[c['upper']==1]),
  Sample("Calc Pre", c[slicecalc * slicepre]),
  Sample("Calc Post", c[slicecalc * slicepost]),
  Sample("Algebra Pre", c[slicealg * slicepre]),
  Sample("Algebra Post", c[slicealg * slicepost]),
  ]
del c

wat = Sample("All Students", catalog)
```

```
topics = [
  Topic("First Law", [str(i) for i in
     [2,3,8,9,10,11,13,14,15,21,23,27,29,30,31,33]]),
  Topic("Second Law", [str(i) for i in
     [11,12,13,14,16,19,24,25,28,29,32,33]]),
  Topic("Entropy", [str(i) for i in
     [1,4,7,12,14,15,16,19,22,24,28,29,32]]),
  Topic("Work", [str(i) for i in
     [2,3,6,8,9,10,14,15,21,22,23,26,27,29,30,31]]),
]
topicsall = [
Topic("All", itemsArr[array([True,True,True,True,True,True,True,
   True,True,True,True,True,True,True,True,True,True,True,True,
   True,True,True,True,True,True,True,True,True,True,True,True,
   True,True])]),
Topic("Reversible", itemsArr[array([True,True,True,True,False,
   False,False,False,False,False,False,False,False,False,True,True
   ,False,False,True,False,False,False,False,False,False,False,
   False,True,False,False,False,False,False])]),
Topic("Irreversible", itemsArr[array([False,False,False,False,
   False,False,False,False,False,False,True,True,True,True,False,
   True,False,False,False,False,False,False,True,True,True,False,
   False,False,True,False,False,True,True])]),
Topic("Cyclic", itemsArr[array([False,False,False,False,True,True,
   True,True,False,False,False,False,False,False,True,True,False,
   True,True,False,False,False,False,False,False,False,False,False
   ,False,False,False,False,False])]),
Topic("Isothermal", itemsArr[array([False,False,True,True,False,
   False,False,False,True,False,False,False,False,False,False,
   False,False,True,False,True,False,False,False,False,False,True,
   True,True,False,False,False,False,False])]),
Topic("Isobaric", itemsArr[array([False,False,False,False,False,
   False,False,False,True,False,False,False,False,False,False,
   False,True,False,False,True,True,False,False,False,False,False,
   False,False,False,False,False,False,False])]),
Topic("Isochoric", itemsArr[array([False,False,False,False,False,
   False,False,False,False,False,False,False,False,False,False,
   False,True,False,False,False,True,False,False,False,False,False
   ,False,False,True,True,False,False,False])]),
Topic("Adiabatic", itemsArr[array([True,True,False,False,False,
   False,False,False,False,False,False,False,False,False,False,
   False,False,True,False,True,False,False,False,False,False,True,
   False,False,False,True,False,False,False])]),
Topic("Systems Universe", itemsArr[array([True,False,False,False,
   False,False,False,False,False,False,True,True,True,True,False,
```

```
    False, False, False, True, False, False, False, True, True, True, False,
    False, True, False, False, True, True, False])]),
Topic("Closed System", itemsArr[array([False, False, False, False,
    False, False, False, False, False, False, True, True, False, True, False,
    True, False, False, False, False, False, False, True, True, False, False,
    False, False, False, False, True, True, False])]),
Topic("Ideal Gas", itemsArr[array([False, False, True, True, False,
    False, False, False, False, False, False, False, False, False, False,
    False, True, False, False, False, False, False, False, False, False,
    False, True, True, False, True, True, True, False])]),
Topic("State Variables", itemsArr[array([False, False, False, False,
    True, True, True, True, False, True, False, False, False, False, True,
    False, False, False, False, False, False, True, False, False, False,
    False, False, False, False, False, False, False, False])]),
Topic("Internal Energy", itemsArr[array([False, True, True, False,
    True, False, False, False, True, False, True, False, False, True, True,
    False, True, False, False, False, True, True, True, False, False, False,
    True, True, True, True, True, False, False])]),
Topic("Relation to T", itemsArr[array([False, True, True, False, False
    , False, False, False, True, False, False, False, False, False, False,
    False, True, False, False, False, False, False, False, False, False,
    False, True, True, False, True, False, False, False])]),
Topic("Heat", itemsArr[array([False, False, False, True, False, False,
    False, True, True, True, False, False, True, False, True, False, False,
    True, True, False, True, False, False, False, False, False, True, False,
    False, False, False, False, False])]),
Topic("Work", itemsArr[array([False, True, True, False, False, True,
    False, True, True, True, False, False, False, True, True, False, False,
