Student Reasoning and Collaboration Networks in Thermal Physics

by

Nathan Philip Crossette

B.S., Muhlenberg College, 2013

M.A.St., University Cambridge, 2014

M.S., University of Colorado Boulder, 2017

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Physics

2022

Committee Members:

Bethany R. Wilcox, Chair

Heather Lewandowski

Steven Pollock

Paul Beale

Christine Lv

Crossette, Nathan Philip (Ph.D., Physics)

Student Reasoning and Collaboration Networks in Thermal Physics

Thesis directed by Assistant Professor Bethany R. Wilcox

Thermal physics courses have received relatively less attention from the field of physics education research than other core physics courses like quantum mechanics or electromagnetism. This thesis is composed of two projects which look at thermal physics courses from different perspectives. The first looks qualitatively at student reasoning in think-aloud interviews on a set of conceptual problems related to entropy. We use a conceptual resources framework to analyze and compare graduate and undergraduate student responses. The set of questions includes both new and previously studied problems and includes a novel system—a string waving in a bath of water—which could be used as a complementary way of introducing students to the concept of entropy. The second project quantitatively examines social networks of students working together on homework assignments in the upper-division thermal physics course at CU Boulder (as well as a middle-division math methods course at the Colorado School of Mines). We calculate the correlation between nodal centrality measures, which quantify how connected a node is to its larger network, and performance to quantify the relationship between collaboration and course grades. Also, we studied the possible effects of systematic errors caused by missing data within networks to better understand the significance of our results.

Acknowledgements

My journey towards this degree was long, with a few extra (interesting) twists and turns along the way, and without the help, support, advice, and mentorship from many people, I would not have completed it. To begin, I would like to thank the PER group at CU Boulder for accepting me back to the PhD program. In particular, I would like to thank Bethany Wilcox for her mentorship and support through my thesis research, and Mike Vignal who was an invaluable source of advice and wisdom on qualitative analysis of think-aloud interviews and on crafting tables in LaTeX. I would also like to thank Mina Fasihi Dolatabadi for helpful conversations about network analysis. An also a thank you to Steve Pollock and Noah Finkelstein for checking in from time to time and for their support and guidance as I prepared to restart my PhD.

In my time away from my graduate work, I worked two jobs simultaneously that, together, gave me the confidence that I could handle anything I put my mind to. At Tech-X Corporation, I would like to thank Jonny Smith, Diana Cheatham, Tom Jenkins, and Jarrod Leddy for their professional mentorship and camaraderie. At Eldora, I would like to thank the Ski Patrol, particularly Travis Brock, Brian Beach, and Jim Barnes, for one of the most meaningful experience of my life. For providing me a start in the ski industry and developing my skiing ability, I would also like to thank the Snowsports team at Eldora.

My first few years of graduate school were definitely the most difficult. I cannot think back at my time at Cambridge without remembering the kindness, support, and encouragement of three friends: Elias Most, Verena Neufeld, and Marise Westbroek. I could not express it at the time, but I am deeply grateful for your friendship as I struggled through a very challenging year. I would also like to thank my research advisers at Cambridge, Silvia Vignolini and Ahu Gumrah-Parry, for their sincere investment in my career and for helping me to see my competence as a scientist.

After my year at Cambridge, I had my first stint in the graduate program at CU Boulder. I would like to thank Loren Hough, my first research adviser at CU Boulder, for his support of my teaching ambitions. During this time Keith Tauscher also helped me through nearly every assigned physics homework set, and has since been a great friend and a just-as-or-even-better resource than StackOverflow for technical programming questions. Before Cambridge, I first began pursuing my career in physics at Muhlenberg College and would like to thank Penny and Bill Dunham, Brett Fadem, and Adam Clark, my advisers and professors who have been incredible sources of guidance throughout my career.

I would also like to to thank all my teachers. Especially my primary school teachers at Glenside Weldon and Copper Beech: Mrs. Snyder, Mrs. Loeffel, Mrs. Burke, Mrs. Tippin, Mrs. Englehardt, Mrs. Miller, and Mrs. McNew. And my junior & senior high school math science teachers: Mrs. Bocian (physical science), Mr. Goldschmidt (earth science), and Mr. Hartwell (biology), Mr. Fiorino (algebra and calculus) Mr. Williams (chemistry), Mr. Wiley (physics), Mr. Palmer (AP chemistry), Dr. Marianne Deakyne (AP Physics), Mr. Tinney (AP Calculus), and Mr. Cragle (AP Biology).

Contents

Tables

Table

Figures

Figure

Chapter 1

Introduction

In the canonical undergraduate physics curriculum, students take a thermal physics course which combines the topics of thermodynamics and statistical mechanics. Thermodynamics typically looks top-down on large, macroscopic systems and includes topics like work, energy, temperature, entropy, the laws of thermodynamics, heat engines, free energy, and chemical processes. Statistical mechanics, on the other hand, takes a complementary, bottom-up approach to calculating thermodynamic quantities starting from microscopic models [1]. Boltzmann factors, partition functions, statistical ensembles, and definitions of thermodynamic quantities as logarithmic derivatives of the partition function form the basis of statistical mechanics.

According to a wide survey of instructors and institutions, most thermal physics courses are taught at the upper-division level to junior and senior physics majors [2]. Somewhat surprisingly given the wide array of concepts and material covered by thermodynamics and statistical mechanics, there was more consensus and uniformity in the content covered by instructors of thermal physics courses than expected. Overall, there have been relatively few studies of this upper-division course within the field of physics education research (PER) both when compared to studies of thermal physics at the lower-division and to other upper-division courses like electromagnetism and quantum mechanics.

Despite the limited research focusing on thermal physics, there are several prior studies directly relevant to the work in this thesis. There is a more detailed literature review in Chapter 2, but the most closely related prior work includes a study investigating student reasoning about entropy and thermal equilibrium which identified several ways students commonly conceptualized entropy [3]. Another examined students' understanding of entropy in the context processes involving ideal gases [4]. We borrowed the system used in this study to include in our investigations of student reasoning. Additionally, research has looked at student difficulties with entropy, heat engines, and the Carnot cycle [5]. These articles mostly focused on the macroscopic perspective in thermal physics. Investigations of the microscopic perspective in thermal physics have examined the consistency of student reasoning on the second law of thermodynamics [6] and introductory-level students ideas of the ideal gas law [7].

Among other upper-division physics courses (like electromagnetism and quantum mechanics), thermal physics has at least as many connections to real-world situations. Thermal physics is not only directly applicable to weather and climate but also to developing technology to address anthropogenic climate change. Everyday appliances like refrigerators, heaters, and air conditioners are part of the thermal physics curriculum. Gas-powered automobiles still use the Otto cycle. Statistical mechanics directly ties into fewer 'everyday' kinds of situations, but has applications to condensed matter, complex system, and information theory all of which are relevant to the very active enterprise of developing quantum computing; though, population inversion, which generates laser light, is a familiar application of statistical mechanics. These connections to familiar systems and its direct applications to condensed matter make thermal physics worthy of deeper research in PER. This thesis is composed of two separate projects that center around thermal physics courses.

The first project qualitatively examines student reasoning on conceptual problems focusing on a notoriously inscrutable quantity: entropy. We posed a set of questions, all involving entropy, to both graduate and undergraduate students in think-aloud interviews in which students solved the problems outloud in the presence of an interviewer. We welcome and encourage instructors to use the flagship question of the interview, which centered around a system of a string waving in a bath of water, as an instructional tool that provides a new way to introduce students to entropy and how it is related to microstates and macrostates. Interviewing both graduate and undergraduate students not only allows for a comparison between the two groups' understanding and conceptualizations of entropy, but also adds to the very limited PER research examining graduate-level student learning. This project is discussed in Chapters 2 and 3.

Chapter 2 focuses exclusively on the interviews with graduate students, and begins by describing the development and explanation of the interview questions, then describes the research methodology before discussing the graduate student reasoning on, and responses to, the interview questions. In Chapter 3, we begin by discussing the results of the undergraduate interviews in the same format as the graduate interviews, and some comparisons to the graduate student responses are made to add some perspective on student reasoning from across these two populations. Then, we take a step back to examine patterns in student reasoning across questions and across the two populations where we draw more comparisons between the reasoning used by the two populations of students. Finally, we discuss some implications we see that may be of interest to instructors of thermal physics courses.

The second project quantitatively explores the correlation between student collaboration on homework assignments and their performance (as measured by grades on homework assignments and exams) in two thermal physics courses and a math methods course using social network analysis (SNA). Having a sense of belonging within a community is associated with students' persistence and achievement [8, 9], so studying a medium through which students interact and develop social interactions, which could help to generate a sense of belonging, may shed light on ways to create more effective teaching environments. This study focused on student collaboration in three courses: two thermal physics courses at the University of Colorado Boulder (CU Boulder) and a math methods course at the Colorado School of Mines (Mines). Furthermore, this research follows up on a study that also examined how collaboration was related to course performance. The prior study examined in-person courses at the Colorado School of Mines that took place before the COVID-19 pandemic. It found significant correlations between students' connectivity to their classmates and their course performance [10]. Our study extends this social networks research to courses at CU Boulder and courses affected by the COVID-19 pandemic. The results of this study are presented in Chapter 4. Also in this chapter is an overview of the relevant network analytic concepts intended to be accessible to readers with no prior experience with network analysis. We also describe details of the data collection and analysis processes.

Chapter 5 explores the effect missing data could have on the correlations we calculate in Chapter 4. Networks are complex, non-linear objects meaning a small perturbation (in the form of missing or extra data) could have anywhere from a negligible to profound effect on the overall network depending on the location of the perturbation. Since the networks examined in this study were constructed from students' self-reports of collaboration it is expected that there will be both some statistical and systematic error. In particular, there was an asymmetry between students reports of receiving help and giving help (students tended to report more instances of receiving help than giving help) indicating that there is missing or extra data. Furthermore data is missing from students who did not give their consent to the data collection process.

To understand how the missing data could affect our results, methods of simulating student collaboration networks were developed. Coincidentally, this process involved randomly generating ensembles of networks that matched macroscopic properties of the real networks observed in physics courses: a technique that borrows the idea of a statistical ensemble from statistical mechanics. Nodes were dropped from the simulated networks in the ensemble to determine the effect on the correlations between centrality and performance. Because the simulated networks, on average, had zero correlation between performance and centrality, this analysis provides insight on how likely the systematic effect of missing data was to inflate correlations.

Alternatively, to study the effect of missing data on networks which initially have large correlations between centrality and performance, a process for dropping nodes from the collaboration network observed in the math methods course at Mines was developed. This network was chosen for this analysis since it was the most complete (in terms of student participation) and also had large correlations between centrality and performance. A total of five methods for dropping students were implemented, each of which took a different assumption about which students were more likely to be missing from the network.

The popularity of network analysis as a tool to study student learning and outcomes is growing within the field of PER. Network-based clustering has been used to study the structure of student responses to multiple choice physics concept inventories to identify common clusters of incorrect responses chosen by students. Also, as is explored in this thesis, *social* network analysis, which applies networks to study interactions among a group of people, can be used to study the effects of students' social connections on their outcomes (grades, persistence, etc.) in physics. It has been noted in the field of network analysis that applied network analysis research rarely analyzes the possible impacts that measurement errors have on calculated centrality measures. With the work discussed in Chapter 5, we hope to increase the awareness of the possible effects of measurement errors and to encourage the exercise within the field of PER.

Much of the work covered in this thesis has been or is in the process of being published. All the results from the interviews with graduate interviews was published in the Physical Review Physics Education Research (PRPER) journal (see [11]), except for the results from one section (Sec 2.6.1) of the interviews which was published in the Physics Education Research Conference (PERC) proceedings (see [12]). Results from the same section of the interviews with undergraduate physics students (covered in Sec. 3.4.1) was presented the 2021 PERC conference proceedings (see [13]). The rest of the results from the undergraduates presented in Chapter 3 are new findings. The results from Chapters 4 and 5 are currently under review for publication in PRPER, but these chapters include an expanded discussion of some network analysis and statistical topics than is covered the submitted manuscript.

Chapter 2

Graduate Student Reasoning About Entropy

2.1 Introduction

Energy and entropy, two core thermal physics concepts, have a wide range of relevancy across the fields of biology, chemistry, physics, and engineering. While energy lacks a general, situational-invariant mathematical definition, it causes less conceptual discomfort among students than entropy. Somewhat paradoxically, entropy has a generalized mathematical definition:

$$S = -k_B \sum_i p_i \ln p_i \tag{2.1}$$

(where p_i is the probability of state i), yet remains a subtle and difficult to understand concept. In many cases, a generalized mathematical definition lends a tool to the understanding of a concept but the concise formula for entropy in Eq. (2.1) belies deep conceptual ideas about microstates, macrostates, and volumes in phase space. Furthermore, despite its multitude of mathematical forms, energy tends to have a more conceptually clear role in how systems behave, especially once the context is specified and the appropriate mathematical form for energy is chosen.

Entropy and free energy optimization play a critical role in explaining nearly all molecular biological processes from membrane formation to protein folding to metabolism. The foundational nature of entropy as a physical quantity maximized at equilibrium makes it a powerful and important tool for predicting and understanding the long-term behavior of complicated systems. In 2015, Dreyfus *et al.* [14] thoroughly catalogued and summarized prior research on student learning in thermodynamics and statistical mechanics across the disciplines of physics, chemistry, and biology. Current Physics Education Research (PER) literature on upper-division thermal physics is limited, and work specifically related to entropy has mostly explored its thermodynamic and macroscopic contexts such as heat engines and the Carnot cycle [14]. One exception is a study by Leinonen *et al.* which compared the consistency of upper-division undergraduate student reasoning about entropy changes of two objects in thermal contact from both microscopic and macroscopic perspectives. The study found a majority of students applied the second law of thermodynamics consistently across the problems from both perspectives, though there was a notable minority of students who treated entropy (or, more often, the number of accessible microstates W) as a conserved quantity [6]. This study expanded on research by Christensen *et al.*, which saw a large number of introductory physics students also treating entropy as a conserved quantity [15].

Other research on undergraduate understanding of entropy from the thermodynamic perspective has examined student reasoning in various contexts such as ideal gases [4, 7, 16] and heat engines [5]. A common theme emerging from these studies points to a tendency of students to 'over apply' the second law of thermodynamics to conclude that entropy cannot decrease even locally, which is possibly related to a finding that students sometimes struggle to disentangle systems, surroundings, and the universe [4, 15]. Additionally, a tendency for physics students to neglect to utilize state function properties when reasoning about ideal gasses has been identified [4, 5].

In a study on heat engines and cyclic processes, Smith *et al.* found that students did not articulate the connection between the constraint imposed by the second law and the Carnot cycle, and demonstrated some difficulty in distinguishing between differential and net changes of state properties [5]. Despite a focus on the macroscopic perspective of entropy, some studies have shed light on student's ideas of the microscopic nature of entropy. In two separate studies, Loverude found that students may struggle to fully differentiate microstates from macrostates, and have not fully connected the idea of entropy with multiplicity [3, 17].

The study presented in this chapter centers around interviews of eight physics graduate students at the University of Colorado Boulder and was intended to provide a qualitative picture of students' understanding of entropy from both macroscopic and microscopic perspectives. Though, due to selection effects, this work can only give perspective on a limited segment of students. The interview consisted of four physics content questions and one short follow-up discussion question. Two questions, which both addressed the macroscopic perspective of entropy, had been used in prior research to study undergraduates' understanding of entropy [4, 3]. Using these questions in interviews with graduate students allows for a more direct comparison and a wider 'triangulation' between undergraduate and graduate student reasoning to uncover a deeper perspective on students' ideas of entropy. The two other, new questions address entropy from a more microscopic perspective which has received less attention by prior research. The findings discussed in Sec. 2.6.1, which presented students with a system of a neutrally buoyant string waving in a water bath, have been reported previously [12]. This question also directly addressed entropy as it relates to probability, which Dreyfus reports has received little coverage in previous literature [14].

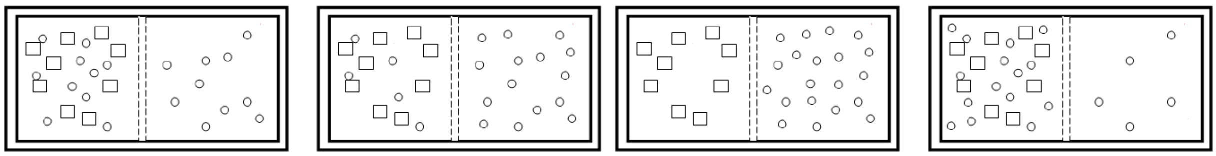
Graduate physics students remain an understudied population within the PER field. Learning more about their conceptual difficulties will indicate truly persistent student struggles, which also provides insight into undergraduate difficulties, and expands the understanding of how graduate students reason and construct models. In addition to potentially improving graduate student learning, future work in this vein will provide more perspective on difficulties experienced by undergraduates. Generally, our research goals were to better understand how graduate students reason with entropy, how graduate student reasoning compares with that of undergraduates, and how graduate students apply their understanding of entropy to an unfamiliar situation. This chapter will focus exclusively on the graduate students. In the following chapter (Chapter 3), we will discuss the undergraduates' responses and compare them to the responses of the graduate students. The entirety of study and results presented in this chapter first appeared in published articles [12, 11].

2.2 Development and Description of Questionnaire

The interview questionnaire contained four physics-content questions and a fifth short question directly asking students about how they conceptualize entropy. See Appendix A for a version of the actual document seen by students in the interviews.

The first question in the interview, the "Partitioned Box" question, presented students four figures representing different states of a system composed of a box with two sections separated by a semi-permeable membrane and two species of particles: a circle species and a square species. The problem statement told students the circles could freely cross the membrane but the squares could not, and to treat the two species

Figure 2.1: The Partitioned Box question



(a) 10 circles: 10 circles (b) 5 circles: 15 circles (c) 0 circles: 20 circles (d) 15 circles: 5 circles

A) Rate the states based on their entropy. Which has the highest?

B) Which of these pictures most closely represents the equilibrium state of the system? How will each state evolve with time?

C) For each of the four states above, which side (left or right) is at a higher pressure?

D) Is your answer to the previous question consistent to any claims made in part B about equilibrium? Is anything maximized, minimized or equilibrated at equilibrium? Please elaborate.

as ideal gasses. These figures and the questions students were asked are shown in Fig. 2.1.

Though it was not explicitly stated in the question prompt, temperature could be assumed to be uniform throughout the boxes (if students asked about temperature, they were told to assume it was uniform). Since the temperature and volume of the squares are identical in each of the four states, the entropy of the squares must also be identical in each of the four states. The addition or removal of a circle has no effect on the entropy of the squares since both species are ideal gasses, making the relative entropies of the 4 systems depend completely on the distribution of circles. So while the state depicted in Fig. 2.1b is a tempting choice for the state with the most entropy —since the pressures on the two sides of the membrane are the most equlibrated — the state in Fig. 2.1a actually has the most entropy. This system is analogous with the osmosis of water across semi-permeable membranes separating two different solutions. Both are examples of a chemical equilibrium where particles are exchanged to equilibrate the chemical potential of each species across the membrane. For mechanical equilibrium to be established, where pressure equalizes, the systems must be able to exchange volume.