    False, False, False, True, True, True, False, False, True, True, False,
    True, True, True, False, False])]),
Topic("Relation to (pdV)", itemsArr[array([False, True, False, False,
    False, True, False, False, True, True, False, False, False, True, True,
    False, False, False, False, False, True, False, True, False, False, True,
    True, False, True, True, True, False, False])]),
Topic("Entropy", itemsArr[array([True, False, False, True, False, False
    , True, False, False, False, False, True, False, True, True, True, False,
    False, True, False, False, True, False, True, False, False, False, True,
    True, False, False, True, False])]),
Topic("Relation to S(dQT)", itemsArr[array([True, False, False, True,
    False, False, False, False, False, False, False, True, False, False,
    False, False, False, False, True, False, False, False, False, True, False
    , False, False, True, True, False, False, False, False])]),
Topic("PV Diagram", itemsArr[array([False, False, False, False, True,
    True, True, True, True, True, False, False, False, False, True, False,
    True, False, False, False, True, False, False, False, False, True, False,
```

```
    False, False, True, False, False, False])]),
Topic("1st Law", itemsArr[array([False, True, True, False, False, False
    , False, True, True, True, True, False, True, True, True, False, False,
    False, False, False, True, False, True, False, True, False, True, False,
    True, True, True, False, True])]),
Topic("2nd Law", itemsArr[array([False, False, False, False, False,
    False, False, False, False, False, True, True, True, True, False, True,
    False, False, True, False, False, False, False, True, True, False, False,
    True, True, False, False, True, True])]),
Topic("Engine Efficiency", itemsArr[array([False, False, False, False
    , False, False, False, False, False, False, False, False, False, False,
    False, True, False, False, True, False, False, False, False, False, False
    , False, False, False, False, False, False, False, False])]),
]
topicsall = [Topic(t.name, list(t.items)) for t in topicsall]
def topicsSummary(topics = topics, samples=samples):
  for t in topics:
    for s in samples:
      print Sample(" ".join([t.name, s.name]), s.cat, t.items, key
        =catalog[t.items][0])

def per_str(x):
  #return ("{0:.0f}\\%".format(100.0 * x))
  return ("{0:.0f}".format(100.0 * x))

def reloadcatfromcsv():
  with open('consolidated.csv', 'r') as f:
    catalog = recfromcsv(f)
  with open('consolidated5.npy', 'w') as f:
    save(f, catalog)
  return catalog

def topicscores(topics=topicsall, samples=samples):
  for s in samples:
    print s.name
    for t in topics:
      ts = t.Sample(s)
      print ", ".join([t.name, per_str(ts.score)])




if __name__ == '__main__':
```

```
if 'appendixtables' in sys.argv[1]:
  printTexDoc(tables, "Appendix")
elif 'tables' in sys.argv[1]:
  printTexDoc(tables, "Summary")
elif 'blocktable' in sys.argv[1]:
  blocktables = []
  for iNum in itemsList: #Each block
    row_template = [surveyItem(iNum, r) for r in
       PossibleResponses]
    correctresponse = [i for i in row_template if i.
       includesCorrect][0]
    incorrectresponses = [i for i in row_template if not i.
       includesCorrect]
    responses = [correctresponse]+incorrectresponses
    column1 = correctresponse.num
    for s in samples:
      row_responses = [correctresponse] + sorted(
         incorrectresponses, key= lambda i: i.responseRate(s),
          reverse=True)
      print latexdelim.join([column1] +
        [per_str(r.responseRate(s)) + " (" + r.response + ")"
           for r in row_responses] +
        [s.name]) + latexEOL + "*"
      column1=""
    print "\\cline{2-6}"
    print "\\multicolumn{7}{c}{}\\\\"
    print "\\cline{2-6}"

elif "diagnostic" in sys.argv[1]:
  for s in samples:
    print s
elif "TopicsDiagnostics" in sys.argv[1]:
  topicsSummary(topics, samples.extend([sampleIntroPre,
     sampleIntroPost]))
else:
  print "Ya done goofed."
```

BIBLIOGRAPHY

[1] John D Bransford and Daniel L Schwartz. Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, pages 61–100, 1999.

[2] George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[3] Michelene TH Chi, Paul J Feltovich, and Robert Glaser. Categorization and representation of physics problems by experts and novices*. *Cognitive science*, 5(2):121–152, 1981.

[4] Lin Ding, Ruth Chabay, Bruce Sherwood, and Robert Beichner. Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical review special Topics-Physics education research*, 2(1):010105, 2006.

[5] David P Maloney, Thomas L O'Kuma, Curtis J Hieggelke, and Alan Van Heuvelen. Surveying students' conceptual knowledge of electricity and magnetism. *American Journal of Physics*, 69(S1):S12–S23, 2001.

[6] Rachel E Pepper, Stephanie V Chasteen, Steven J Pollock, and Katherine K Perkins. Observations on student difficulties with mathematics in upper-division electricity and magnetism. *Physical Review Special Topics-Physics Education Research*, 8(1):010111, 2012.

[7] Chandralekha Singh. Student understanding of symmetry and gauss's law of electricity. *American journal of physics*, 74(10):923–936, 2006.

[8]

[9] Richard Steinberg, Michael C Wittmann, Lei Bao, and Edward F Redish. The influence of student understanding of classical physics when learning quantum mechanics. *Research on teaching and learning quantum mechanics*, pages 41–44, 1999.

[10] Guangtian Zhu and Chandralekha Singh. Surveying students' understanding of quantum mechanics in one spatial dimension. *American Journal of Physics*, 80(3):252–259, 2012.

[11] E. Marshman and C. Singh. Investigating student difficulties with time dependence of expectation values in quantum mechanics. *Proceedings of the 2013 Phys. Ed. Res. Conference*, pages 245–248, 2014.

[12] Chandralekha Singh. Student understanding of quantum mechanics. *American Journal of Physics*, 69(8):885–895, 2001.

[13] Dean A Zollman, N Sanjay Rebello, and Kirsten Hogg. Quantum mechanics for everyone: Hands-on activities integrated with technology. *American Journal of Physics*, 70(3):252–259, 2002.

[14] Michael C Wittmann, Richard N Steinberg, and Edward F Redish. Investigating student understanding of quantum physics: Spontaneous models of conductivity. *American Journal of Physics*, 70(3):218–226, 2002.

[15] Chandralekha Singh. Student understanding of quantum mechanics at the beginning of graduate instruction. *American Journal of Physics*, 76(3):277–287, 2008.

[16] Chandralekha Singh, Mario Belloni, and Wolfgang Christian. Improving students' understanding of quantum mechanics. *Physics Today*, 59(8):43, 2006.

[17] Chandralekha Singh. Interactive learning tutorials on quantum mechanics. *American Journal of Physics*, 76(4):400–405, 2008.

[18] Wolfgang Christian. http://www.opensourcephysics.org/.

[19] Wolfgang Christian, Francisco Esquembre, and Lyle Barbato. Open source physics. *Science*, 334(6059):1077–1078, 2011.

[20] Joan I Heller and Frederick Reif. Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction*, 1(2):177–216, 1984.

[21] Jill H Larkin and F Reif. Understanding and teaching problem-solving in physics. *European Journal of Science Education*, 1(2):191–203, 1979.

[22] Jill Larkin. Cognition of learning physics. *American Journal of Physics*, 49(6):534–541, 1981.

[23] D. P. Maloney. Research in problem solving: Physics. In Gabel, editor, *Handbook of Research on Software and Teaching*, pages 327–354. Macmillan, New York, 1994.

[24] Frederick Reif. Teaching problem solving–a scientific approach. *Physics Teacher*, 19(5):310–16, 1981.

[25] Frederick Reif. Millikan lecture 1994: Understanding and teaching important scientific thought processes. *American Journal of Physics*, 63(1):17–32, 1995.

[26] Frederick Reif. Scientific approaches to science education. *Physics Today*, 39(11):48–54, 2008.

[27] Alan Van Heuvelen. Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics*, 59(10):891–897, 1991.

[28] Robert Dufresne, Jose Mestre, Thomas Thaden-Koch, William Gerace, and William Leonard. Knowledge representation and coordination in the transfer process. *Transfer of learning from a modern multidisciplinary perspective*, pages 155–215, 2005.

[29] Pamela Thibodeau Hardiman, Robert Dufresne, and Jose P Mestre. The relation between problem categorization and problem solving among experts and novices. *Memory & Cognition*, 17(5):627–638, 1989.

[30] Jill Larkin, John McDermott, Dorothea P Simon, and Herbert A Simon. Expert and novice performance in solving physics problems. *Science*, 208(4450):1335–1342, 1980.

[31] Chandralekha Singh. When physical intuition fails. *American Journal of Physics*, 70(11):1103–1109, 2002.