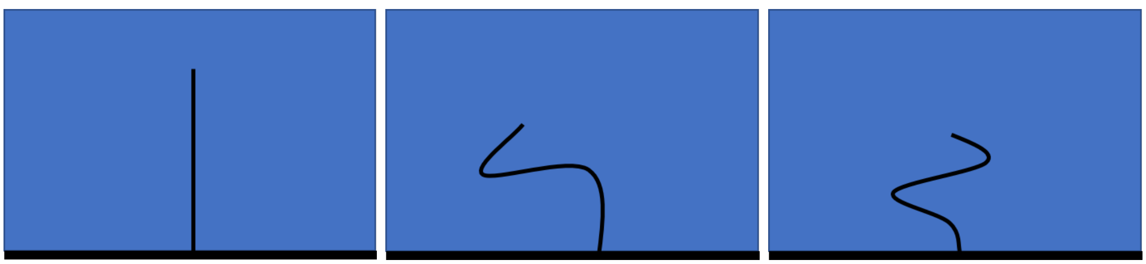
By asking first about entropy and equilibrium, we wanted to see the first, ‘instinctual’ method with which students rank entropy. Then, by asking about pressure, we wanted to see how they would contend with thinking about a property (pressure) that could mislead a student by suggesting a consistent, though incorrect, set of answers to this first question: that the state with the most equal pressures between the left and right side of the box has maximum entropy and therefore the equilibrium state, and that the pressure difference is minimized at equilibrium.

The second interview question, the "Blocks" question (and did not have a figure), was meant to be more familiar to students, and is similar to one used by Christensen *et al.* [15] and Leinonen *et al.* [6], but was developed independently. It presented students with an isolated system of two solids (Block A and Block B). Initially block A was at a higher temperature than block B, then the two blocks are placed into thermal contact. We then asked what happened to the blocks, and what could be said about the entropies of the individual blocks before and a long time in the future.

This question was intended to provide interviewees with a bit of a reprieve in between the more difficult first and third questions, but also to probe to what extent students consider entropy to be a 'substance' that can be transferred from one object to another. Because nothing was specified about the volume, mass, and heat capacities about the two blocks, nothing can be said about the entropies of the blocks before and after except that Block A's entropy decreased, and Block B's entropy increased by an amount more than the decrease in Block A.

The third question, the "Strings" question, centered around a simplified biophysical system, likely unfamiliar to most physics students: a neutrally buoyant string waving in a water bath. The string was attached at one end to a wall of the water bath. The first part of the question started with three snapshots of different string conformations (or arrangements) and asked students about probabilities, microstates, macrostates, and entropy (see Fig. 2.2). The intention with this question was that each of the three snapshots were microstates, and so the probabilities of finding the system in each of the states are all equal. This is a simplification of the system, since the most technical analysis would have to include the 'momentum density' (as a function of length along the string) which describes how each part of the string is moving at a particular time. Furthermore, it is also appropriate to consider finer structure of the string, like the particular arrangements of the atoms/electrons within the string, as the microstates of the system. However, for the sake of creating a more tractable problem (and also a more understandable and illuminating practical demonstration of entropy), we attempted to pitch the question such that students would not be cued to consider these concerns. In terms of defining macrostates, we did not consider there to be only one

Figure 2.2: First Part of the Strings Question.



(a) (b) (c)

A) Based on your intuition, rank the probabilities of finding the string in each of the three conformations shown above.

B) Is there a property of the string that can be used to define a set of distinct macrostates of the string? Are the conformations shown in the figure macrostates or microstates?

C) Based on your answer to part B, how would you rank the probabilities of finding the strand in each of the three conformations above?

D) Can you discuss what is meant by the “entropy” of the string, and how it relates to the possible conformations of the string?

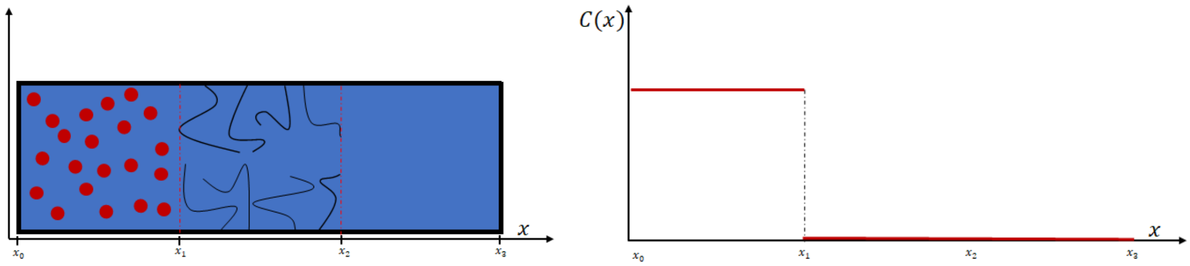
correct answer. As long as the classification property provided by the student generally outlined a means of partitioning microstates, we considered their response to be correct. On this task, we were more interested in the range and creativity of the definitions generated by students.

The final part of the Strings question, Part E, took a deeper step into the string system and was not discussed in our prior article. We presented the students with a new, more complicated system of a channel with three sections and with strings attached to the walls in the middle section (Fig. 2.3a). The first and third sections were free of strings, but the first section was filled with red circular particles (balls) suspended in the water. We asked students what happened to the entropy of the strings when a ball entered the region with the strings, then to draw a plot of the concentration profile of the balls throughout the channel after a long period of time had passed.

This question was structured to guide students' reasoning surrounding the plot of the concentration profile by first having them think about the 'entropic cost' of having balls in the region with the strings. A ball entering the strings region between x_1 and x_2 in Fig. 2.3 decreases the available conformations of the strings, and therefore reduces the entropy of the strings. However, having the balls diffuse through the system increases the entropy of the balls, so there will be a trade-off between the increase in entropy of the balls and a decrease in the entropy of the strings, resulting in a symmetrical concentration profile with a dip in the center region.

The final physics-content question, the "Expansions" question, was taken nearly in its entirety from a study done by Bucy *et al.* [4]. In this question, students were shown diagrams of two ideal processes: an isothermal expansion and a free expansion into vacuum. Both gases began at the same volume, pressure, and temperature and expand to the same final volume (see Fig. 2.4). The first gas, undergoing the isothermal expansion, is in contact with a reservoir (necessary to keep the temperature constant) and the second gas is completely isolated from the outer environment. Students were asked about the signs of the changes in entropy of the two gases, to compare the magnitudes of the changes, and to compare the changes in entropy of the outer environments of the two gases. This final task was a small extension of the question from Bucy's study which was added due to the findings of Christensen [15, 4] which found that undergraduates struggle to disentangle the difference between a system, its surroundings, and the universe.

Figure 2.3: Second Part of the Strings Question



(a) (b)

i) What happens to the number of possible conformations of the strings when the molecules enter the region with the strings?

ii) What will the concentration profile, as a function of the distance along the channel, of the molecules be after a long period of time? Explain your reasoning.

Figure 2.4: Expansions Question Diagram

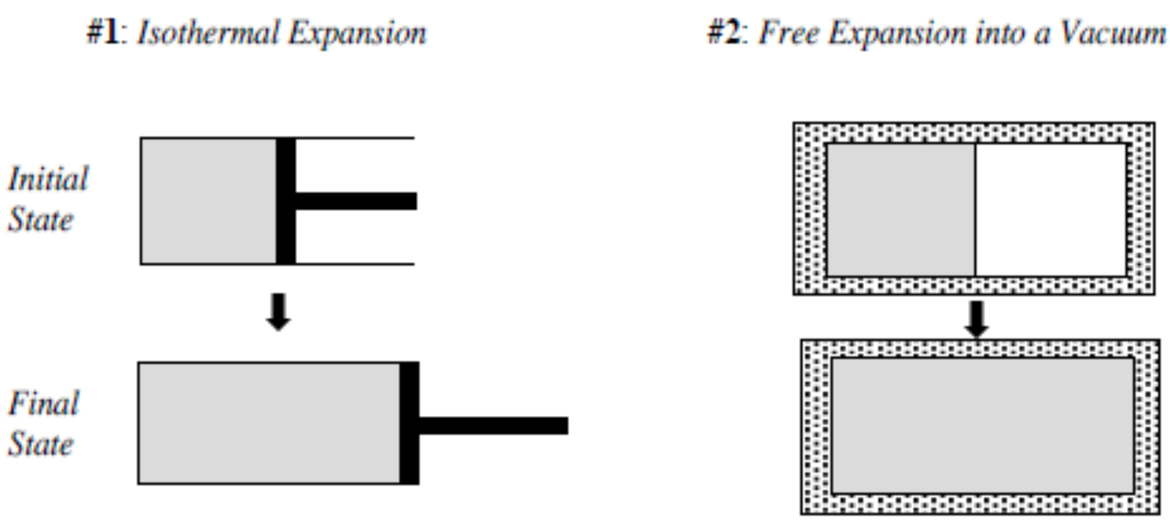


Diagram provided to students in the prompt of the Expansions question. This figure was taken from Bucy [4].

As it turns out, the two gases both increase in entropy by the same amount. The first gas's temperature is constant, so its internal energy, $U = \frac{3}{2}Nk_BT$, remains constant. The second gas does no work, and exchanges no heat with its environment, so its internal energy remains constant as well. Since the two gases also end at the same volume, it must be true that the two gases are in the same final state because their volumes and temperatures, two of their state variables, are identical. The two processes differ in that the isothermal expansion is a reversible process, so the increase in entropy of the gas is canceled by the decrease in entropy of the environment. The second process is irreversible so the total change in entropy of the gas and the outer environment is positive. After completing four the questions described above, students were asked to discuss their conception of entropy as a short follow-up to the interview.

Before interviews were conducted, the strings question was vetted by a physics faculty member familiar with block co-polymer models: the physical system on which this question was inspired. Additionally all questions were vetted with thermodynamics instructors from the chemistry, engineering, biology, and chemical engineering departments. Generally, the instructors from across disciplines found the questions appropriate and relevant to their courses and students. However, one instructor (from Biology) took a mild exception to the question about the strings. His criticism was directed towards our decision to neglect the entropy of the water, since most biology takes place in aqueous solutions where the entropy of the medium plays an important role in the dynamics of the system. A hydrophobic string, for example, would impose constraints on the arrangements of the surrounding water molecules resulting in a profound effect on the entropy of the system. In his words, our formulation of the question was a "purely physics view which actually does violence to your understanding of the real world." His criticism was acknowledged and disregarded. The questions were also piloted in two interviews: one with a content expert, and one with a PER graduate student.

2.3 Methodology

2.3.1 Interview Participants and Format

Interviews took place in the lead up to and the first several weeks of the Spring 2020 semester (before campus closures due to the COVID-19 pandemic). A total of eight graduate students participated, most of whom ($N = 5$) were beginning the second semester of their first year. Of the remaining three students, one was a second year, one was a transfer student with a Master's degree, and one was in their sixth year. Two of the eight participants were international students. Four students (Beth, Chris, Garth, and Harry) had previously completed a graduate-level course in statistical mechanics and two more (Daana and Erik) were taking a graduate-level statistical mechanics course in the Spring of 2020 at CU Boulder. Seven participants had a physics (or engineering physics) undergraduate background, and one had a background in computer science. All students had taken at least one undergraduate-level thermal physics course except Erik[1] , the student with the computer science background.

An overview of the individual participants is provided in Table 2.1. Our sample of students roughly reflects the demographics of the graduate student population at the University of Colorado Boulder which is a predominantly white institution that under-represents some demographic groups with respect to the larger populations of both the state of Colorado and the United States. Thus, we expect there to be selection effects that limit the range of student reasoning we will be able to observe. Interviewees were paid volunteers who responded to an email request to the physics graduate student population at the University of Colorado Boulder. All students are referred to by pseudonyms.

Interviews were conducted in a think-aloud setting [18] where participants were asked to verbalize their thoughts and reasoning as much as possible while working through the questions. The interview questions were printed on LiveScribe paper. In addition to synced written work and audio from a LiveScribe pen, interviews were audio and video recorded. An interviewer (author NC) was present to answer questions, prompt the students to verbalize their thinking, and ask students to further explain their reasoning.

[1] While Erik had not taken an undergraduate thermal physics course he was enrolled in a graduate thermal physics course at the time of the interview. As completing an undergraduate thermal physics course is not a requirement for enrolling in a graduate physics program, some graduate students do not have experience with upper-division thermal physics. So we retain Erik in our analysis to make our study more representative of the general graduate physics student population.

Table 2.1: Graduate Student Participants

Pseudonym	Notes	Thermal Research Experience (Y/N)
Alex	First year	Y
Beth	Transferred w/ MS[2]	N
Chris	First year	Y
Daana	First year, international	N
Erik	First year, CS[3] background	N
Fred	First year	Y
Garth	Second year, international	Y
Harry	Sixth year	N

Student participant pseudonyms with brief summaries of their graduate-level standing and whether they self-reported any thermal physics (that is, either thermodynamics or statistical mechanics) related research experience. [2]Masters of Science. [2]Computer Science.

The interviewer also asked students a few questions about their research experience, prior thermal physics coursework, and time in the graduate program before participants began the content related questions.

2.3.2 Analysis

To analyze graduate student responses, written drafts of the interviews were obtained from Otter.ai transcription software. The drafts were then manually checked and edited using the recorded video. Interview transcripts were emergently coded based on student responses to question prompts. Our coding methodology was inspired by Hammer's resources framework [19], and thus focused on categorizing common responses and reasoning elements to identify patterns and themes in students approaches and answers to each question. After the initial coding, the coded transcripts were collaboratively reviewed[4] to verify code assignments and reach agreement on difficult-to-code passages.

The interviews were coded and studied question by question to break the analysis up into more easily processed pieces. Summaries of student responses to each question were produced which allowed for easy identification of the range of responses as well as the commonalities and differences between the students' reasoning both within and across questions. Approaching the analysis with the resources framework allowed for inter-question patterns in student reasoning to emerge from the coded transcripts and summaries. Additionally, this approach allowed for a more direct comparison with a study by Loverude [3] which identified entropy-related resources used by undergraduate students. The resources framework facilitates an understanding of students' reasoning on entropy by providing a standardized way to discuss when and how students use particular ideas. Insight into which ideas—from all the information students hear in lecture, read in textbooks, and encounter in homework sets—students both absorb and utilize will be uniquely valuable for thermal physics instructors. We have a deeper discussion of the definition of resources used in this study in the following chapter at the beginning of Sec. 3.7 in which we will also discuss the specific resources we observed students using on the questionnaire.

[4] with Bethany Wilcox and Michael Vignal

Table 2.2: Graduate Students' Entropy Rankings of Partitioned Box States

Ranking	Initial Ranking	Final Ranking
$\mathbf{A > D = B > C}$	Chris, Erik, Garth, and Harry	Chris, Erik, Garth, Harry, and Daana
$A > D > B > C$	Fred and Daana[5]	Fred
$D > A > B > C$	Beth and Daana	-
$A > B > D > C$	-	Beth
$B > A > D > C$	Alex	Alex

Summary of graduate student entropy rankings for part A of the first question 2.1. All the rankings generated by students are listed in the leftmost column. The correct response ($A > D = B > C$) is bolded. [5]given as intermediate ranking between initial and final answers.

2.4 Results: Partitioned Box

Students' ranking of the four states (in Fig. 2.1) of the system by entropy are summarized in Table 2.2. Five out of the eight students settled on the correct ranking (for ideal gasses) of $A > D = B > C$ by the end of the question, and seven out of eight correctly identified state A as having the greatest entropy. Since the question involved ideal, non-interacting gasses rather than real gasses, the entropy of states D and B are identical. However, in the case of real gasses, whose particles take up volume, case D would have a negligibly lower entropy than case B due to a volume exclusion effect of real gas particles. In case D, the excess of particles on the left results in fewer positional microstates available to the system, since two (non-ideal) particles cannot occupy the same space. Therefore the entropy of the system of *real* gasses increase (negligibly) by moving to a macrostate with fewer circles on the left. This sense, coupled with an intuition to have the pressure difference between the two sides minimized at equilibrium, led us to expect the dominant misconception seen in this question to be ranking the entropy of state B to be higher than the entropy of state D.

The graduate students had relatively little trouble identifying states A and C as the states with the most and least entropy, respectively, but had more trouble determining where to rank states D and B. Most students (five out of eight) gave the correct ranking, but interestingly, the second most common ranking (including both initial and final) given by graduate students was to rank state D as having a larger entropy than state B, with three students giving this form of ranking a total of four times at some point in the interview. An appeal to mixing was the reason leading Fred and Daana for the entropy of state D's being greater than the entropy of state B:

> ***Daana:*** *The ones which have a more even mixture of squares and circles should have a higher entropy. And, uhh, ones which are more uniform, that is like, if they have all circles or all squares, those should have less entropy.*

Fred had a similar sense that seemed to indicate an intuition that entropy is greater the closer the particle ratio is to 50:50 on the left side of the partition.

Slightly less prevalent were rankings with state B having greater entropy than state D, with two students (Alex and Beth) giving this ranking at some point in their interviews. Beth gave this ranking initially, before re-ranking at the end of the question after recognizing state A as equilibrium and a prompt from the interviewer which asked whether anything was maximized, minimized, or equilibrated at equilibrium. Interestingly, Alex appealed to mixing as justification for the ranking of $B > A > D > C$, explicitly stating that the state with the highest entropy was the one "that is most well-mixed." This demonstrates that terms used to describe entropy like "mixing" mean different things to different people.

Graduate students utilized a variety of terms as descriptive proxies for entropy in the context of this question. The most common terms had to do with 'mixing' or that entropy had to do with how the particles were distributed. Previous research has found that even when the term 'disorder' is avoided in the physics classroom, undergraduate students still frequently use the term as a resource for thinking of entropy [3], possibly due to encountering the term in prior high school or college level science courses. Interestingly, explicit use of terms like 'order/disorder' only came up with two students on this question, possibly indicating that the term loses favor as a moniker for entropy as physics students advance to the graduate level. Other proxy terms and definitions included connections between entropy and volume/spreading out, the number of microstates, information, and 'the thing maximized in the long run.'