[32] Andrew Elby. Helping physics students learn how to learn. *American Journal of Physics*, 69(S1):S54–S64, 2001.

[33] Bat-Sheva Eylon and F Reif. Effects of knowledge organization on task performance. *Cognition and Instruction*, 1(1):5–44, 1984.

[34] David E Meltzer. Relation between students' problem-solving performance and representational format. *American Journal of Physics*, 73(5):463–478, 2005.

[35] Darryl J Ozimek, Paula V Engelhardt, Andrew G Bennett, and N Sanjay Rebello. Retention and transfer from trigonometry to physics. In *AIP Conference Proceedings*, volume 790, page 173. IOP Institue of Physics Publishing Ltd, 2005.

[36] Alan Van Heuvelen. Overview, case study physics. *American Journal of Physics*, 59(10):898–907, 1991.

[37] Lei Bao and Edward F Redish. Understanding probabilistic interpretations of physical systems: A prerequisite to learning quantum physics. *American Journal of Physics*, 70(3):210–217, 2002.

[38] Mario Belloni, Wolfgang Christian, and Douglas Brown. Open source physics curricular material for quantum mechanics. *Computing in Science & Engineering*, 9(4):24–31, 2007.

[39] A Mason, E Cohen, E Yerushalmi, and C Singh. Identifying differences in diagnostic skills between physics students: Developing a rubric. In *2008 Physics Education Research Conference*, volume 1064, pages 147–150. AIP Publishing, 2008.

[40] E Cohen, A Mason, C Singh, and E Yerushalmi. Identifying differences in diagnostic skills between physics students: Students' self-diagnostic performance given alternative scaffolding. In *2008 Physics Education Research Conference*, volume 1064, pages 99–102. AIP Publishing, 2008.

[41] Chandralekha Singh, Edit Yerushalmi, and Bat Sheva Eylon. Physics learning in the context of scaffolded diagnostic tasks (ii): Preliminary results. In *2007 Physics Education Research Conference*, volume 951, pages 31–34. AIP Publishing, 2007.

[42] Edit Yerushalmi, Chandralekha Singh, and Bat Sheva Eylon. Physics learning in the context of scaffolded diagnostic tasks (i): The experimental setup. In *2007 Physics Education Research Conference*, volume 951, pages 27–30. AIP Publishing, 2007.

[43] Edit Yerushalmi, Andrew Mason, Elisheva Cohen, and Chandralekha Singh. Effect of self diagnosis on subsequent problem solving performance. In *2008 Physics Education Research Conference*, volume 1064, pages 53–56. AIP Publishing, 2008.

[44] Edit Yerushalmi, Elisheva Cohen, Andrew Mason, and Chandralekha Singh. What do students do when asked to diagnose their mistakes? does it help them? i. an atypical quiz context. *Physical Review Special Topics-Physics Education Research*, 8(2):020109, 2012.

[45] Edit Yerushalmi, Elisheva Cohen, Andrew Mason, and Chandralekha Singh. What do students do when asked to diagnose their mistakes? does it help them? ii. a more typical quiz context. *Physical Review Special Topics-Physics Education Research*, 8(2):020110, 2012.

[46] Andrew Mason and Chandralekha Singh. Do advanced physics students learn from their mistakes without explicit intervention? *American Journal of Physics*, 78(7):760–767, 2010.

[47] Eugenia Etkina, Anna Karelina, and Maria Ruibal-Villasenor. How long does it take? a study of student acquisition of scientific abilities. *Physical Review Special Topics-Physics Education Research*, 4(2):020108, 2008.

[48] Anna Karelina and Eugenia Etkina. Acting like a physicist: Student approach study to experimental design. *Physical Review Special Topics-Physics Education Research*, 3(2):020106, 2007.

[49] Carole Ames. Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, 84(3):261, 1992.

[50] Carol S Dweck and Ellen L Leggett. A social-cognitive approach to motivation and personality. *Psychological review*, 95(2):256, 1988.

[51] Paul R Pintrich. The role of motivation in promoting and sustaining self-regulated learning. *International journal of educational research*, 31(6):459–470, 1999.

[52] Deborah J Stipek. *Motivation to learn: Integrating theory and practice*. Allyn & Bacon, 2002.

[53] Michelene TH Chi, Miriam Bassok, Matthew W Lewis, Peter Reimann, and Robert Glaser. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, 13(2):145–182, 1989.

[54] President's Council of Advisors on Science Executive Office of the President and Technology. Report to the president engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. February 2012.

[55] Robert J Beichner. Testing student interpretation of kinematics graphs. *American journal of Physics*, 62(8):750–762, 1994.

[56] Chandralekha Singh and David Rosengrant. Multiple-choice test of energy and momentum concepts. *American Journal of Physics*, 71(6):607–617, 2003.

[57] Lorenzo G Rimoldini and Chandralekha Singh. Student understanding of rotational and rolling motion concepts. *Physical Review Special Topics-Physics Education Research*, 1(1):010102, 2005.

[58] Paula Vetter Engelhardt and Robert J Beichner. Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, 72(1):98–115, 2004.

[59] BS Bloom, MD Engelhart, EJ Furst, WH Hill, and DR Krathwohl. Taxonomy of educational objectives: The classification of education goals. handbook i: Cognitive domain (vol. 1). new york, ny: David mckay company, 1956.

[60] P Kline. A handbook of test construction: Introduction to psychometric design. london: Methuen. *DM Carter et aL/Schizophrenia Research 16 (1995) 157*, 165(165):581–586, 1986.

[61] G Frederic Kuder and Marion W Richardson. The theory of the estimation of test reliability. *Psychometrika*, 2(3):151–160, 1937.

[62] MW Richardson and GF Kuder. The calculation of test reliability coefficients based on the method of rational equivalence. *Journal of Educational Psychology*, 30(9):681, 1939.

[63] Benjamin W Dreyfus, Benjamin D Geller, David E Meltzer, and Vashti Sawtelle. Resource letter ttsm-1: Teaching thermodynamics and statistical mechanics in introductory physics, chemistry, and biology. *American Journal of Physics*, 83(1):5–21, 2015.

[64] Peter L Thomas and Richard W Schwenz. College physical chemistry students' conceptions of equilibrium and fundamental thermodynamics. *Journal of Research in Science Teaching*, 35(10):1151–1160, 1998.

[65] Matthew J Cochran and Paula RL Heron. Development and assessment of research-based tutorials on heat engines and the second law of thermodynamics. *American Journal of Physics*, 74(8):734–741, 2006.

[66] Elon Langbeheim, Samuel A Safran, Shelly Livne, and Edit Yerushalmi. Evolution in students' understanding of thermal physics with increasing complexity. *Physical Review Special Topics-Physics Education Research*, 9(2):020117, 2013.

[67] David E Meltzer. Investigation of students' reasoning regarding heat, work, and the first law of thermodynamics in an introductory calculus-based general physics course. *American Journal of Physics*, 72(11):1432–1446, 2004.

[68] Risto Leinonen, Mervi A Asikainen, and Pekka E Hirvonen. University students explaining adiabatic compression of an ideal gas—a new phenomenon in introductory thermal physics. *Research in Science Education*, 42(6):1165–1182, 2012.

[69] Michael E Loverude, Christian H Kautz, and Paula RL Heron. Student understanding of the first law of thermodynamics: Relating work to the adiabatic compression of an ideal gas. *American Journal of Physics*, 70(2):137–148, 2002.

[70] Risto Leinonen, Esa Räsänen, Mervi Asikainen, and Pekka E Hirvonen. Students' pre-knowledge as a guideline in the teaching of introductory thermal physics at university. *European Journal of Physics*, 30(3):593, 2009.

[71] Christian H Kautz and Gerhard Schmitz. Probing student understanding of basic concepts and principles in introductory engineering thermodynamics. In *ASME 2007 International Mechanical Engineering Congress and Exposition*, pages 473–480. American Society of Mechanical Engineers, 2007.

[72] David E Meltzer. Observations of general learning patterns in an upper-level thermal physics course. In *AIP Conf. Proc*, volume 1179, page 31, 2009.

[73] Linda C Hadfield and Carl E Wieman. Student interpretations of equations related to the first law of thermodynamics. *Journal of Chemical Education*, 87(7):750–755, 2010.

[74] Jack Barbera and Carl E Wieman. Effect of a dynamic learning tutorial on undergraduate students' understanding of heat and the first law of thermodynamics. *The Chemical Educator*, 14(2):45–48, 2009.

[75] H Goldring and J Osborne. Students' difficulties with energy and related concepts. *Physics Education*, 29(1):26, 1994.

[76] Kinsey Bain, Alena Moon, Michael R Mack, and Marcy H Towns. A review of research on the teaching and learning of thermodynamics at the university level. *Chemistry Education Research and Practice*, 2014.

[77] Mark F Granville. Student misconceptions in thermodynamics. *Journal of Chemical Education*, 62(10):847, 1985.