Common among the students giving the correct ranking was a strategy of ignoring the square particles for the purposes of ranking entropy. This indicates a competency among the graduate students to simplify problems to the most salient form and understand that the entropies of ideal gasses are additive. Fred was the only student who articulated this reasoning and did not give the $A > D = B > C$ ranking. For Fred, the presence of the squares in a sense, broke a symmetry through some 'mixing' mechanism:

> ***Fred:*** *If the squares were irrelevant, then [states B and D] would be equal... I understand that mixing two gases across a membrane should increase entropy... I think I also have to consider the volume not just mixing, so that's why I went with A for the first. But I guess, intuitively, it's more mixed in D than B, umm, so I guess if I'm considering mixing and the squares to be relevant than I would say D before B because there are more circles mixed in in D than B.*

The graduate students in the sample had no trouble ranking the pressures on either side of the partition. Most appealed to the ideal gas law, though one student uniquely reasoned through it by thinking about what the particle flux of the circles would be. When it came to confronting whether their pressure rankings were consistent with their choice for equilibrium, three students recognized and explicitly stated (correctly) that the pressure would not necessarily equilibrate between the two sides of the box. Two students experienced a little dissonance, but did not change their choice of equilibrium and accepted that the left and right pressures would be unequal at equilibrium. One of the students concisely articulated this discomfort:

> ***Harry:*** *I guess it does seem kind of strange to me if I'm claiming that [state] A represents the, umm, equilibrium case, it does seem a bit odd to me as I'm looking at it now that there would be an unequal pressure, it feels like those should be... should cancel out. But I guess there is also a gradient in chemical potential between the barriers... Umm, so, I'm not exactly sure how to think about that.*

Three students (Alex, Beth, and Fred) explicitly stated that the left and right pressures should be equilibrated at equilibrium, which is interesting since both Beth and Fred choose state A as the equilibrium state. Beth later changed her ranking (from $D > A > B > C$ to $A > B > D > C$), but not until prompted for an answer about what would be maximized or minimized at equilibrium. Fred admitted that his pressure rankings were not consistent with his choice for equilibrium and that state B would look more like equilibrium, but then expressed an uncertainty with the relationship between entropy and pressure. For Alex, the pressure ranking turned out to be consistent with his choice of state B as the equilibrium state.

Additionally, while reasoning about the pressures, two students considered whether pressure depended on the masses of the particle species, which has been noticed in prior work on undergraduate understanding of the ideal gas law [7].

2.5 Results: Blocks

Much of the previous research examining student reasoning on two-object systems have found that some undergraduates inappropriately try to conserve entropy, or over apply the second law by thinking entropy must increase everywhere [3, 6, 15]. Our framing of this task specifically cued students to think about the two blocks separately and did not make any statements about the relative sizes of the two blocks. Our question is quite similar to the macroscopic question from the diagnostic test from the study by Leinonen *et al.*, but was intentionally less structured and posed to graduate students in a think-aloud format.

In their responses to this question, all interviewees correctly stated something about how the temperatures of the two blocks will equalize to a common temperature. If this was not the first thing they considered, a statement about the temperatures occurred very early in their reasoning about the question.

When considering the entropies of the two blocks, six out of the eight students began by claiming that total entropy of the combined two-block system would increase, possibly as a way of getting oriented and grounding themselves. Notably, none of them explicitly mentioned the second law by name.

> ***Beth:*** *The entropy in the final state will be higher than it was in the initial state.* ***Interviewer:*** *The entropy of what will be higher?* ***Beth:*** *The entropy of this system.*

> ***Harry:*** *And the individual entropies... So the entropy of the entire system... the entropy is got to increase.*

A seventh student, Alex, began thinking about the total entropy before trailing off then providing answers for the individual blocks with out making any statement about the total entropy. Interestingly, Alex was the only student who explicitly claimed that the changes in entropies of the two blocks would be equal and opposite (similar to reasoning observed by by Leinonen *et al.*) implying a conservation of entropy. The eighth student, Garth, reasoned directly about the individual entropies by relating entropy to temperature. We will discuss this association in more depth later in this section (see quote from Garth below).

This pattern of first considering the total entropy was one of the more striking observations of this task, particularly since the question prompt asked students to reason about the entropy changes of the individual blocks and that students naturally and effectively thought about constituents of a system (the two gasses) independently in the previous problem. Furthermore, the second law does not provide insight into the changes in entropy of the constituents of a system, so students had to turn towards other means of reasoning about the entropies of the two blocks.

A common alternative mode of reasoning about the entropies of the blocks was through an association between entropy and temperature. Garth mentioned this idea early on in the question when considering the initial entropies of the two blocks:

> ***Garth:*** *If something is at a higher temperature, its atoms are vibrating more. So it should have greater entropy...*

and also as one of his final thoughts after realizing there wasn't enough information to know the initial entropies of the blocks:

> ***Garth:*** *So, I'm definitely associating entropy with temperature and I think if [block] A is cooling then its entropy should be less, and if [block] B is heating up its entropy should increase.*

Four students made a connection that higher/increasing temperature means more/increasing entropy, as if there were a monotonic function relating entropy and temperature. Such intuition that entropy is functionally related to temperature is not poorly grounded since this connection is consistent with the fact that the entropies of Blocks A and B will decrease and increase, respectively. Though it lead some students to the correct answer, this intuition oversimplifies the true nature of the process: the sign of the change in entropy, $\Delta S = Q/T$, depends on whether the heat flow, Q, is positive or negative.

A positive heat flow often results in increases in both temperature and entropy, but overlooking heat flow and considering temperature instead can lead to incorrect conclusions. For example, in spontaneous, endothermic, chemical reactions the heat flow into the system is positive, but the temperature can decrease while entropy increases. In an Appendix (see Sec. B), we discuss a process in the familiar context of ideal gasses that demonstrates the risk of directly relating entropy and temperature. The process involves a non-quasistatic expansion of an insulated gas in which the gas does work causing its temperature to decrease, while an increase in volume increases entropy more than the drop in temperature reduces entropy.

A non-quasistatic process, something which generally is not covered deeply in thermal physics courses, was required for this counterexample. The little emphasis placed on non-quasistatic processes, along with the fact that for most cases the heuristic that increasing temperature increases entropy is true leaves little mystery why students would lean on this association. However, two students (Alex and Fred) who made this connection between entropy and temperature reasoned that the initial total entropy of Block A was greater than the total entropy of Block B, which would be true if the two blocks were identical.

2.6 Results: Strings

The strings question (see Fig. 2.2) was analyzed in two parts. The first part consisted of the first four questions which only involved the three snapshots of a single string. The second part consisted of all the questions involving the channel with strings and circle particles. This was a natural splitting point in the question since the physical system under consideration changed between the two parts.

2.6.1 Part I: Microstates & Macrostates

In the following sections, we will present student responses by topic. First, students' responses related to microstates and probability (which corresponds to parts A, B, and C in Fig. 2.2) are discussed, followed by reasoning about macrostates (part B), then students' reasoning about the entropy of the string and whether they connected the idea to the concept of multiplicity (part D). The results presented in this section have been published in the proceedings of the 2020 Physics Education Research Conference [12].

Identification of Microstates and Probabilities

When first asked to rank the probabilities of the images in Fig. 2.2, three of the eight students correctly responded that the probabilities were equal using either the fundamental assumption of statistical mechanics or an appeal to symmetry. A total of four students gave the ranking of $P(c) > P(b) > P(a)$, which from here on will be referred to as the 'intuitive ranking'. For a summary of student responses, see Tab. 2.3.

Student reasoning behind the 'intuitive ranking' favored qualitative, vague metrics like how 'wavy,' 'wiggly,' or 'centered' (i.e., how much the string deviated from center) the string looked. In more quantitative forms, these metrics more appropriately describe macrostates. Usage of these metrics to rank microstates may suggest that students initially assumed the images represented macrostates. However, we see evidence

Table 2.3: Graduate Student Responses to Part I of Strings Question

Student	Initial Ranking (A)	Consideration in Initial Ranking (A)	Ideas for Macrostates (B)	Final Ranking (C)
Alex	$P(c) > P(b) > P(a)$	deviation from center	net deviation, conformations	$\boldsymbol{P(a) > P(c) > P(b)}$
Daana	$P(c) = P(b) = P(a)$	fundamental assumption of statistical mechanics	# of turns, net deviation, 'range' between points	$P(c) = P(b) = P(a)$
Beth	$P(c) > P(b) > P(a)$	'wavyness'	# of turns	$\boldsymbol{P(c) = P(b) = P(a)}$
Erik	$P(c) = P(b) = P(a)$	there is no preference for a particular conformation	conformations	$P(c) = P(b) = P(a)$
Chris	$P(c) > P(b) > P(a)$	'wiggliness', amount of deviation	distance between ends	$\boldsymbol{P(c) = P(b) = P(a)}$
Fred	$P(a) > P(c) > P(b)$	net 'closeness' of string to center	net deviation, conformation, # of turns	$P(a) > P(c) > P(b)$
Garth	$P(c) > P(b) > P(a)$	distance between string ends	distance between ends, position of end	$\boldsymbol{P(c) = P(b) = P(a)}$
Harry	$P(c) = P(b) = P(a)$	fundamental assumption of statistical mechanics	distance between ends, net deviation, # of turns	$P(c) = P(b) = P(a)$

Summary of graduate student responses to parts A, B, and C of the strings question (third question of the interview). Letters in parentheses on title row refer to prompts in Fig. 2.2. Bold entries in the final column indicate rankings that changed after students considered macrostates and microstates.

that students consciously or unconsciously recognize the relative probabilities of overarching macrostates, then project these probabilities onto string conformations: microstates belonging to those macrostates.

After calling the images *microstates*, Chris employed this projection when considering whether to re-rank the states in part C:

> ***Chris:*** *I guess going along the lines of the macrostates, you gotta think of which like, length of string is more likely. I would imagine the full length of the string is the least likely.*

This reasoning gets the causal link for macrostate probabilities backwards, and overlooks the equal probability between microstates. The graduate student in their sixth year, Harry, discussed the allure of this 'intuitive ranking':

> ***Harry:*** *Um, I feel like if this was actually set up in the water, and I like looked into it and saw that the string was exactly straight I would think that was really weird. But then B and C? You know, I wouldn't think that was weird. I feel like that's just how it randomly went around.*
>
> *But I think the answer to this question is probably that each of the three microstates shown here is equally probable. And if I were to like bunch those into macrostates where I say like there's a string perfectly straight or string bunched up somehow, like, I would expect it to be more likely that it was bunched up somehow. But like having it bunched up in exactly the way shown is a very small section of that particular macrostate.*

Of the four students initially giving the 'intuitive ranking', three (Beth, Chris, and Garth) later changed their ranking to equally probable after considering whether the states in the figures were microstates or macrostates. Beth made this correction spontaneously after articulating that the images were microstates so their probabilities should be equal. Chris and Garth, however, recognized the images were microstates on their own, but struggled to settle on the equal ranking until the interviewer asked if they knew anything about the relative probabilities of microstates.

After fully considering the third prompt, a total of five students had decided the figures were microstates. Alex, Erik, and Fred stated the images represented *macrostates.* In terms of probability rankings, Erik vacillated between the 'intuitive ranking' and ranking the states equally. Interestingly, Alex and Fred both settled on the ranking $P(a) > P(c) > P(b)$ based on reasoning that it was most probable for the string to have less net deviation from the center. Their reasoning, illustrated in the following quotes, suggested a conflation between a macrostate and a macroscopic object.

> ***Fred:*** *I look at that as a macrostate, because it's kind of like the combined behavior of the whole string... So macro, because macro means large.* ***Interviewer:*** *Can you talk about how you would define a macrostate versus a microstate?* ***Fred:*** *Well, I guess when I think of a macrostate, I think of like the collective behavior of a bunch of individual parts.... I don't know exactly what I could think of a microstate as, besides, like, the individual location of a piece of string and its displacement from the center.*

Alex had similar reasoning, though not as explicitly:

> ***Alex:*** *So we're talking about like positions in this string like each of these individual positions I would consider as a microstate. Whereas the macrostate is, like... the configuration itself.*

This choice of the straight string in Fig. 2.2a as the most likely configuration was unexpected and resembles findings from Loverude and Geller that some students have an unexpected association of equilibrium (and high entropy) with an 'ordered' state [3, 20]. While our interview question did not directly relate to equilibrium and students did not bring up ideas about equilibrium, it is possible students followed a similar line of reasoning; though, instead of associating equilibrium with an 'ordered' state, students here ranked an 'ordered' state with a higher probability of occurring. It is important to note, however, that both Alex and Fred arrived at the conclusion that the straight string was the most likely configuration as a logical conclusion of their 'centered-ness' metric. Further study is warranted to determine whether the similarity between Loverude's finding and ours is more than superficial.

Definition of Macrostates

Defining macrostates for the string required some creativity, and all the students interviewed articulated appropriate classification systems (see Tab. 2.3). A total of five students brought up multiple relevant ideas for defining macrostates, though not all represented fully formed definitions.

Macrostate classifications given by the interview students fell into one of five emergent groups: the net deviation of the string from center (four out of eight); a displacement between two points on the string (four out of eight); the number of 'turns,' 'kinks,' or 'bends' in the string (four out of eight); the actual conformation of the string shown in the figures (three out of eight); and there was one mention of the location of the string's end—a position as opposed to a displacement.

In the three cases where the actual conformation of the string was identified as a macrostate, two answers revealed conceptual errors in differentiating microstates and macrostates, as discussed in the previous section. The other response, from Erik, was accompanied by an appropriate argument that microstates were the precise arrangement of atoms and electrons in the string: that there are many ways to arrange the atoms and electrons in the string to form that specific conformation. While not incorrect, this does reveal an unproductive scale from which to reason about this system.

For the majority of students who initially gave an unequal probability ranking (Alex, Beth, Fred, and Garth), the initial ranking metric went on to form the basis for their macrostate definitions. For Alex and Fred, ideas about net deviation and net 'closeness' to center lead to macrostate definitions involving the net deviation of the string to one side or the other. Beth's metric of 'wavyness' preceded a definition based on the number of turning points on the string. Garth's reasoning about net expected displacement in a random walk foreshadowed a definition based on the distance between the two ends of the string.

The conversion of these metrics into macrostate definitions demonstrated a skill in turning intuition into models. It also demonstrated an expert-like ability to shift into more correct modes of reasoning. However, this also supports our interpretation that students, at least initially, projected properties of macrostates onto constituent microstates.

Reasoning about the Entropy of the String

Student responses to this question reflected the open-ended nature of the prompt. And the prompt itself could be confusing depending on how students interpreted the question. For example, if students thought it asked about the entropy of the string in a particular microstate, which is a bit of a red herring. In any case, this proved the most difficult task of those discussed here, which echoes findings from Loverude and Bucy that students struggle connecting entropy and multiplicity [3, 21]. Four students responded in some way that related multiplicity of states to entropy. Three students clearly articulated this connection.

For example,

> ***Daana:*** *Entropy of the string [...] in some macrostate would be related to the number of microstates in which that system belongs to that macrostate.*

A fourth student came close to this answer, but didn't articulate the connection as concisely as the others:

> ***Chris:*** *Like so, basically what I'm thinking is there's like, there's, I guess theoretically infinite ways that the string could be at this length, but it has to lie on that semicircle. Whereas there's even more infinite ways that the string could be this length because it could snake in some other weird way. And I suppose this length is slightly shorter. I think that's what they were going for. There are slightly more ways that it could sort of snake into that position, I guess is what they're kind of asking.*

Two students, Erik and Fred, reasoned that the entropy of the string was a constant. Erik argued that since each of the conformations were "not structurally distinguishable" that there was some kind of equivalence between each conformation that meant entropy didn't change if the conformation of the string changed. Fred stated he couldn't think of a reason why the entropies of the states should be different. He recognized that this contradicted his unequal ranking of the states, but ultimately stuck with the ranking for the sake of consistency with the 'centered-ness' argument. These students neglected the deeper connection between entropy and multiplicity, but seemed to nearly activate the fundamental assumption of statistical mechanics.

The final two students, Alex and Beth, both reasoned about entropy by breaking the string up into chunks and considering which way each chunk bent (i.e., did the chunk deviate left, right, or not at all). Alex reasoned that a "mixed" sequence of chunks was more favorable, and Beth did not expect most of the chunks in a sub-sequence to 'do the same thing'. For Alex, the idea came from an analogy with a two-state Ising magnet system with which he was familiar.

Given that other researchers have found a strong preference for the association of entropy with 'disorder' among undergraduates [3, 4], perhaps it is surprising that only two graduate students employed this particular association when asked directly about the entropy of the string. However, terms of similar ambiguity like 'wavy,' 'wiggly,' and 'chaotic' were used by students in this study, though mostly as a qualitative handle in developing more precise descriptions. The fact that half of the students initially ranked the more 'disordered' states (i.e. the two states that were not perfectly straight) as more likely configurations suggests at least some association between probability and 'disorder.'

2.6.2 Part II: Channel

In part E, we sought to see how well the students could connect the ideas about entropy and conformations to a more complicated system involving the strings. When considering what happens to the number of conformations of the strings, seven students (correctly) decided it would decrease when the circles and strings interact (See Table 2.4 for a summary of each student's responses). The primary explanation for the decrease, given by six students, was that the circle particles would occupy space/exclude volume from the strings, which will limit their conformations. The seventh student, Erik, decided to recast the system into the more familiar context of an ideal gas expanding in to a vacuum, then claimed the decrease would occur transiently, as the circles created a sort of wind that would blow the strings into rightward pointing conformations.

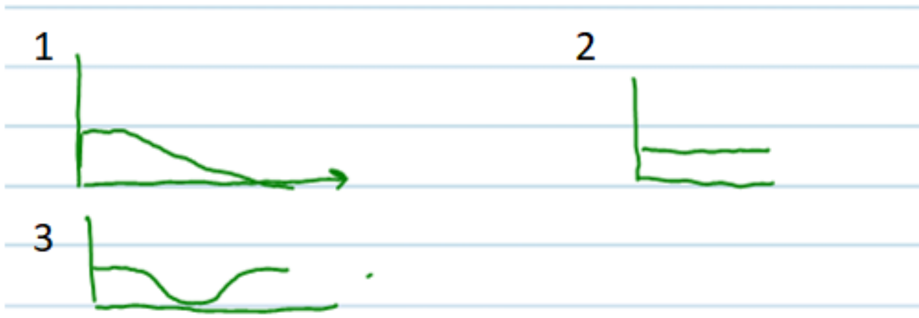
Erik: *I would expect gale force winds in this direction [right] which would limit the number of conformations of the strings because all the strings will be flapping in that [right] direction...*

What do I expect to the possible conformations of the string when I remove this barrier and let the gas [circles] through: I expect it to decrease because all the strings are going to curve to the right.

Alex was the only student who did not explicitly state the number of conformations would decrease. Like Erik, he made a claim on the conformations themselves rather than the number of conformations. Alex invoked the idea of pressure, similar to Erik's idea of wind. When drawing concentration profiles, both Alex and Erik argued for flat, uniform profiles as would be the case without the strings, though Alex did briefly consider a profile with a dip in the middle region.

On the whole, five students (not including Alex) decided the concentration profile would have a dip, which we considered the correct answer, similar to Beth's third sketch in Fig. 2.5. Beth's sketches, coincidentally, make a good summary of the whole field of responses. Several students took a moment to consider a transient profile (similar to Beth's first sketch) early in the system's evolution despite no cue to do so. Two other students (Garth and Harry) explicitly considered the concentration profile for a system

Figure 2.5: Concentration Profiles Drawn by Beth



Three profiles drawn by Beth while answering the prompt about the final concentration profile. The numbers indicate the order in which the plots were drawn. When drawing the second plot, Beth was considering a case in which there were no strings. In the third sketch, Beth said she would not expect the concentration profile to go completely to zero in the middle region despite drawing it that way.

without the strings. Furthermore, as was mentioned previously, Alex and Erik argued for uniform profile matching Beth's second sketch.