[78] Risto Leinonen, Mervi A Asikainen, and Pekka E Hirvonen. Overcoming students' misconceptions concerning thermal physics with the aid of hints and peer interaction during a lecture course. *Physical Review Special Topics-Physics Education Research*, 9(2):020112, 2013.

[79] Brandon R Bucy, John R Thompson, and Donald B Mountcastle. What is entropy? advanced undergraduate performance comparing ideal gas processes. In *2005 Physics Education Research Conference*, volume 818, pages 81–84, 2006.

[80] Christian H Kautz, Paula RL Heron, Peter S Shaffer, and Lillian C McDermott. Student understanding of the ideal gas law, part ii: A microscopic perspective. *American Journal of Physics*, 73(11):1064–1071, 2005.

[81] Christian H Kautz, Paula RL Heron, Michael E Loverude, and Lillian C McDermott. Student understanding of the ideal gas law, part i: A macroscopic perspective. *American Journal of Physics*, 73(11):1055–1063, 2005.

[82] Sylvie Rozier and Laurence Viennot. Students' reasonings in thermodynamics. *International Journal of Science Education*, 13(2):159–170, 1991.

[83] Evan B. Pollock. Student understanding of the physics and mathematics of process variables in pv diagrams. In *AIP Conference Proceedings*, volume 951, pages 168–171, 2007.

[84] Tor Nilsson and Hans Niedderer. An analytical tool to determine undergraduate students' use of volume and pressure when describing expansion work and technical work. *Chemistry Education Research and Practice*, 13(3):348–356, 2012.

[85] PH Van Roon, HF Van Sprang, and AH Verdonk. 'work'and 'heat': On a road towards thermodynamics. *International Journal of Science Education*, 16(2):131–144, 1994.

[86] David E Meltzer. Investigation of student learning in thermodynamics and implications for instruction in chemistry and engineering. In *2006 Physics Education Research Conference*, volume 883, pages 38–41, 2007.

[87] Warren M Christensen, David E Meltzer, and CA Ogilvie. Student ideas regarding entropy and the second law of thermodynamics in an introductory physics course. *American Journal of Physics*, 77(10):907–917, 2009.

[88] Judith M Bennett and Mustafa Sözbilir. A study of turkish chemistry undergraduates' understanding of entropy. *Journal of chemical education*, 84(7):1204, 2007.

[89] GTC Ribeiro. Entropy and the second principle of thermodynamics—fourth year undergraduates' ideas. *Research in Assessment*, 9:23–36, 1992.

[90] Trevor I Smith, Warren M Christensen, and John R Thompson. Addressing student difficulties with concepts related to entropy, heat engines and the carnot cycle. In *2009*

Physics Education Research Conference, volume 1179, pages 277–280. AIP Publishing, 2009.

[91] Jill H Larkin. *The role of problem representation in physics.* Carnegie-Mellon University, Department of Psychology Pittsburgh, PA, 1981.

[92] Jill H Larkin and Herbert A Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, 11(1):65–100, 1987.

[93] Yulin Qin and Herbert A Simon. Imagery and mental models in problem solving. In *Proc. AAAI Symposium on Reasoning with Diagrammatic Representations., Stanford, CA*, 1992.

[94] Jiaje Zhang and Donald A Norman. Representations in distributed cognitive tasks. *Cognitive science*, 18(1):87–122, 1994.

[95] Jiajie Zhang. The nature of external representations in problem solving. *Cognitive science*, 21(2):179–217, 1997.

[96] Lillian C McDermott. A view from physics. *M. Gardner, J. Greeno, F. Reif, AH Schoenfeld, A. diSessa, and E. Stage (Eds.), Toward a scientific practice of science education*, pages 3–30, 1991.

[97] Donald R Jones and David A Schkade. Choosing and translating between problem representations. *Organizational Behavior and Human Decision Processes*, 61(2):214–223, 1995.

[98] Robert J Dufresne, William J Gerace, and William J Leonard. Solving physics problems with multiple representations. *Physics Teacher*, 35:270–275, 1997.

[99] For example, see L. C. McDermott, *Millikan Lecture 1990: What we teach and what is learned-Closing the gap*, Am. J. Phys. **59**, 301, 1991; L. C. McDermott, *Guest Comment: How we teach and how students learn–A mismatch?*, Am. J. Phys. **61**, 295, 1993.E. F. Redish, *Implications of cognitive studies for teaching physics*, Am. J. Phys. **62**, 796, 1994; E. Redish, *Teaching Physics with the physics suite*, Wiley, 2003.

[100] National Reserach Council, *How people learn: Bridging research and practice*, Committee on learning research and educational practices, Eds. M. Donovan, J. Bransford, and J. Pellegrino, National Academy Press, Washington, DC, 1999; National Reserach Council, *How people learn: Brain, mind, experience and school*, Committee on developments in science on learning, Eds. J. Bransford, A. Brown and R. Cocking, National Academy Press, Washington, DC, 1999; H. A. Simon, and C. A. Kaplan, *Foundations of cognitive science*, (M. I. Posner editor, MIT press, 1989); J. R. Anderson, *The adaptive character of thought*, Hillsdale, NJ: Lawrence Erlbaum, 1990; J. Anderson and C. Lebiere, *Atomic components of thoughts*, Mahwah NJ: Lawrence Erlbaum, 1998; A. Newell, *Unified theories of cognition*, cambridge, MA: Harvard University Press, 1990; A. Newell and H. Simon, *Human Problem Solving*, Englewood Cliffs NJ: Prentice Hall,

1972; E. Thorndike, *Human Learning*, New York: Century, 1931; R. D. Tennyson and R. L. Elmore, *Learning theory foundations for instructional design* in Instructional Design: International Perspectives, (Vol. 1: Theory, Research, Models), R. D. Tennyson, F. Schott, N. Seel, and S. Dijkstra, Mhawah, NJ: Lawr. Erlb., 1997; R. G. Glaser, *Education and thinking: The role of knowledge*, Am. Psychologist, **39**, 93-104, (1984).

[101] F. Reif, *Millikan Lecture 1994: Understanding and teaching important scientific thought processes*, Am. J. Phys. **63**, 17 (1995); F. Reif, *Scientific approaches to science education*, Phys. Today **39**, 48, 1986; F. Reif, *Interpretation of scientific or mathematical concepts: Cognitive issues and instructional implications* **11**, 395, 1987; F. Reif and S. Allen, *Cognition of interpreting scientific concepts: A study of acceleration*, Cognition and Instruction **9(1)**, 1-44, (1992); J. Mestre and J. Touger, *Cognitive Research-what's in it for physics teachers?*, Phys. Teach. **27**, 447 (1989).

[102] H. Ginsberg, and S. Opper, *Piaget's theory of intellectual development*, Englewood cliffs, Prentice Hall, N.J., 1969; R. Gorman, *Discovering Piaget: A guide for teachers*, Merrill, Columbus, 1972; J. Piaget, *The language and thought of a child*, NY: Meridian, (1974); J. Piaget, *Success and understanding*, Cambridge, MA: Harvard University Press, 1978; A. Lawson, *A review of research on formal reasoning and science teaching*, J. Research in Science Teaching, **22(7)**, 569-617, 1985.

[103] L. Vygotsky, *Mind in Society: The development of higher psychological processes*, Harvard University Press, 1978; J. Wertsch, *Mind in Action*, NY: Oxford University Press, (1998); J. Valsiner, and R. van der Veer, *The encoding of distance: The concept of the zone of proximal development and its interpretations*. In (R. R. Cocking and K. A. Renninger Eds.) The development and meaning of psychological distance (pp 35-62), Hillsdale, NJ: Lawrence Erlbaum Associates; B. Litowitz, *Deconstruction in the zone of proximal development*, same book as above (pp. 184-196); C. A. Stone, *What is missing in the metaphor of scaffolding?*, In (E. Forman, N. Minick, and C. A. Stone Eds.) Contexts for learning: Sociocultural dynamics in children's development, (pp. 169-183), NY: Oxford University Press, 1993; J. Lave and E. Wegner, *Situated learning: Legitimate peripheral participation*, Cambridge University Press, 1991; J. Lave, *Cognition in practice*, Cambridge, England, Cambridge University Press, 1988.

[104] H. Simon, *How big is a chunk?*, Science, **183** (4124), 482-488, 1974; P. Kyllonen and R. Christal, *Reasoning ability is (little more than) working memory capacity?!*, Intelligence, **14**, 389-433, 1990; A. Baddeley, *Working Memory*, Oxford: Clarendon Press/Oxford University Press, 1986; M. Miyake, M. Just and P. Carpenter, *Working memory contraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts*, J. Memory and Language, **33(2)**, 175-202, 1994; A. Fry, S. Hale, *Processing speed, working memory and fluid intelligence: Evidence for a developmental cascade*, Psychological Science, **7(4)**, 237-241, 1996; D. Neves and J. Anderson, *Knowledge Compilation: Mechanisms for automatization of cognitive skills*, in Cognitive skills & their acquisition, Ed. J. Anderson, Hills., NJ: Lawr. Erl., 1981; R. Kail and T. Salthouse, *Processing speed as a mental capacity*, Acta Psychologica, **86**, 199-225, 1994.