A sketch drawn by Chris, however, was unique. Chris argued for a concentration profile of circles that was peaked in the middle region (see Fig. 2.6). He argued that the circles would get trapped in with the strings causing the flux of circles out of the middle region to be less than the flux into the region. This profile would occur if the strings and circles could bind strongly enough to release a sufficient amount energy to increase the overall entropy of the system.

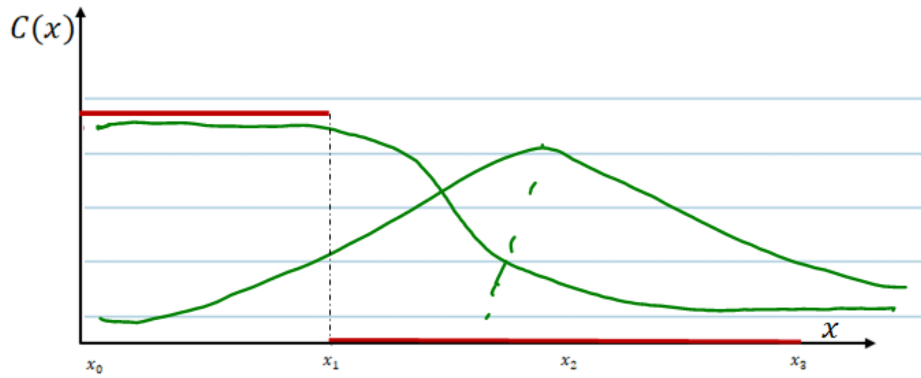
When asked to discuss what the decrease in configurations of the strings due to the presence of the circles (which Chris calls "molecules") meant for the entropy of the strings, Chris concisely articulated the relevant effect:

> ***Chris:*** *The entropy of the strings has to go down because there's fewer states that they can access. But the entropy of the molecules goes way up because, you know, once I take that wall out, there's a ton more places the molecules could be.*

This reasoning did not cause Chris to change his profile, however, despite one last question from the interviewer asking whether the peaked profile was his final answer.

Throughout the discussion of the question, seven students made a connection between the number of configurations and the entropy of the strings, though three of them did so only after a suggestive question

Figure 2.6: Peaked Concentration Profile Drawn by Chris.



The slanted, mostly vertical dotted line was drawn by Chris to indicate the maximum of the peak should be at the center of the channel.

Table 2.4: Graduate Student Responses to Part II of Strings Question

Student	Number of Conformations (i)	Reasoning (i)	Profile Drawing (ii)	Entropy Connection
Alex	deviate right	pressure of circles 'push' strings	flat, briefly considers dip	None (not prompted)
Daana	decrease	circles exclude volume	dipped	spontaneous
Beth	decrease	circles exclude volume	dipped	prompted
Erik	decrease	deviate right by pressure of circles	flat	prompted
Chris	decrease	circles exclude volume	peaked	prompted
Fred	decrease	circles exclude volume	dipped	spontaneous
Garth	decrease	circles exclude volume	dipped	spontaneous
Harry	decrease	circles exclude volume	dipped	spontaneous

Summary of graduate student responses to part E of the strings question. Lower case roman numerals in parentheses on title row refer to prompts in Fig. 2.3. For example sketches of the flat and dipped profiles drawn by students see Fig. 2.5.

from the interviewer. Alex did not make this connection, and Erik made the connection after some prompting, but still settled on a flat concentration profile.

Harry's discussion of the entropy trade-off between the circles and strings lead to a direct connection to a kind of "force tending to push the balls out of the region" with the strings. Indeed, there is an analogy between the change in entropy of the strings (ΔS) associated per loss in volume (ΔV) due to a circle which parallels the thermodynamic definition of pressure: $P = T\frac{\partial S}{\partial V}$.

This system with the strings, circles, and channel is also directly analogous to the first question on the interview with the partitioned boxes. The circle and square particles in the partitioned box question play similar roles as the circular particles and strings, respectively, in the channel system. In both systems the circular species are free to move throughout, while the square particles and strings are confined to one single region. The key difference between these two systems, however, is that the species in the channel question take up volume, unlike the ideal gases in the partitioned box system which are point-like. As discussed in Sec. 2.2, the non-trivial volume of the circles and strings is what results in the entropic force causing the dip in the concentration of the circles in the middle region. If, instead, the circles were point-like and the strings were line-like, the circles would have a uniform concentration profile, similar to how the circle ideal gas particles in the partitioned box question equally distribute themselves between the two halves of the box despite the presence of the square particles.

The vast majority of graduate students picked up on the relevant fact that the circles in the strings question took up volume, which lead to a non-trivial concentration profile. This is noteworthy, especially since students also were able, for the most part, to reason correctly about the ideal gases in the first question, which were also drawn (misleadingly) with non-zero volume. None of the students recognized the connection between these two questions, though Chris and Daana explicitly asked the interviewer whether they should consider the 2D area of the circles as they were drawn. Nonetheless, the ability of the graduate students, as a whole, to recognize the most salient features of each question and reason accordingly given the different situations demonstrates well-developed physical reasoning skills.

2.7 Results: Expansions

The dependencies of entropy on volume and temperature have emerged as two themes woven through this set of interview questions. While the preceding tasks only required students to think about one dependence at a time, this task focusing on expanding gases required students to contend with both the volumetric and thermal aspects simultaneously. Several students commented on recognizing either the isothermal or free expansion in a prior context:

Beth: *I'm really, really trying to remember... I feel like I've talked about this exact problem in a class before, and I remember it being a very obvious thing that I just don't remember*

Despite the students' familiarity with the expansions, this task proved to be the most challenging of the interview. Students struggled to form sound arguments, they expressed less confidence in their answers, and their responses were largely idiosyncratic with little consistency from student to student.

In their study exploring this task, Bucy *et al.* gave this prompt to students both at the beginning and end of an upper-division thermal physics course. Generally, they found students relied heavily on mathematical equations to approach this task. On the post-test, they found that undergraduates applied the thermodynamic definition of entropy to correctly identify the sign of ΔS in the isothermal expansion, but struggled with determining the sign in the free expansion. Many students on the post-test made a claim that the temperature of the gas would decrease, despite the prompt explicitly stating the temperature was constant[6] . On the comparison task, they saw only two (out of seven) students correctly identify that the changes in entropy were identical between the two processes on the post-test, with only these two students explicitly using entropy's state function property.

Our study found very similar trends with how graduate students responded to this question. On the first task of determining the sign of the change in entropy, ΔS_{iso}, of the gas undergoing the isothermal expansion, most students reasoned it would be positive. However, only two students were able to both activate and successfully reason by using the thermodynamic definition of entropy: $\Delta S = Q/T$. Most other students relied on qualitative intuitions on how entropy is related to volume and temperature. The intuition that entropy depended on temperature lead Beth to her answer of $\Delta S_{\text{iso}} = 0$ (since the temperature was

[6] In our version of this question, we did not give students this hint

Table 2.5: Graduate Student Responses to the Expansions Question.

Student	Sign ΔS_{iso}	Sign ΔS_{FE}	Entropy Change Comparison	Claimed $\Delta T_{\text{FE}} < 0$
Alex	+	+ → - → 0 → +	$\Delta S_{\text{iso}} < \Delta S_{\text{FE}}$	yes
Daana	+	0 → +	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	no
Beth	0	-	$\Delta S_{\text{iso}} < \Delta S_{\text{FE}}$	yes
Erik	0 → +	+ → 0 → + → 0	$\Delta S_{\text{iso}} > \Delta S_{\text{FE}}$[7]	yes
Chris	+	0 → +	$\Delta S_{\text{iso}} > \Delta S_{\text{FE}}$	yes
Fred	+	+	$\Delta S_{\text{iso}} < \Delta S_{\text{FE}}$	no
Garth	+	0 → +	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	yes
Harry	+	0	$\Delta S_{\text{iso}} > \Delta S_{\text{FE}}$	no

Summary of graduate student responses to the expansions question. In this table, the quantities ΔS_{iso} and ΔS_{FE} are the magnitudes of the changes in entropy gases undergoing the isothermal and free expansions (respectively). ΔT_{FE} is the change in temperature of the gas undergoing the free expansion. A "yes" in this column indicates the student considered this at any point while reasoning about the task, regardless of their final conclusion. [7]inferred from final answers to prior questions constant). Erik's initial response of $\Delta S_{\text{iso}} = 0$ came from recognizing this expansion as a reversible process, where $\Delta S_{\text{tot}} = 0$, and thinking this meant that ΔS_{iso} had to be zero (he later corrected this response).

Again mirroring Bucy's findings, the graduate students we interviewed struggled more with determining the sign of ΔS_{FE}, the change in entropy of gas undergoing the free expansion. This is somewhat surprising, since the free expansion is a canonical example of an irreversible process (in which entropy must increase), but six out of the eight graduate students interviewed at one point in the interview stated that ΔS_{FE} was zero. Two students even considered that it was negative due to an intuition that the gas would cool and that entropy was related to temperature.

Three students (Garth, Harry, and Daana) considered that ΔS_{FE} was zero after realizing that the change in energy of the gas was zero. Daana and Harry, the only two students to explicitly invoke the equation $\Delta S = Q/T$ in the prior part, incorrectly applied it to the free expansion (though Daana later corrected). This could be due to an implicit assumption that the process could be treated as quasistatic where the relationship equating the change in entropy to heat flow is true, a possible artifact of a heavy emphasis on quasistatic processes in many thermal physics courses.

Garth reasoned about the free expansion more conceptually by thinking about how changes in the volume and temperature would affect the entropy:

Garth: *Although the gas has more volume, its temperature is decreasing.*

This intuition that an expanding gas must cool was brought up by most of the students at some point while engaging with this question, and five brought it up while thinking specifically about the free expansion. Chris justified this intuition with a real-world experience:

Chris: *It ends up colder, I guess, by the gas law. In the same way that your propane gets cold when you turn on your stove.*

Many scientists are likely to admit having a similar, visceral sense for an expanding gas to cool. While a free expansion is an idealization and special case that can not truly happen in nature, it helps to illustrate the cause for temperature changes of expanding and contracting gases: work on the environment. However, it should be noted that in the case of a propane tank releasing a gas, the primary effect causing the tank to cool is the latent heat of the endothermic phase transition of liquid propane vaporizing.

Only Daana, Erik, and Garth explicitly identified the free expansion as an irreversible process. For Daana and Garth, this realization seemed to unlock the whole problem since it preceded them correcting their answers for the sign of ΔS_{FE} and correctly comparing the changes in entropy of the two processes. In Erik's case, it did not provide sufficient confidence to finalize his choice of sign, and he went on to change his choice multiple times before finally settling on $\Delta S_{\mathrm{FE}} = 0$.

As can be seen in Table 2.5, there was very little consensus when it came to comparing the magnitudes of ΔS_{iso} and ΔS_{FE}. In addition to recognizing the free expansion as an irreversible process, Daana and Garth were the only two students who recognized that the two gases started and ended in the same initial and final states, respectively. No students explicitly mentioned that entropy was a state function, though Daana and Garth demonstrated an understanding of the relevant reasoning. For the students saying that ΔS_{iso} was greater, most arrived at the conclusion as a result of deciding that ΔS_{FE} was zero. Alex and Fred settled on ΔS_{FE} as being greater with both mentioning the ability of the gas to exchange energy with, and lose energy to, the environment. For Beth, the magnitude of ΔS_{FE} was greater simply as the logical conclusion of deciding that ΔS_{iso} was zero.

Students fared better on the question about the changes in entropy of the surrounding environment. All students recognized that the environment outside of the thermally isolated free expansion would be zero since it was decoupled from the gas and no heat could flow. Only six students, though, concluded that the entropy of the environment outside the gas isothermally expanding was negative. Fred claimed that the entropy of the environment would increase, and Erik struggled separating the changes in entropy of the gas, outer environment, and universe, a difficulty also found in prior studies [4, 15].

2.8 Discussion: Graduate Students' Conceptualizations of Entropy

The interview transcripts provide two ways of examining the graduate students' conceptualizations of entropy. First, we can study the way students talked about entropy while solving the content-problems. Second, we can analyze their answers to the final interview question which directly asked them how they conceptualized entropy. A deeper discussion of the graduates' conceptualizations will come in following chapter (in Sec. 3.7), but to conclude this chapter we will first highlight some trends among the content questions, then discuss the students' responses to the final question of the interview.

Two students had unique conceptualizations of entropy that they brought up multiple times in their interviews. First, Alex had a strong affinity for thinking of entropy as information, and spontaneously brought up the term in the Partitioned Box and Expansions questions.

Alex: *Entropy itself is directly related to information by the Shannon Entropy.*

This proxy did not appear to help in the Partitioned Box task since Alex settled on an incorrect entropy ranking, but this intuition did appear to help him realize that the entropy of the freely expanding gas increased in the Expansions question.

The second student, Fred, had an affinity for the term chaos, bringing it up spontaneously in the Blocks, Strings, and Expansions questions. However, despite raising this association in the Strings question, Fred defied this intuition. He ranked the probabilities of the arrangements as $P(a) > P(c) > P(b)$ based on an expectation that strings with less deviation from the center were more likely. Furthermore, he admitted to thinking that the *entropy* of the three configurations were the same despite his intuition that the more "squiggly" strings would have more entropy.

A consensus from our interviews with thermodynamics instructors lead us to expect that when asked about entropy students would default to the mantra that entropy is disorder. Indeed, students brought up the idea of disorder, but much less than expected. On the four physics content questions, Beth used the term in the Partitioned Box and Strings questions, Garth used the term in the Blocks question, and Erik used the term "ordered" in the Partitioned Box question. Two additional students, Daana and Fred, brought up disorder when directly asked how they conceptualized entropy. Exhaustively, that is a total of five students mentioning either disorder or order in six different situations (nine situations if Fred's use of the term 'chaos' is included). This observation echoed a finding in Leinonen's study which noted that few undergraduates used the association of entropy with disorder [6].

To close the interviews, we asked students a fifth question: how they personally conceptualized entropy. As a proportion of the total time, the amount of discussion on students conceptualization of entropy was much less than any single content question, but was the most common place for students to bring up the association between entropy and disorder. When taken with the rest of the results of the interviews, this suggests that graduate students have a robust and somewhat varied set of tools with which to practically reason about entropy, but still have trouble conceptualizing the quantity abstractly.

The most common response, given by six students, in some way mentioned that entropy was related to the number of configurations or (micro)states. The two remaining students, Alex and Fred, mentioned information and disorder, respectively. In addition to Fred, only Daana brought up that entropy was related to disorder on this question, but went on to say that disorder was an incomplete way of thinking about entropy. Three students, Alex, Beth, and Erik, admitted to struggling with conceptualizing entropy.

Alex's unique association between entropy and information also seemed to have a connection to prior research experience. When explaining his conceptualization, Alex discussed how it applied in the Ising magnet model, a system with which he had prior research experience. Garth also had a somewhat unique association of entropy as "the arrow of time" since systems always evolve towards higher entropy, though Garth's association could be a manifestation of the "entropy of a system must always increase" resource. Fred had a similar association of entropy as something that always increases or stays the same.

Overall, the strong preference for conceptually relating entropy to the multiplicity of microstates mirrors a finding from Bucy's study that finds undergraduates favored the statistical definition of entropy over the thermodynamic description [4]. This stated preference for microstates is consistent with our pool of students frequently thinking about microstates while solving problems and thus demonstrates a high level of metacognitive awareness indicative of more advanced physicists [22].

Chapter 3

Undergraduate Student Reasoning About Entropy

3.1 Introduction

The second half of our study on student reasoning with entropy investigated undergraduate reasoning using the same set of interview questions used with the graduate students. As noted in the previous chapter, entropy is a notoriously difficult concept for students to learn so by extending the study to include undergraduates, we increase the perspective on student reasoning about this enigmatic concept. In this chapter, the undergraduate responses will be presented and discussed question-by-question along with some comparison to the graduates' responses. Then, we will discuss broader themes which emerged throughout both the graduate and undergraduate interviews where we will draw more comparisons between the two groups of students interviewed in this study. The themes we found woven throughout the interviews will be discussed within the conceptual resource framework [19]. This framework will be described in more detail in Sec 3.7, but briefly, a conceptual resource is a 'tool' or 'element' of reasoning (that, depending on the context, may or may not be appropriate). In this part of the chapter on resources, many comparisons will be made between the graduates' and undergraduates' approaches to the interview questions.

Though we make comparisons between the graduates' and undergraduates' reasoning, this research was not intended to make quantitative claims or draw general conclusions about differences between the two populations. Due to the small sample size and selection effects caused by mostly interviewing students at the University of Colorado Boulder, any comparisons made refer specifically to the two sets of students interviewed. When numbers of students giving a particular answer are presented, it is to satisfy the curiosity

Table 3.1: Undergraduate Student Participants

Pseudonym	CU Student (Y/N)	Thermo Course	Interview Semester
Abby	N	Sp20	Fa20
Bill	N	Fa20	Fa20
Carson	Y	Fa20	Fa20
Dewei	Y	Fa20	Fa20
Emma	Y	Fa20	Fa20
Frank	Y	Fa20	Fa20
Greg	Y	Fa20	Fa20
Henry	Y	Fa20	Fa20
Isaiah	Y	Sp21	Sp21
Jessie	Y	Sp21	Sp21

Undergraduate participant pseudonyms with information on the timing and context of their thermal physics course and interview.

of the reader and not to make any statistical claims. Furthermore, references to 'graduates,' 'graduate students,' 'undergraduates,' or 'undergraduate students' in this chapter refer specifically to the interviewed students and not to the larger populations of graduate and undergraduate populations.

Due to the COVID-19 pandemic, think-aloud interviews with undergraduate students occurred via Zoom. The questions were shared via a Google Jamboard, which allowed both the student and interviewer to view and edit a single document without having to share screens. The interviews were recorded by Zoom, which automatically produced transcripts of the interview. As with the graduate interviews, the generated transcripts were manually checked and edited against the video recording of the interview. The transcripts were then coded with a combination of *a priori* codes generated from the analysis of the graduate student interviews and new emergent codes based on the undergraduates' responses to the questions to capture new ideas and patterns not observed in the graduate interviews.

In total, ten undergraduates were interviewed: eight interviews were conducted in the Fall 2020 semester, and the remaining two were conducted in the Spring of 2021 (see Table 3.1). An interviewer (NC) was present to answer questions from the interviewees, prompt the students to verbalize their thinking, and ask students to further explain their reasoning. Before beginning the interview questions, students were encouraged to ask clarifying questions, to verbalize their thinking, and to read the interview questions out loud.

Table 3.2: Undergraduate Students' Entropy Rankings of Partitioned Box States

Ranking	Students
$\mathbf{A > D = B > C}$	Bill[1], Dewei, Emma, Frank, Greg, Henry, Jessie
$D > B > A > C$	Carson
$D > A > B > C$	Isaiah
$B > A > D > C$	Abby
$B > A > C > D$	Bill[2]

Summary of undergraduate entropy rankings for part A of the partitioned box question. All the rankings generated by students are listed in the leftmost column. The correct response, $A > D = B > C$, is bolded. [1]initial ranking, [2]final ranking.

The first two students interviewed in the Fall 2020 semester were conducted with non-CU students. Abby was a senior at a medium-sized research university (and was the only student who had fully completed a thermal physics course prior to the interview), and Bill was senior at a large R1 university. The remaining eight students (six from the Fall of 2020, and two from the Spring of 2021), were students currently enrolled in the thermal physics course at the University of Colorado Boulder. All relevant content had been covered and tested in the course by the time students participated in the interviews. Both semesters were taught by the same instructor. All students are referred to by pseudonyms.

3.2 Results: Partitioned Box

The undergraduates' rankings of the entropies of the four states in Fig. 2.1 appear in Table 3.2. Six out of the ten undergraduates settled on the canonically correct ranking of $A > D = B > C$ as their final answer. There was no single second ranking given more frequently by students since all four of the remaining undergraduates gave unique rankings. The four other rankings, however, can be grouped into a dichotomy of state D ranked highest (Carson and Isaiah) or state B ranked highest (Abby and Bill).

The two students ranking state D as having the largest entropy both based their reasoning on an association between entropy and mixing. For Carson, the association with mixing connected to the idea that entropy is related to the number of available microstates:

> ***Carson:*** *You want to maximize the amount of microstates you have, and in system D, by having the most circles on the left hand side [in the side with the squares], I would think that's when that system*

is fully mixed.

This reasoning about mixing almost lead Carson to the conclusion that the state with all 20 circle particles on the left (with the squares) would have the most entropy, but we walked this back after realizing that would leave the right side of the box empty and in a state with very low entropy. Isaiah had very similar reasoning:

Isaiah: *D should be first because it has the most, the most particles mixed together. What that tells me is that, in terms of the multiplicity there should be the most number of microstates for each particle.*

Despite ranking state D as the state with the most entropy, when answering the second part of the question (which state most closely resembles the equilibrium state) Isaiah claimed state B was the equilibrium state. In his reasoning he seemed to be attempting to balance the idea of mixing with the idea that the entropy of a side will increase if the number of particles on that side will increase. When asked in the final question whether his answers were consistent, Isaiah even invoked the second law of thermodynamics but did not notice the inconsistency.

Both Isaiah and Carson were aware they had to consider the entropy on both sides of the partition, but considering the amount of mixing seemed to distract them from the fact that each gas will increase its entropy independently of the other gas by having its partial pressure equilibrate across the partition. This association of entropy with mixing seemed to be over-applied by Carson and Isaiah since the actual equilibrium state, A, is (in a sense) more 'mixed' than the state with equal pressures (state B) which we expected would be the most common distracter. Though the idea of 'mixing' may be helpful in reasoning about entropy in certain contexts, we generally observed students over-applying the concept which lead them to incorrect answers. See Sec. 3.7.3 for a deeper discussion of how we observed students use this idea of mixing.

Two students, Abby and Bill, gave rankings that put state B as the highest entropy state. Abby reasoned from an intuition that the lowest entropy state would take the most "effort" to organize. For example, state C would take the least amount of effort to organize, so it has the least entropy. This reasoning is very closely related to the idea that entropy is related to disorder. In fact, disorder was one of the first ideas mentioned by Abby on this question, though she immediately gave a caveat that she did not think it was exactly appropriate to the question.

Bill originally ranked the states correctly, stating that the circle particles should be "equally distributed on both sides." Despite this reasoning, Bill expressed some uncertainty over the effect the square particle species would have on the entropy of the system. After considering the question on pressure, however, Bill became much more certain that the entropy and equilibrium state of the system would be determined by the pressure difference between the two halves of the box. Bill generated his final ranking after making this claim, and a prompt from the interviewer asking whether he wanted to re-rank the states. So, despite the similarity in rankings between Abby and Bill, their primary modes of reasoning were different.

In similarity with the interviewed graduate students, few undergraduates changed their initial ranking after considering the pressures in the two compartments of the box. Only one student, Bill, changed their answer and it was from the correct response to the ranking of $B > A > C > D$. Additionally, on the whole students were generally able to identify state A as the as the equilibrium state of the system. All but one (seven out of eight) of the graduate students choose A as the equilibrium state and six of the ten undergraduates did as well. Both groups also had little trouble identifying state C as the one with the least entropy.

Overall, the undergraduates in our sample also had more difficulty than the interviewed graduate students determining which half of the box would have a higher pressure. Four of the undergraduates began this question by attempting to derive a differential equation with which they could solve for the pressure, or to reason based on the thermodynamic identity ($dU = TdS - PdV$). Most of the undergraduates realized they could simply apply the ideal gas law ($PV = NkT$ or $PV = nRT$), though at least two undergraduates struggled to reason productively about the pressures.

In reasoning about state A, Greg rated the sides as having equal pressures, despite reasoning the pressure would be related to the number of collisions against the walls. If Greg was considering only the partial pressure of the circle species, this rating would be correct, but if he was considering the partial pressure he did not explicitly state so. Henry tried reasoning through the thermodynamic identity and the counter-factual situation in which the barrier could move. He eventually stated that the left side (the side with the squares) would have the higher pressure in every case. In contrast to Greg, Henry's conclusion is consistent with only considering the partial pressure of the square species. A third student, Dewei, also rated both sides in state A as having equal pressures. However, due to a language barrier, it was unclear whether his answer was for the total pressure or for the partial pressure of just the circle species.

3.3 Results: Blocks

The interviewed undergraduates' reasoning on the blocks question was very similar to that of the interviewed graduate students. Though perhaps unsurprisingly, the undergraduates had a few more struggles than the graduates. All but one of the undergraduates made a statement that the temperature of the two blocks would equalize after a long period of time. Most undergraduates also correctly identified that the entropy of the hotter object would decrease and the entropy of the colder object would increase. Their reasoning also largely paralleled that of the graduate students. Students in both groups had a strong association that entropy was positively correlated with temperature and it was not unusual for a student to bring up the concept of microstates and consider what was happening to the number of accessible microstates. For example, when thinking about the entropy of block A, Jessie stated:

Jessie: *...because it's going to be* k *natural log of the multiplicity, um, and there's a lot more ways that* T_A *initial has to move its stuff around, so* S_A *initial is going to be greater.*

As with the graduate students, the undergraduates' intuition that increasing temperature increases entropy and that decreasing temperature decreases entropy was not strongly connected to the Clausius Theorem: $dS = \delta Q/T$. Three undergraduates did mention this equation, though only two, Dewei and Henry, seemed aware of the nuance that the sign of the change in energy Q determined the sign in the change in entropy. More often, however, students would either gloss over the fact that heat (i.e. energy) was exchanged between the two blocks and default to the reasoning that temperature was associated with entropy, or discuss how the increase in energy of cooler block would allow the block to access more microstates, thus increasing its entropy. Carson, who employed this latter reasoning, drew a connection with the hydrogen atom which has more states accessible at higher energies (since the electron can reach states with higher principal quantum numbers n).

A minor difference between the undergraduate and graduate responses was that more undergraduates (five out of ten) explicitly mentioned the second law of thermodynamics when reasoning about this question. Graduates generally did not mention the law by name, but would state that the entropy of the system would increase. This difference could have resulted from the undergraduates' closer contact with thermal physics concepts since most of them were taking the course concurrently with the interviews.

Another observation unique to the undergraduates was that they often stated that the entropies of the two blocks would be equal in the final state. In many of the cases where a student made this claim, the interviewer would ask if that would be true if the blocks were not identical. In every case the question was asked, the students would realize they had implicitly assumed the blocks were identical, perhaps because the question was so general and making this assumption helped more specify the problem.

Two of the students (Bill and Greg) who claimed that the entropy of the two blocks would equalize, seemed to conclude that the entropy of system would remain constant. This implies a conservation of entropy in this system, a claim students have been observed making in other studies [3, 15]. The conclusion (that entropy would remain constant) from Bill and Greg also seemed related to a separate difficulty in distinguishing between systems, surroundings and the Universe [4, 15, 23]. In Bill's case, he seemed aware that he was struggling to think about the two blocks separately:

> ***Bill:*** *It's hard for me because I don't want to think about, like, the entropy of the individual objects when they are together. I think about, like, the entropy of the system at the end.*

Greg seemed to conclude the entropy remained constant despite explicitly mentioning the second law and stating the entropy might increase. After being asked about what would happen to the total entropy of the system Greg responded:

> ***Greg:*** *Um, increases. So, it should be greater than or equal to S_A plus S_B. Um, it depends on the whole situation, but most likely increase, I think. And that's because of the second law. Well, It's isolated. Okay, well maybe since it's isolated... the entropy of the universe, there's no change, but the total entropy should still be S_A plus S_B.*

After making this statement he proceeded to write $S_{tot}(\text{before}) = S_{tot}(\text{after})$. When using the terms 'universe' and 'total,' it is unclear whether these both refer to the entropy of the system, or whether one of the terms might refer to the entropy of the entire Universe[3] . This ambiguity is very similar to an observation made by Christensen *et al.* that students sometimes struggle to distinguish the entropy of systems from the the entropy of the environment [15]. For a deeper discussion of this and possible implications for instructors, see Sec. 3.8.1.

We can make another interesting comparison to prior research. In his study involving a nearly identical question about two blocks coming to thermal equilibrium, Loverude observed that students struggled to connect changes in temperature with changes in entropy [3]. While we also observed some students struggle, most had a strong association between how changes in temperature relate to changes in entropy, however, it was not (at least directly) grounded in the Clausius Theorem which is the most relevant relationship to this physical situation.

In this simple scenario with the two blocks, this association between entropy and temperature generally lead students to correct responses. However, relating entropy directly to temperature is an oversimplification. Generally, a thermodynamic quantity like energy or entropy, depends on more than one other independent variable. In the context of this question, the entropy of the blocks could be Einstein solid depends on the number of oscillators and the energy. Since the number of oscillators in an object typically remains fixed, this allows for the simplification that entropy is related to only one other independent variable (either temperature or energy). In Appendix B, we discuss a case where changes in two independent variables must both be considered when reasoning about the change in entropy of a system.

3.4 Results: Strings

Similar to the discussion of the graduate students responses to the strings question (in Sec. 2.6), we broke the analysis of the undergraduates' responses into two parts. In this chapter, these two parts reflect the same split made when discussing the graduate student responses. The first examines the responses to the questions about the three snapshots of the single string, and is further broken down by topics as described in Sec. 3.4.1. This finer breakdown of the first part of the question does not exactly match the breakdown made in the graduate interviews due to differences in how the two groups responded to the question. The

[3] In this thesis, when capitalized word 'Universe' refers to all of existence.

second part (discussed in Sec. 3.4.2), just as with the graduate students, examines responses to the questions about the channel with strings and circle particles.

3.4.1 Part I: Microstates & Macrostates

In the following sections, we first present students' rankings of the probabilities of the three string configurations (covering student responses to parts A and C of the strings task), then we explore the students' reasoning about microstates and possible macrostate definitions (covering parts A and B). Finally, we discuss the students' reasoning about the entropy of the string and the extent to which they connected the idea of entropy to the concept of multiplicity (part D). The research discussed in this section first appeared in an article published in the Proceedings of the 2021 Physics Education Research Conference [13].

Probability Rankings

As summarized in Tab. 3.3, five out of ten students settled on the correct ranking in which all states are equally likely. Furthermore, all five of these students (Emma, Frank, Greg, Henry, and Jessie) alluded to the fundamental assumption of statistical mechanics by correctly claiming that all microstates of a system are equally likely, though none of these students referred to the fundamental assumption by name. This proportion was roughly similar to the proportion of graduate students implicitly or explicitly utilizing this idea.

Initially, the second most preferred ranking was a version that put the state in Fig. 2.2a, the perfectly straight string, as the most likely state. Both Abby and Bill argued for this based on the idea that forces from the water on the left and right side of the string should cancel out. This symmetry argument somewhat resembles an association between equilibrium and 'order' found in prior research [3, 20]. In the final rankings, Dewei and Bill joined Isaiah in choosing the *intuitive ranking* of $C > B > A$. In his reasoning for the switch, Bill appeared to think that the microstates associated with a more probable macrostate were themselves more probable:

Bill: *If I think these are all microstate pictures [...] then like, I don't think any microstate is necessarily— But then I guess it's more likely that it's in this macrostate, so then it's the microstates of it that are* *more likely.*

This reasoning echos our prior observation that multiple graduate students appeared to project macrostate characteristics onto constituent microstates [12].

In the interviews with graduates, four out of the eight students initially ranked the probabilities of the three snapshots from Fig. 2.2 by the intuitive ranking. Though, after considering part B of the question, many graduate students migrated to the correct answer (in total, six of the graduates settled on the correct ranking), and none settled on the intuitive ranking [12]. Three out of the ten undergraduates changed their answer, with two settling on the intuitive ranking. None of the undergraduates switched from an incorrect to a correct ranking after considering part B of the prompt.

Table 3.3: Undergraduate Responses to Part I of the Strings Question

Student	Initial Ranking (A)	Micro/Macro (B)	Ideas for Macrostates (B)	Final Ranking (C)
Abby	$P(a) > P(c) > P(b)$	Micro	net deviation	$\boldsymbol{P(a) = P(c) > P(b)}$
Bill	$P(a) > P(b) > P(c)$	Micro	height of string end	$\boldsymbol{P(c) > P(b) > P(a)}$
Carson	$P(a) > P(b) = P(c)$	*Macro*	tautness of string	$P(a) > P(b) = P(c)$
Dewei	$P(c) = P(b) = P(a)$	*Macro*	energy, pressure, mass	$\boldsymbol{P(c) > P(b) > P(a)}$
Emma	$P(c) = P(b) = P(a)$	Micro	number of turns	$P(c) = P(b) = P(a)$
Frank	$P(c) = P(b) = P(a)$	Micro	distance between ends	$P(c) = P(b) = P(a)$
Greg	$P(c) = P(b) = P(a)$	Micro	*no answer*	$P(c) = P(b) = P(a)$
Henry	$P(c) = P(b) = P(a)$	Micro	number of turns/inflection points	$P(c) = P(b) = P(a)$
Isaiah	$P(c) > P(b) > P(a)$	Micro	number of turns/changes in derivative	$P(c) > P(b) > P(a)$
Jessie	$P(c) = P(b) = P(a)$	Micro	number of turns/location derivative is zero	$P(c) = P(b) = P(a)$

Summary of undergraduate student responses. Parentheses in headers refer to prompts in Fig. 2.2. The second column, labeled ‘Micro/Macro,’ indicates whether the student identified the states from Fig. 2.2 as microstates or macrostates, and the instances of ‘macrostate’ were italicized for visibility. Bold entries in the final column indicate rankings that changed after students considered part B of the question.

Reasoning about Microstates and Macrostates

The majority of interviewed students identified the states in Fig. 2.2 as microstates. This was often accompanied by the correct reasoning that a microstate is a particular, fully-specified configuration of a system, and that a macrostate is a grouping of configurations that are in some way the same. For most students, this understanding was sufficient to generate a macrostate property for the string system. Greg, however, articulated a largely correct, abstract description of the difference between macrostates and microstates, but could not think of a macrostate property for the string:

> ***Greg:*** *A macrostate, is like, if all the configurations were the same somehow, then the multiplicity of that macrostate would be greater than one. So, if there's multiple ways of arranging it, but if the way that they're arranged are the same, then that's a macrostate. But the concept that I'm having trouble with is, this probably applies to gases, too, but like, how they can fit into a macrostate.*

Dewei, the only student who mentioned multiple macrostate properties, seemed to be reasoning by making a direct analogy to an ideal gas since earlier in the question, Dewei had invoked the ideal gas law when reasoning about the probability ranking, and talked about particles when discussing the difference between microstates and macrostates:

> ***Dewei:*** *A microstate is like a particle, or an energy of a particle. It's single part. And a macrostate is like a group of particles or the group of particles' behaviors.*

Furthermore, this reasoning seems to conflate macrostates with macroscopic objects, a difficulty also observed in the graduate interviews. When discussing whether the cartoons were microstates or macrostates, Carson also appeared to associate macrostates with macroscopic objects:

> ***Carson:*** *I would say they're macrostates only because they're something that the entire—If we define a system as the string, and [...] we're able to identify the entirety of the string and what it's experiencing—not just what is the end of the string experiencing, what is the middle of the string experiencing—instead we're just saying the entire string is this, is experiencing this force of tension.*

Carson's macrostate definition, the tautness (or tension) in the string, was also briefly considered by one of the graduate students before thinking of a more appropriate definition. In reality, the force of tension varies along the length of the string, so tension is a somewhat indefinite way to define macrostates. However, Carson's reasoning that a macrostate property involves some macroscopic aspect of the system is valid, though it seemed to take for granted that the tension in the string was uniform across the entire string, as is typically the case in introductory physics.[4]

[4] It should be noted that when reasoning about this part of the strings question, Carson claimed the three conformations

The seven remaining undergraduates (Abby, Bill, Emma, Frank, Henry, Isaiah, and Jessie) all identified an appropriate property that could define macrostates of the string (see Table 3.3). Each of these seven students, plus Carson, only gave a single macrostate definition and did not attempt to generate any additional ways of classifying macrostates. On the other hand, the graduate students we interviewed often generated multiple macrostate definitions spontaneously. As detailed in Table 3.3, the number of turns/curves/bends in the string was the most common property generated by the undergraduates, with four students giving this definition. Variations on the distance between two points on the string, given by Bill and Frank, was the next most common and one student, Abby, came up with the net horizontal deviation of the string.

Reasoning about the Entropy of the String

In total, seven undergraduates either explicitly stated that entropy was related to the number of microstates or mentioned Boltzmann's law to discuss the entropy of the string in part D. Several of these undergraduates, though, expressed some discomfort about the (seemingly) infinite space of possible conformations:

> ***Emma:*** *So, like, I don't think this is correct, but, I almost feel that defining entropy doesn't have a lot of meaning, right, because if we were to define it as our usual* $k_B \ln \Omega$ *then kind of saying like, there's infinite entropy, but—I guess, like, that doesn't sit very well with talking about how we add entropy...*

In addition to Emma, Frank and Henry shared this concern that there is infinite entropy when an infinite number of microstates correspond to a single macrostate. All three of these students were among the strongest performing students on the entire interview, so this difficulty is not linked to poor performance. None of the graduate students expressed this same concern regarding an infinite number of microstates resulting in infinite entropy. It is unclear if this is because the graduate students understood a resolution to issue of a seemingly infinite number of microstates (e.g., via a quantization argument) or did not realize there was a problem.