[105] W. Chase and H. Simon, *The mind's eye in chess*, In Visual information processing, Ed. W. Chase, New York: Academic Press, 1973 W. Chase and H. Simon, *Perception in chess*, Cognitive Psychology, **4**, 55-81, 1973; F. Gobert and H. Simon, *Recall of random and distorted chess positions: Implication of the theory of expertise*, Memory and Cognition, **24(4)**, 493-503, 1996; P. Rosenbloom and A. Newell, *Learning by chunking: A production system model of practice*, in Production system models of learning and development, Eds. D. Klahr and P. Langley, Cambridge, MA:MIT press, pp. 221-286, 1987; D. Broadbent, *The magic number seven after fifteen years*, in Eds. A. Kennedy and A. Wilkes, Studies in long term memory, NY: Wilet, 1975.

[106] K. Ericsson and H. Simon, *Protocol Analysis: Verbal reports as data*, MIT Press, Revised Edition, 1993; M. T. H. Chi, in *Thinking Aloud* Chapter 1, Eds. Van Someren, Barnard, and Sandberg, 1, Erlbaum, 1994; M. T. H. Chi, *Quantifying qualitative analysis of verbal data: A practical guide*, J. Learning Sci. **6(3)**, 271, 1997.

[107] H. Simon, and C. Kaplan, *Foundations of cognitive science*, (M. I. Posner editor, MIT press, 1989), A. Newell and H. Simon, *Human Problem Solving*, Englewood Cliffs NJ: Prentice Hall, 1972; J. Anderson, J. Greeno, L. Reder, and H. Simon, *Perspective on learning, thinking, and activity*, Educational Researcher, **29(4)**, 11-13, 2000; H. Simon, *How big is a chunk?*, Science, **183** (4124), 482-488, 1974; W. Chase and H. Simon, *Perception in chess*, Cognitive Psychology, **4**, 55-81, 1973; F. Gobert and H. Simon, *Recall of random and distorted chess positions: Implication ofthe theory of expertise*, Memory and Cognition, **24(4)**, 493-503, 1996; Qin and H. Simon, *Imagery and mental models in problem solving*, In Pragmatic reasoning: Cognitive and conceptual perspectives, Eds. Glasgow, Chandrasekaran and Narayanan, Melno Park, CA, AAAI press/MIT press, pp. 403-434, 1995; H. Simon, *On the forms of mental representation*, In C. Savage (Ed.), Minnesota studies in the philosophy of science (Univ. Minnesota Press, Minneapolis), Vol. IX, *Perception and cognition: Issues in the foundations of psychology* (1978); H. Simon and D. Simon, *Individual differences in solving physics problems*, in Models of thought, Vol. II, Yale Univ. Press, (1989); H. Simon, *Reason in human affairs*, Basil Blackwell, (1983); H. Simon, *Rational choice and the structure of the environment*, Psych. Rev., **63**, 129-138, (1956).

[108] A. diSessa, *Knowledge in Pieces*, in Constructivism in the Computer Age, edited by G. Forman and P. Pufall (Lawrence Erlbaum, Hillsdale, NH, 1988), Chap. 4, pp. 49-70, A. diSessa, *What changes in conceptual change?* International Journal of Science Education, **20**, 1155 (1998).

[109] J. Sweller, *Cognitive load during problem solving: Effects on learning*, Cognitive science **12**, 257, 1988; J. Sweller, R. Mawer, and M. Ward, *Development of expertise in mathematical problem solving*, J. Exptal. Psychology: General **112**, 639, 1983.

[110] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," Phys. Teach. 30, 141-151 (1992); R. K. Thornton and D. R. Sokoloff, "Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active

learning laboratory and lecture curricula," Am. J. Phys. 66, 338-352 (1998); R. J. Beichner, "Testing student interpretation of kinematics graphs," ibid. 62, 750-762 (1994); R. Thornton, "Questions on heat and temperature", (1992); D. R. Sokoloff, "Electric circuit concept test", (1993); D. P. Maloney, T. L. O'Kuma, C. J. Hieggelke, and A. V. Heuvelen, "Surveying students' conceptual knowledge of electricity and magnetism," Am. J. Phys. 69, S12-S23 (2001).