The other three students not mentioning Boltzmann's Law or multiplicity made associations between maximizing entropy and finding the system in "its most natural state," "most probable state," or the "chaos

were macrostates. Though, after moving on to the second part of this question (discussed in Sec. 3.4.2), Carson would realize this answer was incorrect and say the figures were microstates. Since this occurred after he finished answering the questions on this part of the questionnaire, we kept his answer of 'macrostates' as his final answer in Table 3.3.

state" (Abby, Carson, and Dewei respectively). Interestingly, Emma also worried about what the equilibrium state of this system would look like:

Emma: *There is no, like, preferred configuration then because there are no net forces and the string will probably just keep moving at random and probably never reach an equilibrium state.*

The equilibrium state of a system is a macrostate. Thus, it is possible this difficulty arises from the fact that at any moment an observer sees the string in a microstate as a result of the extra step required in conceptualizing the macrostates. Two graduates appeared to have a related difficulty in thinking that the entropy of the string was constant and did not change as the string's conformation changed [12].

In the graduate student interviews, four out of the eight students made the connection and articulated the link between entropy and multiplicity, and seven of the ten undergrads also made the connection. Other studies have seen upper-division undergraduates struggle to connect entropy with multiplicity [3, 21], so the observation here that many undergraduates successfully made this connection stands out. Furthermore, in similarity with the graduates, besides a brief comment by Isaiah and Dewei's mention of the "chaos state," the undergraduates did not commonly associate entropy of the string with "disorder."

3.4.2 Part II: Channel

When reasoning about the system with strings and circle particles in Fig. 2.3, the undergraduate students had much less overall consensus on what happened to the number of conformations of the strings when the circle particles entered the middle of the channel than the graduate students. All the graduate students realized the presence of the circles would constrain the strings in some way (and all but one graduate student explicitly stated the the conformations would decrease). Not only did the graduate students have a wide consensus on the effect, their explanations were also very similar: that the number of conformations decreased because of a space/volume exclusion effect. A plurality of the undergraduates (four out of ten) also said the the conformations would decrease, but three said the number would stay the same and another three thought the conformations would increase (see Table 3.4).

For each of the possible answers (increase, decrease, or remain the same), all the undergraduates who chose the particular answer all had very similar reasoning for their answer. So, the four students (Dewei,

Table 3.4: Undergraduate Responses to Part II of Strings Question

Student	Number of Conformations (i)	Reasoning (i)	Profile Drawing (ii)
Abby	remains the same	no new degrees of freedom	dipped
Bill	increase	circles provide new "influence" to strings	flat
Carson	increase	more unique arrangements b/c of circles	peaked
Dewei	decrease[5]	circles reduce available volume	flat
Emma	remains the same	infinite number of states in either case	dipped
Frank	decrease	circles decrease area for strings	dipped
Greg	increase	collisions result in more arrangements	flat
Henry	remains the same	same available states with and without circles	flat
Isaiah	decrease	circles occupy space which restricts strings	dipped[6]
Jessie	decrease	string reduce space available to strings	flat

Summary of undergraduate student responses to part E of the strings question. Lower case roman numerals in parentheses on title row refer to prompts in Fig. 2.3. For example sketches of the flat and dipped profiles drawn by students see Fig. 2.5. [4]Dewei initially said conformations would increase, but changed his answer. [5]see Fig. 3.2.

Frank, Isaiah, and Jessie) who said the number of conformations would decrease all, in some way, argued it was because the circles would restrict the strings or exclude some volume from the strings, limiting the conformations they could take.

The three students (Abby, Emma, and Henry) who thought the conformations would remain the same all said so because they argued that any of the possible arrangements of the strings without circles are also, in principle, still possible when the circles are present. The circles introduce a kind of 'weak' restriction on the strings; the circles don't constantly preclude any particular arrangements of the strings since any configuration of a string blocked by a circle could be available later if the circle moves far enough away. These three students seemed to think this 'weak' restriction[7] was not sufficient to decrease the number of conformations of the string. This is supported by Abby claim that the likelihoods of some configurations would change, and Emma's acknowledgement that there would be at least some restriction placed on the strings due to the circle particles:

> ***Emma:*** *...even once these larger particles [the circles] move into that region [with the strings], like, it definitely restricts some of the possible conformations.*

Though, Emma went on to say that despite the restriction, there were still an infinite number of configurations with and without the strings. She struggled with the idea that the circles took away a "small number of infinite ways to arrange the string" from a "really large number of infinite ways" which would still just result in an infinite number of arrangements.

The three students (Bill, Carson, and Greg) who claimed that the number of conformations would increase all seemed to think the presence of the circles would 'unlock' new arrangements for the strings. Carson at first simply stated that the circles would create more unique arrangements for the string (but later, when thinking about the concentration profile, returned to reasoning about mixing he used in the Partitioned Boxes question). Greg claimed that collisions with the circles would increase the conformations, and reinforced this with an argument that entropy always increases (see Sec. 3.7.6 for more on this mode of reasoning). And Bill thought the circles would add a new sort of "degree of freedom" that could influence the strings which would cause the number conformations to increase.

[7] as opposed to a 'strong' restriction that would more constantly preclude a definite set of configurations

Carson and Greg appeared to be employing reasoning analogous to conflating the total entropy of a system (in this case: the entropy of the circles plus the entropy of the strings) which does increase, with the entropy of just a part of the system (the strings) which actually decreases. Both these students subtly mentioned ideas associated with the total entropy of the system. Carson mentions mixing with the other species and how the number of conformations increases, which is true for the system as a whole. This conflation was a little more clear in Greg's reasoning, since he mentioned the second law of thermodynamics which only applies to the system as a whole. Though it is possible Bill was also making this conflation, he did not mention any ideas about the whole system as explicitly Carson and Greg. Interestingly, Abby also brought up the idea of degrees of freedom, but used it to arrive at a different conclusion.

Another, more general comparison between the graduates' and undergraduates' reasoning is in the variety among their explanations. The graduate students, despite their more varied undergraduate experiences, had much more uniformity in the way they talked about the space/volume restriction/exclusion. This may reflect the more uniformity in their answers to the question as well. Overall, the undergraduates brought up ideas like 'degrees of freedom,' space exclusion, influences/potentials, and collisions. But, in comparison with the graduate students, this variety in reasoning underlies a variety in answers.

Having students reason about the number of conformations of the strings was intended to help set them up for thinking about the long-term concentration profile of the circles. The reduction in conformations of the strings results in a decrease in entropy for the strings and an 'entropic cost' to the presence of the circles in the middle region. Balancing the increase in entropy for the circles diffusing into the middle region and the decrease in entropy of the strings causes the long-term concentration profile to have a dip in the center region.

In similarity to the responses on the number of conformations, overall the undergraduates had less consensus than the graduate students (though the difference between the graduates and undergraduates responses was less in this question on the concentration profile). A slight majority of the graduate students (five out of eight) decided on a dipped profile, some decided on a flat profile, and one picked a peaked profile.

Five out of ten of the undergraduates decided on a flat profile, and a large minority (four out of ten) decided on a dipped profile. One student, Carson, picked a peaked profile (see Fig. 3.1). Many of

Figure 3.1: Peaked concentration profile drawn by Carson.

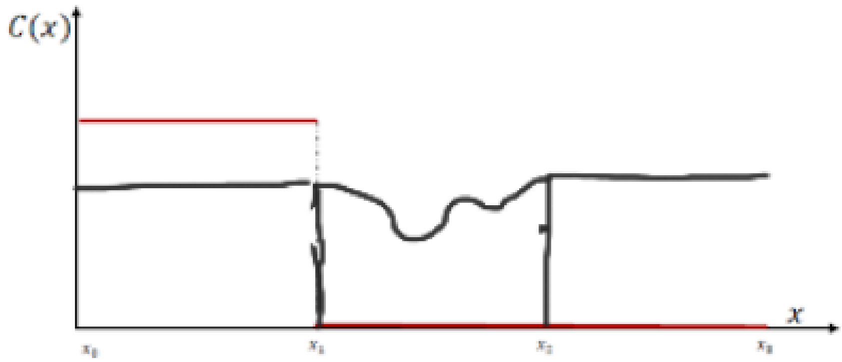
the students who choose a flat profile seemed to be defaulting to the 'free' case, in which the circles diffuse without the presence of the strings. Henry drew an appropriate analogy to the heat equation (which also describes diffusion). Henry did not think the strings created any sort of "potential" that would effect the concentration of the circles:

Henry: *Well, the thing I immediately thought of was, um, the heat equation... The particles will behave very similarly to heat... Assuming that middle region doesn't, I don't know, try to force the molecules out, essentially having a higher potential there, a higher energy potential, then the molecules should be evenly distributed in the channel after a while.*

In Greg's case, he drew a direct analogy with the Partitioned Box question in which the circle species spread out uniformly throughout the system regardless of the presence of the other species. Abby also made a connection to the Partitioned Box question, though it was in support of a dipped profile, her reasoning on both questions was consistent. This provides an interesting point of comparison with the graduate students who did not make explicit connections between these two questions.

Of the four students who decided on a dipped profile, two (Emma and Frank) initially were leaning towards a flat profile. Both changed their answer to dipped after considering the size-scale of the circles or strings and whether they took up a non-zero area. Isaiah also considered the scale of the particles and decided on a dipped profile (see Fig. 3.2), though this profile was unique and is discussed in more depth later. Explicitly considering the scale of the species in the problem was not, universally a key to picking a dipped profile. Abby picked a dipped profile, reasoning that the amount of 'stuff' in each compartment

Figure 3.2: Final state concentration profile drawn by Isaiah.



would be constant. She claimed that the total density of matter in each section of the channel would be the same which, because the strings were present in the middle section, would result in fewer circles in the middle region. And Carson did explicitly consider the sizes of the particles, but chose a peaked profile.

The two most unique profiles were drawn by Carson (Fig. 3.1) and Isaiah (Fig. 3.2). Though Carson drew a similar profile as one of the graduate students, Chris (see Fig. 2.6), their reasoning was different. Chris thought the strings would physically trap the circle particles in the center region, but for Carson the peak was because of an idea that mixing always increases entropy:

Carson: *I know that whenever you mix two species together you're going to be increasing entropy...*

This reasoning is consistent with his response to the Partitioned Box question in which the 'amount of mixing' was his primary mode of thinking about the system.

Isaiah's profile was the most unique among all (both graduate and undergraduate) interviews. His profile was dipped, but had four inflection points (as opposed to the typical two) which resulted in a profile with more wiggles: one peak and two valleys (see Fig. 3.2). It was unclear whether the wiggles represented random fluctuations in concentration that could naturally occur (It is reasonable to think that the strings could introduce temporary, local fluctuations to the concentration profile that would be larger than the fluctuations in the regions without the strings); or whether Isaiah might be thinking about the concentration profile as the solution to some diffusive differential equation, like Henry. The latter is perhaps more likely since Isaiah explicitly considered the boundary conditions at x_1 and x_2. In this case, the wiggles might represent a superposition of sinusoidal basis functions which is the general solution to the relevant differential equation.

3.5 Results: Expansions

As noted before (in Sec. 2.2 and Sec. 2.7), this question about expanding gases was developed in a prior study [4]. In the study by Bucy *et al.*, this question was given twice: once as a pre-test, and once as a post-test, at the beginning and end of a thermal physics course. When analyzing the undergraduate interviews in our study, we found differences and similarities with the prior research (both to Bucy's study and the graduate responses).

Unlike the prior studies, the undergraduates we interviewed appeared to struggle more with determining the sign of the isothermal expansion than to determine the sign of the entropy change in the free expansion. As shown in Table 3.5, the undergraduates were more likely to change their answer for the sign change in the isothermal expansion. Though most students (nine out of ten) answered correctly initially, only six students kept their correct answer. Bill's initial intuition was that the change was positive, but could not think of compelling reasoning to justify his intuition. He expressed much uncertainty, but eventually reasoned that since the gas was doing work, the number of microstates would decrease causing the entropy to decrease.

Table 3.5: Undergraduate Student Responses to the Expansions Question.

Student	Sign ΔS_{iso}	Sign ΔS_{FE}	Entropy Change Comparison	Claimed $\Delta T_{\text{FE}} < 0$
Abby	- $\rightarrow$ +	-	$\Delta S_{\text{iso}} > \Delta S_{\text{FE}}$	yes
Bill	+ $\rightarrow$ -	+	*no answer*	no
Carson	+	+	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	no
Dewei	+ $\rightarrow$ 0	+	$\Delta S_{\text{iso}} < \Delta S_{\text{FE}}$	yes
Emma	+	+	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	no
Frank	+ $\rightarrow$ 0 $\rightarrow$ +	+	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	no
Greg	+ $\rightarrow$ 0[8]	+	$\Delta S_{\text{iso}} > \Delta S_{\text{FE}}$	no
Henry	+	+	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	no
Isaiah	+	+	$\Delta S_{\text{iso}} < \Delta S_{\text{FE}}$	no[9]
Jessie	+	+	$\Delta S_{\text{iso}} = \Delta S_{\text{FE}}$	yes

Summary of the undergraduates' responses to the expansions question. In this table, the quantities ΔS_{iso} and ΔS_{FE} are the magnitudes of the changes in entropy of the gases undergoing the isothermal and free expansions (respectively). ΔT_{FE} is the change in temperature of the gas undergoing the free expansion. A "yes" in this column indicates the student considered this at any point while reasoning about the task, regardless of their final conclusion. [8]Greg changed his answer for ΔS_{iso} to zero after comparing the entropy changes of the gases but did not revisit his answer to the entropy comparison. [9]Isaiah claimed the temperature of the freely expanding *increased*.

Two students (Greg and Dewei) changed their answer from positive to zero after realizing the process was reversible, meaning the total change in entropy was zero, but conflated the total change in entropy with the change in entropy of just the gas (see Sec. 3.8.1 for a deeper discussion of this difficulty in distinguishing system, environments, and the Universe). For Greg, this conclusion was supported by an intuition about heat engines:

> ***Greg:*** *S equals zero, and it's similar to an engine... the entropy of the universe increases, but not the actual system, so the gas's entropy doesn't really change.*

Greg seemed be thinking about a cyclic process in which the entropy of the working substance returns to its original value, while the entropy of the surroundings has increased. Frank also nearly made a similar error (assuming the entropy change of the working substance in a reversible process was zero), but realized his mistake and that the heat flow had to increase the entropy of the gas. Uniquely, Abby was able to correct her initial answer by reasoning about the number of microstates available to the system, and this case is discussed in Sec. 3.7.1.

Similarly to the students in Bucy's study (and dissimilarly to the graduate students), many of the undergraduates activated the Clausius relationship between heat flow and entropy: $\Delta S = Q/T$. Five students (Emma, Frank, Greg, Henry, and Isaiah) reasoned productively with this equation, and two more (Carson and Bill) activated the idea but struggled to use it correctly. This difference between undergraduate and graduate students may be due to the undergraduates' closer proximity to thermal physics concepts at the time of the interviews, a hypothesis mentioned in Sec. 3.3. Though many of the undergraduates used the appropriate macroscopic/thermodynamic idea (i.e. the Clausius relationship), another similarity between all three populations of students was a preference for a more statistical definition of entropy.

On the whole, the undergraduates we interviewed had little trouble determining the sign of the change in entropy of the free expansion (especially compared to the graduate students). They used a variety of arguments, from more incomplete reasoning that 'entropy always increases,' to more appropriate claims on the changes to the volume or number of microstates, to explicit recognition that a free expansion is an irreversible process. The graduates were more likely to initially state that there was no change in entropy in this process, but several students demonstrated an ability to correct this mistake. This observation of the undergraduates' initial responses being correct, and the graduates correcting their responses matches a pattern seen in the strings question (see Sec. 3.4.1).

When ranking the changes in entropy of the two processes, the undergraduates also outperformed the graduate students. Five out of the ten undergraduates correctly reasoned that the two entropy changes were equal, as opposed to just two out of eight graduate students. This reflects the fact that more undergraduates recognized the symmetry between the two process. Three students (Frank, Henry, and Jessie), explicitly recognized that entropy was a state function and the other two students (Carson and Emma) noticed the two gases were in the same final state and had the same changes in volume but did not explicitly mention that entropy was a state function (i.e. changes in entropy only depend on the initial and final state of the gas and is not process dependent). There were few patterns among the students saying the entropy changes were unequal, and typically expressed explicit uncertainty in their answer.

In both Bucy's study and the interviews with graduates, students were observed claiming that the temperature of the freely expanding gas would decrease. Students making this claim either had a strong intuition that expanding gases cooled or assumed that the freely expanding gas did work during the expansion (lowering its internal energy). We noticed far fewer undergraduates making this claim in these interviews. Many undergraduates, appropriately, did not concern themselves too much with the temperature of the freely expanding gas. However, two of the students (Abby and Isaiah) giving unequal comparisons came to opposite conclusions about the temperature change in the free expansion, which lead to opposite answers to the comparison question.

Abby thought the temperature decreased, and noted that an increase in volume and a decrease in temperature are changes that would tend to increase and decrease the entropy, respectively. Thus, according to Abby, these counter opposing effects would mean the free expansion had a smaller magnitude of change than the isothermal expansion where the volume increased (increasing entropy), but in which temperature remained constant. Isaiah had a similar line of reasoning, but began with the claim that the temperature of the freely expanding gas increased. Then since increases in temperature and volume both increase entropy, the free expansion would have a larger change in entropy than the isothermal expansion in which the temperature was constant, but in which volume increased.

3.6 Discussion: Undergraduates' Conceptualizations of Entropy

The final question on the interview asked students about their conceptualizations of entropy. As was the case with the graduate interviews, this was the place associations between entropy and disorder were most common among the undergraduates. And among the undergraduates, mentions of disorder were more common than among the graduate students: seven of the ten undergraduates mentioned either disorder or order.

Another common conceptualization, also given by seven students, was the association with microstates, the possible ways to arrange a system, or Boltzmann's Law: $S = k \ln \Omega$. As this was a common conceptualization throughout the undergrad and graduate interviews, it will be discussed in more depth in Sec. 3.7.1.

Four of the undergraduates mentioned thinking about entropy as a 'force' or a 'driver' of processes. Isaiah talked about thinking of entropy as a "propellant:"

> ***Isaiah:*** *I kind of conceptualize entropy as a propellant. It acts as a potential for heat flow. When I think about, like, heat transfer between two solids, what I think happens there is that the entropy is kind of what is propelling the heat to be transferred.*

This conceptualization will be discussed in a little more depth in Sec. 3.7.6.

3.7 Common Resources Used by Students

In the previous sections discussing students' responses to the interview questions, we focused our analysis narrowly on individual questions, independently of the other tasks on the interview questionnaire. In this section, we will take a step back to look more broadly at themes that emerge when considering the interviews among both the graduate and undergraduate students as a whole. In examining common elements of reasoning employed by students across questions, we will identify student ideas that can be categorized as conceptual resources, as introduced by Hammer [19].

According to Hammer, a conceptual resource is analogous to a block of computer code that performs a specific task, and is often pasted into new code without any thought given to the inner workings of the block. In parallel to this computer science metaphor, empirically, students use conceptual resources as tools, often times without justification for why the tool is appropriate or correct. Consequently, a particular resource may be correct or appropriate in one context, but incorrect or inappropriate in a different context. From an instructional perspective, an understanding of the resources students use to think about the concepts used and taught in classrooms helps an instructor better relate and communicate with students.

3.7.1 Entropy is Related to the Number of (Micro)states

In both the graduate and undergraduate interviews, at least one student employed the reasoning that entropy is related to a number of states (or multiplicity) in every single question, even in the Blocks and Expansions questions where there was neither a cue nor a need to think about microscopic quantities. In his study of undergraduate student reasoning about entropy and equilibrium, Loverude identified a resource he labeled as “entropy is related to multiplicity” [3]. Though we consider the resource discussed in this section to be identical to the one identified by Loverude, we chose to refer to this reasoning (that entropy is related to a number of states) with language that reflected the phrasing used by the students we interviewed. Students, especially the graduate students, talked about a number of states or a number of microstates more frequently than they mentioned the word ‘multiplicity.’

Additionally, Bucy noticed that students have a preference for the microscopic/statistical definition of entropy over the macroscopic/thermodynamic definition [4] which is consistent with students in our interview thinking about microstates in the Blocks and Expansions questions. One possible reason for this affinity stems from the statistical definition of entropy

$$S = -k_B \sum_i p_i \ln p_i \tag{3.1}$$

which provides a more *a priori* conceptualization of entropy without relationship to other quantities. Incidentally, the graduate students in this study seemed to invoke the general form above in Eq. 3.1 about as often as the more standard form: $S = k_B \ln W$. Furthermore, this definition holds true more generally than the thermodynamic definition of entropy, $\Delta S = Q/T$, which does not apply to processes in which entropy changes, but the heat flow is zero.

Many students used this association between entropy and multiplicity appropriately and productively.

In the Partitioned Box question, Daana invoked Eq. 3.1 and Garth (both graduate students) thought about the number of ways to split up the particles with combinations and binomial coefficients to reason about the entropy of each of the four states. In the expansions question these same two students utilized reasoning about microstates and configurations, respectively, to decide that the entropy of the gas expanding freely increased despite the heat flow being zero.

Half the undergraduate students (five out of ten) brought up microstates when reasoning about the Partitioned Box question. If references to arrangements, distributions, or spreading out are included, the total goes up to seven out of ten. As noted in Sec. 3.2, for both Isaiah and Carson, this reasoning about microstates was tied to a resource about mixing which led them to (different) incorrect responses.

The undergraduates used this microstates resource more effectively than the graduate students on the Expansions question. Though not always with rigorous correctness, many students recognized that the increase in volume increases the number of microstates which result in an increase in entropy. For one undergraduate, Abby, activating this resource helped her correct her answer to the question asking about entropy change of the isothermal expansion:

> ***Abby:*** *Increasing the volume would increase the microstates because it's like... increasing the number of places that a particle could be at any given time. But, like, that isn't, like, a line of logic we ever followed in class, and would sort of lead to the conclusion that the change in entropy is positive.*

In the Strings question, half of both the graduate and undergraduate students made a substantive connection between the configurations of a string and the entropy of the string. Harry (a graduate student) even made a deeper connection between the entropy of the string and the density of states:

> ***Harry:*** *So there's only one microstate that corresponds to this string reaching as far as it can. And then this one, there's a lot more, um, a much higher density of microstates that lead to the string reaching that far because it can be like this, or this....*

Emma, an undergraduate, made a similarly profound connection about the entropic interaction between the strings and the circles in the second part of the strings question, though didn't quite fully tie it all together:

> ***Emma:*** *We are trying to maximize the entropy of the [circle] particles, you know, like probably some balance between them having access to a greater number of microstates by expanding through the volumne,*

but also through their interactions with the strings. And yeah, I don't know exactly how to visualize that.

And as mentioned in Sec. 2.5, Chris (a graduate student) effectively used microstates to reason about the changes in entropy of the two blocks in the Blocks question:

Chris: *I suppose the entropy of the higher temperature one will decrease because it has to give up some of its energy to solid B and the entropy of solid B will increase because it's sort of received some energy has more microstates allowed because there's more momentum possible in the little particles. But it will increase more than the decrease in entropy of solid A, so that the total entropy is... at the end is bigger than at the beginning.*

However, Fred and Beth (graduate students) also brought up this resource about microstates in the Blocks question, but it did not help them reason productively. Fred said the entropy of both blocks would increase. Beth thought the entropy of the two block system would remain constant because she could not see how the total number of states would change. A tendency for students to treat entropy as a conserved quantity has been noted in other studies [6, 15, 24]. In this case with Beth, it seemed like the idea of constant entropy was due to a sense that the total space of states possible to the system (independent of energy constraints) remains static, instead of considering entropy changes as a reflection of changes in the number of *accessible* microstates as the macrostate of the system changes. Her reasoning also dovetailed with the relationship between entropy and temperature which will be discussed in the next section (see 3.7.2).

Beth: *I remember S being the sum of states. I think there was a log in there somewhere, but whatever... Yeah, I think that was log of the microstates. Um, yeah, so like naively, the way we think about it in condensed matter is the number of states is related to the temperature because the more things can move the more states they have, so higher temperature would be more possible states and lower temperature means less... states.*

In the interviews with undergraduates, a similarly-sized minority (as the graduates) brought up the idea of microstates on the blocks question. For one student, Carson, this microstate resource was particularly strong. As mentioned in Sec. 3.3, he made a connection to the states of the hydrogen atom system:

Carson: *Because there is going to be more microstates readily available... And just like the basic hydrogen atom, you have the energy levels... So if solid A is at a high temperature and is losing some heat and losing internal energy, you know, it's dropping down those energies, but as solid B is getting internal energy there's way more microstates that it's going to be able to jump to....*

This reasoning from Carson helped him understand how the increase in entropy of the cold block overcame the decrease in entropy of the hot block so that the total entropy of the system increased.

3.7.2 Entropy is Related to Temperature

The second conceptual resource we observed students use frequently was an association between entropy and temperature, as if a monotonic relationship existed between the two quantities. This association did not appear in the Partitioned Box question in either the undergraduate or graduate interviews. This is not surprising since temperature was not a salient feature of the task. However, in the graduate interviews it did appear in the three other physics-content tasks and frequently in the undergraduate interviews on the Blocks and Expansions questions. Since such an association can be problematic in some circumstances, we have included an Appendix (see App. B) with a deeper discussion and counterexample showing a case were this association can fail, which might serve as a tool to help instructors who may want to address this association with students.

In the graduate interviews, we noticed two primary use cases for this resource. In the first, graduate students used temperature to directly compare the entropy of two different objects, as Fred does in this quote from the Expansions task:

Fred: *I would think that the one [gas] with the higher kinetic energy has more entropy, um, just cuz of my, like elementary understanding entropy as being like chaos. I think of the rapidly moving molecules as having more entropy.*

On the surface, such reasoning is valid if temperature is the *only* difference between two objects. However, directly associating an object's entropy and temperature lead some students astray as it did for Beth in the Expansions task:

Beth: *Also, vaguely, entropy has something to do with temperature. I don't remember what, there was* *a formula I learned at some point, but if temperature is not changing then, naively, the entropy shouldn't change either. Maybe.*

This quote from Beth also demonstrates the second way in which students used this resource: to think about how changes in temperature affect the entropy of a single object.

As already discussed in Sec. 2.5, students also used this resource heavily in the Blocks question. But somewhat surprisingly, this resource also briefly appeared in the Strings question. Erik brought up this resource when thinking about the changes in entropy of the circles and strings as the circles diffused through the channel. He seemed to wonder whether the circles being at some temperature might impart some energy to the strings and increase the entropy of the strings:

Interviewer: *So, does that [having a circle in the strings region] do anything to the entropy of the strings?*

Erik: *The entropy of the strings itself? Okay, let me think about it this way... If, if I had a very hot gas here and nothing in here, and the strings were here, I let this loose, the temperature of this, the strings, must come up. So, I'm guessing the entropy goes up?? But the entropy is not related to the temperature directly...*

After a brief discussion, Erik did conclude that the entropy of the strings would decrease due to the decrease in number of conformations. Two other students, Harry and Chris, also worried about how the temperature of the water bath or circles might effect the strings, though more tangentially.

The undergraduates also frequently used an association between temperature and entropy to reason about the Blocks and Expansions questions. Sometimes, as in the Blocks question, this reasoning would be productive in leading students to the correct answer. In the Expansions question, this association lead some students astray, but only because they made a previous error. For example, Abby reasoned that entropy would decrease in the free expansion, but the initial error was thinking that the temperature of the gas decreased in the expansion:

Abby: *The temperature is going to be going down, so there's less sort of excited states that the particles can be at. So that would decrease entropy.*

Furthermore, Isaiah reasoned that the change in entropy of the free expansion was greater than the change in entropy of the isothermal expansion after incorrectly reasoning that the temperature of the free expansion increased. In his thinking, Isaiah implicitly assumed the pressure of the free expansion remained constant, so according to the ideal gas law ($PV = NkT$) the temperature has to increase if the volume increases. Then, using the Sakur-Tetrode equation, reasoned that the increase in volume and temperature in the free expansion meant it had a greater change in entropy than the isothermal expansion which just had an increase in volume.

3.7.3 Entropy is Related to Mixing

Appeals to mixing came up with relatively similar frequency in both the graduate and undergraduate interviews, typically on the Partitioned Box and Strings questions. Mixing two substances is one of the primary examples of an irreversible process in which entropy increases. A typical example of mixing occurs in a system of two different gases initially separated from each other in two halves of a sealed container. When the gases are allowed to diffuse, they eventually mix and both gases will double the volume they occupy with no change in pressure or temperature, which increases the total entropy of the system.

However, the increase in entropy for this process of two gases mixing is identical to the total change in entropy of having the same two initial gases undergo free expansions in separate containers. In other words, if the change in entropy in the mixing process is ΔS_{mix}, and the change in entropy of the free expansion are ΔS_1 and ΔS_2, then $\Delta S_{\text{mix}} = \Delta S_1 + \Delta S_2$. There is no additional entropy created by a ‘mixing’ interaction (unless there is some type of energetic interaction like an attraction or repulsion between the separate species). Thus, it is best to think about ‘mixing’ as an *ad hoc* description of a process in which entropy increases due to the mutual, simultaneous increases in volume of multiple pure substances.

Many undergraduates, however, seemed to think that there actually was an extra, unique ‘entropy of mixing’ that occurred when two species mixed. This reasoning lead Carson to his peaked concentration profile in the second half of the Strings question. Additionally, this reasoning seemed to appear in Emma’s reasoning too:

Emma: *I guess in a way, like, by having them interact... like the interactions, so like, the collisions* *themselves would probably create entropy,*

though she would later go on to draw a dipped concentration profile. In the Partitioned Box question, both Isaiah and Carson brought up the idea of mixing (see Sec 3.2) and choose state D, the state with the most circle particles mixed in with the square particles (see Fig. 2.1d), as the state with maximum entropy which is consistent with treating mixing as a process that generates extra entropy.

Furthermore, there was little consistency between students who invoked the 'mixing' resource. In the graduate interviews both Alex and Fred claimed the most mixed state would have the highest entropy but choose different rankings. A third student, Daana, also mentioned mixing and initially gave a third ranking distinct from either Alex or Fred's, but went on to determine the correct ranking after realizing the partial pressures of each species will equilibrate (see Sec. 2.4). Similarly in the undergraduate interviews, Carson and Isaiah both invoked mixing as their metric for ranking the states, but choose different rankings (see Sec. 3.2). Though Carson and Isaiah both agreed on the states with the most and least entropy, the graduate students invoking the mixing argument chose states A and state B as the states with the most entropy.

3.7.4 Entropy is Related to Disorder

One of the interesting observations from the graduate interviews was that the graduate students did not commonly mention that entropy was related to disorder, though it did come up in a few situations. For example, when asked to discuss what might be meant by the entropy of a string Beth initially stated:

Beth: *So you expect to find it in a higher entropy state where things are disordered.*

Though disorder appeared to be Beth's default conceptualization of entropy and likely influenced her unequal initial ranking of the string microstates, she did eventually realize all three configurations were microstates, and thus were equally likely to occur. Daana was another graduate student who brought up the idea of disorder, but mentioned that it was not a complete way of talking about entropy

Daana: *It is some measure of disorder, but that also is not, it's not complete. Yeah, it's a measure of disorder and also the number of states a system can assume.*

Graduate students would also bring up other terms which did not appear consistently enough to, on their own, be particularly interesting, but have the same sort of vague ambiguity as ‘disorder’. In the Partitioned Box question, many graduate students described entropy as something related to the “distribution” of particles. In the Strings question, students seemed to favor terms like “chaotic,” “squiggly,” and “wavy.”

Perhaps unsurprisingly, references to disorder were more common among the undergraduate students but the undergraduates would almost always state that disorder wasn’t the best way to talk about entropy. For example, when reasoning about the Partitioned Box question, Jessie claimed:

Jessie: *[State] C is the least entropic, as there’s, I don’t think any way I could make it more organized, or, you know. I guess that’s not, like, the best way to view it.*

Furthermore, Abby made a connection between the order in the system, and the temperature

Abby: *cooling thing, like, increased the sort of order in them,*

which relates to the resource discussed in Sec. 3.7.2. Similar to the graduate students, though, this was not the most common heuristic undergraduates had for entropy. Thinking of entropy as a number of microstates (or multiplicity) was far more common among the students we interviewed.

3.7.5 Work is the Opposite of Entropy

This resource was more prevalent among the undergraduates, but it also appeared in the graduate student interviews. This was Abby’s primary conceptualization of entropy

Abby: *I think of entropy as, like, the amount of work, like sort of the antithesis of work... Like maximizing entropy is something that, sort of, like, leads things towards equilibrium because it takes, like, less work to be in that way.*

Bill used a similar idea when reasoning about the Blocks question:

Bill: *Entropy should increase in this system as time progresses, because there’s nowhere for the, like, work done needed to reverse the entropy to come from.*

Bill also brought up this idea in the Expansions Question to conclude that the gas undergoing the isothermal expansion has a decrease in entropy because it is doing work. He ignored entropy change caused by the heat flow into the system (needed to keep the temperature constant) and reasoned that the work done by the system would decrease the entropy of the system.

Also in the Blocks question, Carson brought up the idea of 'intentionality' when talking about spontaneous processes; that some things happen naturally without "us trying to do anything intentionally here." This appears at least somewhat related the idea of work, or more generally, effort. This matches Bill's reasoning in that an increase in entropy is an 'effortless' thing, but a decrease in entropy is an 'effortful' thing.

This idea came up in at least one of the graduate interviews. When reasoning about the expansions question, Fred claimed that the work done by the gas undergoing the isothermal expansion would likely decrease the entropy of the gas. In this case, Fred was considering how energy was leaving the system as work and was relating the expenditure of energy to a decrease in entropy. It seemed, though, that Fred was making an oversimplification in his reasoning by relating a decrease in energy with a decrease in "chaos" (one of Fred's preferred conceptualizations of entropy) which would mean a decrease in entropy, which contrasts with the undergraduates who made more explicit connections between work and entropy.

On the whole, this conceptualization of work as the opposite of entropy seems like a slightly-more-sophisticated corollary of the 'entropy is related to disorder' resource. It relates to the 'messy room' image oftentimes associated with entropy, especially in the analogy that it takes 'work' to clean the room and bring it to a more ordered (less 'entropic') state.

3.7.6 Other Resources

The following two resources are honorable mentions in comparison to the resources discussed above as there is relatively less to say about them. The first, 'entropy always increases,' is almost as ubiquitous as the association between entropy and disorder, and so the pitfalls in over applying this concept is also likely to be well-known among practitioners. It is also closely related to the difficulty in distinguishing systems, environments, and the Universe, which is discussed in more detail in Sec. 3.8.1. The second is an association between entropy and an abstract concept of a force or driver of processes.

Entropy (of a System) Always Increases

The statement that the entropy of a system always increases is essentially the second law of thermodynamics. Therefore, this is a largely correct resource for students to use. Though, as one might expect, there is a risk of over applying this concept to reach the conclusion that entropy of every thing always increases. It is more appropriate to say 'entropy tends to increase,' but this restatement alone likely is not enough to steer students away from this pitfall. This error in over-applying this concept is generally connected to students' neglecting to fully distinguish systems, surroundings, and the universe.

We saw students committing this over-application in a few places throughout the interviews, but not to a particularly worrisome degree. One example was Carson reasoning through the change in entropy of the environment outside the freely expanding gas:

Carson: *Also, I intuitively want to say it's positive just because* ΔS *of the universe is always increasing.*

Students also often used the language that 'entropy is maximized at equilibrium,' but it may be more appropriate to classify this as a separate resource.

Entropy as a Driver (or Force)

In his book, Daniel Schroeder states "the net increase in entropy is the driving force behind the flow of heat" in the section: The Macroscopic View of Entropy. In this section, Schroeder speaks very loosely about how sometimes he thinks of entropy as a fluid, but a fluid that can only be created and not destroyed. It provides great context and good handle to help explain what entropy is (and see Sec. 3.8.2 for a less gentle, but complementary explanation of entropy).

This discussion seemed to be impactful to students since many undergraduates made statements about how entropy was the driver of processes, particularly when asked about their conceptualizations of entropy. Four undergraduates mentioned that they conceptualized entropy was a force or driver of processes when directly asked in the fifth and final question of the interview. Two undergraduates brought up this idea in the Partitioned Box and Blocks questions as well. Greg brought it up the most frequently:

Greg: *Entropy is the driver of temperature... The real driver here is entropy.*

Graduate students brought it up less frequently, but it still appeared in a few places throughout their interviews. Erik mentioned how entropy "drives all changes in temperature" in the Blocks question, and in the Strings question Harry brought up the idea that the space exclusion effect from the strings creates a "force tending to push the balls [circle particles] out of the region between x_1 and x_2."

3.8 Implications for Instructors

It is important to stress that that resources discussed in the previous section are not necessarily misconceptions. More precisely, they are conceptual models for understanding and reasoning about physical concepts. Many of the resources used by students were productive at leading students to appropriate conclusion whether or not the resource was precisely applicable to the particular situation. As with any model, it is important to understand its strengths and limitations, i.e., under what conditions it is or is not appropriate to use. For instructors, an understanding of what avenues of productive reasoning (or troublesome misconceptions) a resource may encourage will be valuable for effectively communicating in ways which resonate with students' understanding and help them develop productive reasoning skills.

However, the implications we discuss in the following sections should be taken with some limitations in mind. First, the majority of our interviews took place among students at CU Boulder. All but two of the interviewed undergraduates took the same thermal physics course at CU Boulder with the same instructor. And though the CU Boulder graduate students did represent a larger variety of undergraduate institutions, they are still subject to a selection effect from the graduate admissions process. This results in a somewhat limited perspective on the range of student reasoning, so the resources employed by other groups of students may differ from the ones we observed and identified. For example, a prior study found that among students at California State University Fullerton (CSUF), the connection between entropy and multiplicity was not strong, whereas we noticed students making a strong connection between entropy and multiplicity [3]. In other ways, however, there were parallels between the studies at CSUF and CU Boulder such as students struggling to relate changes in temperature to changes in entropy (see Sec. 3.3).

Despite these limitations, we do discuss some recommendations for instructors who may want to improve their thermal physics courses. While writing these recommendations I oftentimes remembered a statement made by the Head of the Central Bureaucracy from Futurama: "you are technically correct, the best kind of correct." It is often cumbersome and tedious to be technically correct, and sometimes the extra steps needed to make a statement more technically correct can introduce more confusion than it's worth by creating more places to stumble. Though, when it comes to entropy, which is often talked about in a more vague, hand-wavy sort of way, perhaps a little more technical precision could benefit the discussion of the topic. The following sections are a bit pedantic, and at times go out of the way to be 'the best kind of correct,' which may be grating to some, especially those with extensive knowledge of thermal physics. The intention, however, is for the precision to be helpful to those who are still grappling with the concept of entropy earlier in their education.

3.8.1 Systems, Environments, and the Universe

Across many studies on student reasoning in thermal physics, students have been observed struggling to differentiate the entropies of systems, environments, and the Universe/universe[10] . Or, in a situation where a system is comprised of multiple parts, to differentiate the entropy of the parts from the entropy of the whole [4, 15, 23]. Typically in situation where this distinction is important, there is a system with S_{sys}, and an environment with S_{env}, which together compose a universe which has $S_{\text{unv}} = S_{\text{sys}} + S_{\text{env}}$. The second law of thermodynamics states that the entropy of the universe, S_{unv} tends to increase. The second law says nothing about what the entropy of the system and environment will do individually beyond that the sum of the two will (tend to) increase (in a system of a large enough size). Entropy of either the system or the environment may decrease, but in this case the increase in entropy of the other object will typically be larger than the decrease in the first.

Thinking of a system, environment, and the Universe/a universe is the typical setup when thinking about heat engines (but there is an added twist that there may be more than one 'environment' like in situations with both hot and cold reservoirs). Some of students' struggles may stem from trying to relate this situation to situations similar to the blocks question. In principle, the blocks are surrounded by an environment and are only a *part* of the Universe. But since the blocks were isolated from the rest of the

[10] When capitalized Universe refers to all of existence (i.e., the Universe), when in lowercase it refers to an isolated, self-contained system (e.g., a universe)

Universe, the two blocks can be treated as their **own** universe. So, when trying to map this situation to the $S_{\text{unv}} = S_{\text{sys}} + S_{\text{env}}$ equation, it is tempting to still think of S_{unv} as the entropy of everything, i.e., both the blocks and the surrounding environment. However, since we can treat the blocks as their own universe, it is more appropriate to treat the entropy of the first block (block A) as $S_A \equiv S_{\text{sys}}$ and the entropy of the second block (block B) as $S_B \equiv S_{\text{env}}$, or vice versa. This makes $S_{\text{unv}} = S_A + S_B$, with the second law stating that S_{unv} tends to increase which will permit a decrease in either S_A or S_B, as long as the other increases sufficiently.

In some situations, we noticed students writing things like $S_{\text{tot}} = S_A + S_B$ in the blocks question, and then considering how S_{tot} might be connected to S_{unv} (e.g. treating S_{tot} as S_{sys}). Therefore, we recommend instructors have more intentional discussions on how to identify systems and environments, and to which quantities the second law of thermodynamics applies. Perhaps the term 'Universe/universe,' with its common connotation as all of existence, could be replaced with something more neutral like 'total.'

3.8.2 What exactly IS Entropy?

Entropy is a physical quantity just like temperature, energy, or pressure. Like all other physical quantities it has units. In SI system, entropy is measured in units of Joules per Kelvin[11] . But unlike the other quantities, it is much harder to boil down to a simple, easily understood aphorism. For example, temperature can be reasonably oversimplified to an object's willingness to give up energy; pressure is a system's way of 'pushing' on another over some surface. The best oversimplification of entropy is perhaps 'the thing maximized when a system reaches equilibrium,' though this definition may raise more questions than it answers.

As stated in the introduction, entropy does have a concise mathematical expression that generalizes very well to almost (if not all) possible systems, the Boltzmann Equation:

$$S = -k_B \ln W. \tag{3.2}$$

In systems where the energy of each of the microstates are not equal, Eq. (3.2) becomes a sum over proba-

[11] Which, for extra wildness, are the same units as a heat capacity.

bilities of each microstate (see Eq. (3.1)). And in systems with a continuous distribution of microstates (like the string system), the sum in Eq. (3.1) becomes an integral over some volume of configuration space.

Another aspect about the difference between entropy and other physical quantities, like temperature or pressure, is that you can think about the temperature or the pressure of the system at a particular moment in time without worrying about other ways the system could be at other times. Entropy, on the other hand, basically **is** a relationship between the system in its current state and all other possible arrangements of the system. To think about entropy, a means of classifying macrostates of the system must be defined, then the entropy is the Boltzmann constant times the natural log of the number of microstates which manifest the macrostate. So, when talking about the entropy of a system, it is most appropriate to talk about the entropy of the *macrostate* in which the system is in.

We saw students grappling with this subtlety of entropy in the interviews. In particular there were instances where a student would say that the probability of a particular arrangement (i.e. a microstate) would be more probable because it belongs to a macrostate that is more likely (see Sec. 2.6.1 and Sec. 3.4.1). By the fundamental assumption of statistical mechanics, all the microstates of the string are equally likely, so this logic used by some students was incorrect. However, it begs the question: Can we talk about the entropy of a *microstate*? If so, it is a bit awkward to discuss. Since a microstate belongs to a macrostate, it may be reasonable to attach the entropy of the macrostate to the microstate *ad hoc*. Again, since at any given moment a system is in a particular microstate, we may want to talk about the entropy of the system at any given moment. But a clearer answer is: at any given moment the system is also in a macrostate (and it doesn't really matter what the microstate actually is) and the *macrostate* has an entropy associated with it. This is one of the more useful things about entropy: dealing with the indeterminacy of the particular microstate.

Another nuance of entropy is that, most technically, entropy is **not** something that is exchanged between objects. Entropy *exists* only *in* a particular object. An object can exchange energy, volume, or matter, with other objects, but it does not, in the most technically rigorous sense, transfer entropy. The same is true for temperature. Objects don't exchange temperature, they exchange energy and temperatures change according to the amount of energy transferred. If something decreases its temperature, that lost temperature doesn't go anywhere.

As in the blocks question, for a specific example, as the hot block gives up energy to the cold block, the entropy (or temperature) it is losing does not 'appear in' or 'transfer to' the cold block. The energy is what transfers. The entropy of the hot block doesn't *go* anywhere, it just decreases. Similarly, the entropy of the cold block increases because it is gaining energy but that entropy didn't come from anywhere, it just increased. If one does want to talk about it as the creation or annihilation of entropy within particular objects, entropy 'appears' from and 'disappears' into non-existance in a very non-Parmenidean[12] sort of way.

One, less technical, way I have come to think about entropy is as the Universe's way of doing a cost/benefit analysis, similar to the way Daniel Schroeder describes entropy in his 'Silly Analogy' [1]. If you think of a business, in principle, it is trying to maximize profit, but there are many trade offs it has to consider. For example, will cutting prices to sell more product increase profit, or will raising prices increase profit? Will cutting workers' wages save money (even at the risk of losing productivity), or will increasing workers' pay increase productivity and thus profits?

As with entropy, a system (i.e., an isolated system composed of separate parts) may have many options for changing its entropy. Different parts of the system can exchange volume, energy, or particles/stuff (similar to how a business could increase/decrease prices or pay). Unlike a business, however, which may make poor (or irrational) decisions that will not result in an increase in profit, a system will, with overwhelming likelihood, move toward the state which maximizes entropy. Loosely speaking, as a simple matter of pure probability, nature naturally/automatically implements an optimal cost/benefit analysis that maximizes entropy.

Hopefully this discussion of entropy is helpful for those teaching or grappling with what entropy is. Though we don't have specific recommendations for instructors, having explicit conversations about what entropy is and how it differs with other physical quantities with students may help assuage some of entropy's notoriety. This section was intended to help frame such discussions.

[12] Parmenides was an ancient Greek philosopher who, to grossly oversimplify, argued that all change is an illusion because the non-existence of something that exists is impossible. Thus, anything that exists must always have existed since change from non-existence to existence is impossible.

3.8.3 How do Functions of Many Variables Change?

This section was included because, in the Expansions Question, we saw students struggle to determine changes in a quantity, like the entropy of an ideal gas, that is a function of multiple independent variables (e.g. $S(T, V)$). We saw many students fail to realize that since the temperature in the isothermal expansion is constant, only the change in volume mattered for determining the change in entropy. Many students worried, unproductively, about how the work done, or the heat flow would affect the entropy in conjunction with the increase in volume. This struggle with the expansions problem made this question appear to be one of the most difficult sections of the interview for both the graduate and undergraduate students.

In the case of the Expansions Question, the entropy of the system can be derived mathematically. For ideal gases, the Sakur-Tetrode equation defines entropy as a function of energy, volume, and number of particles: $S(U, V, N)$ or $S(T, V, N)$ by using the equipartition theorem to replace U with T. Students seemed more inclined to reason through the change in entropy by thinking in more of a multivariable scalar calculus/co-variational way of, effectively, attempting to determine the signs an relative magnitudes of the partial derivatives of entropy with respect to volume or energy. The most direct way of approaching this problem, however, is to use the concept that the heat flow into the system, Q, is related to the change in entropy through the equation $\Delta S \geq Q/T$.

A speculative explanation for why students commonly utilized the co-variational reasoning is that is, generally, a more applicable technique that applies in many situations. The direct approach, on the other hand, is much more context dependent as it typically only applies to thermodynamic processes. Physics tends to value general problem solving strategies over 'memorization' of specific methods for specific situations.

Since this question about expansions seemed to be the most difficult question on the interview, it warranted extra discussion which perhaps could be beneficial and interesting to instructors. However, the cases of isothermal and free expansions are also quite idiosyncratic, and our population of students is small and affected by selection bias. This makes it difficult to make any kind of recommendation, but perhaps a conceptual review of multivariable integration, the multivariable chain rule, and partial derivatives could help students in situations like this when they must determine the change in a quantity that depends on multiple other quantities.

3.9 Conclusion

The study discussed in the two prior chapters qualitatively analyzed students responses to a questionnaire of conceptual problems related to entropy. We detailed the methodology of the study, discussed student's responses to individual questions, made comparisons between undergraduate and graduate student reasoning, identified common conceptual resources employed by students, and discussed some things thermal physics instructors might be able to take away from this study. The following two chapters will discuss an entirely different project: the quantitative analysis of student collaboration on homework assignments. The two projects are related by their focus on upper-division thermal physics courses.

Chapter 4

Network Analysis of Student Collaboration

4.1 Introduction

Many studies have demonstrated that interactive engagement, which encourages learning through discussion and collaboration, improves student understanding of physics concepts [25, 26]. Additionally, having a sense of belonging within an academic community is associated with persistence and achievement among students, especially among students from underrepresented backgrounds [8, 9]. Homework assignments in physics courses are one of the primary situations in which students can collaborate and form bonds and a sense of belonging, so studying the relationship between collaboration on homework assignments, connection within a community, and performance may shed light on effective ways to create supportive environments within physics courses.

Social network analysis (SNA) provides a quantitative method for analyzing how individuals can be connected to a larger group. Key concepts related to network analysis are discussed in more detail in Sec. 4.2, but briefly, a social network is a complex network (or graph) with individuals represented by nodes (or vertices), while connections between individuals are represented by edges. In a general network both nodes and edges carry weights (indicating characteristics of an individual or the strength of a connection between individuals), and edges can have a direction indicating a non-reciprocal type of connection.

SNA is an analytical tool that is gaining traction within the field of PER, and there is a small but growing collection of studies examining students' connections to their peers and how it relates to student experiences and outcomes. For example, a pair of studies investigated whether students' connection to their classmates (measured via a series of surveys given periodically throughout the semester) predicts their persistence within the introductory physics sequence [27, 28]. The first study found that centrality (i.e., a student's level of connectivity to other students) was a good predictor[1] of persistence. The second study provided a more nuanced analysis that investigated social networks that developed in the physical classroom (the in-class network) and networks the formed outside of the physical classroom (the out-of class network) which incorporated collaboration on homework assignments. This second study found that course grade was more correlated with persistence for students with high final grades (or lack of persistence for students with low final grades), and that centrality in the out-of-class network more correlated with persistence for "middle-of-the-pack" students. With these observations, they concluded that developing social connections outside of the classroom either helped create or reflected an already existing commitment to their studies [28].

Additionally, SNA has been used to explore whether changes in students' feelings of self-efficacy in physics is related their connection to other students [29]. Though this study found that students left their introductory physics courses with a lower average sense of self-efficacy, centrality within the network predicted post-course self-efficacy after controlling for pre-course self-efficacy. Furthermore, centrality measures were associated with various sources of self-efficacy [29]. Another study used SNA to analyze student interactions in a help-room setting and determined that the environment was equitable because gender and ethnicity were not predictors of participation [30]. SNA is a subdivision of the larger field of network analysis which uses graphs and networks to analyze complex systems. Outside of the context of social interactions, networks have been used in PER to study the patterns of student responses to multiple choice surveys [31, 32, 33].

In a direct precursor to this study, Vargas *et al.* created social networks from students' reports of collaboration on homework assignments in three upper-division courses at the Colorado School of Mines (Mines). Various measures of a student's centrality within the network were then correlated with performance on exams and homework assignments. In all three courses, homework scores were positively correlated with several centrality measures, but negatively correlated with measures representing whether student collaborated with only a few versus many other students. The findings suggested that students who collaborate both frequently and with many others tended to perform better on graded assignments [10]. Another closely

[1] Though the cited article used the term 'predictor,' no analysis of causality was performed.

related study examined this same connection between students' performance and their connections to their peers in a highly collaborative introductory physics course and also found a significant link between students' centrality and performance using regression analysis [34].

The study presented in this and the following chapters builds on the study by Vargas *et al.* to examine the relationship between self-reported student collaboration and performance during the COVID-19 pandemic. Data was collected from one course at Mines and two courses at the University of Colorado Boulder (CU Boulder). This allows for a direct comparison between networks from pre-pandemic and pandemic-affected courses at Mines. Furthermore, the data collection in the CU Boulder courses adds an additional perspective on student collaboration by investigating a different student population and educational context.

In this chapter, the methodology of the study will be discussed, including a brief overview of key network analytic concepts. Then the context and structure of each of the courses in the study, data collection methods, and our data analysis process are described. The results of the correlation analysis characterizing the relationship between students' collaboration and performance will be discussed. Finally, the chapter will conclude with a discussion of the distinction between two concepts (reciprocity and symmetry) relevant to the nature of the directed networks constructed in this study. This work also appears in an article currently submitted for publication.

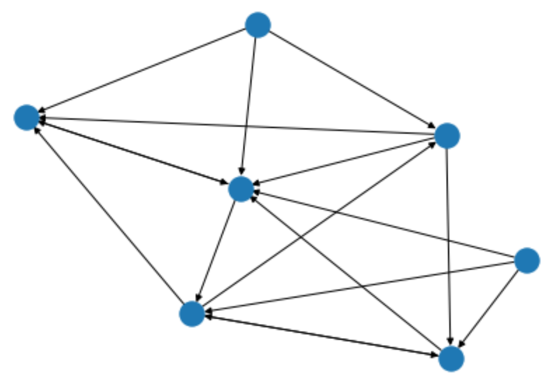
4.2 Methodology

In this section, we provide an introduction to relevant network analytic concepts accessible to readers with no prior experience in network analysis. This includes centrality measures, which quantify a student's connection to their peers (Sec. 4.2.1); the context of the courses from which data was collected including course format, grading structure, and data collection methods (Sec. 4.2.2); and our data analysis process (Sec. 4.2.3).

4.2.1 Overview of Important Network Analysis Concepts

An example of a complex network with seven nodes is shown in Fig. 4.1. An edge connecting two students has both a weight, representing the number of times the pair of students reported working together

Figure 4.1: Example of a Complex Network



while completing a homework assignment, and a direction indicating which student was giving/receiving help. A network in which the edges have a direction is called a *directed network*.

The information about the edges in a network is encoded into an *adjacency matrix*. The adjacency matrix, A, is an $N \times N$ matrix, where N is the number of nodes, and the matrix elements, a_{ij}, are the weights of the edges that connect a node i with the node j. In a directed network, the adjacency matrix is generally not symmetric since directed connections between individuals are not necessarily reciprocal. Additionally, nodes in a network may also contain extra information. In our case, this nodal information includes the homework and exam grades of the student represented by the node.

From the networks, we calculate a number of centrality measures, which quantify a node's connection to the rest of the network. Each centrality measure captures a different way in which a node can be connected to the larger network, which in turn can represent different ways in which the node may contribute to the flow of information within the network.

The notation used to describe centrality measures can be a bit cumbersome, so it is worth a brief description. This thesis will use the same notation as the prior study by Vargas *et al.* which uses a character to label the general 'class' of centrality (e.g. s, c, and Y), a superscript to explicitly disambiguate the specific centrality measure in question, and a subscript to indicate that the centrality measures for all the nodes in a network can be collected in a list, or vector, of values enumerated by nodes. This notation convention appears common within the network analysis community since it also used in two well-known textbooks [35, 36]. However, it is not unusual to see a notation with the subscript index replaced with function notation which treats the node as the argument of the function.

The simplest of these centrality measures are the in-strength and out-strength. In-strength s_i^{in} and out-strength s_i^{out} quantify the total weight of edges terminating and beginning on a node, respectively. With an adjacency matrix, these centrality measures can be calculated with the following equations:

$$\begin{aligned} s_i^{\text{in}} &= \sum_j^N a_{ij}^T \\ s_i^{\text{out}} &= \sum_j^N a_{ij} \end{aligned} \tag{4.1}$$

where N is the total number of nodes in the network. It is very important to note that these equations establish an 'orientation' to the adjacency matrix in which the entries in row i represent the edges pointing *out of* the node i. Furthermore, the entries in column j represent the edges that point into node j (from all the other nodes in the network). So, to get the in-strength, we sum over the elements in each column or, equivalently, elements in the *rows* of the transpose of the adjacency matrix. To get the out-strength, we must sum over the elements in each row.

This convention of having the inward edges in the rows is somewhat arbitrary; we could just as easily used the opposite convention. Extreme care was taken in keeping track of the proper orientation of the matrices containing our data so that it was correctly handled by NetworkX, the python package used to construct and analyze the networks. Fortunately, the solution to correctly orienting the adjacency matrix is simply a transposition, but still, my quantum field theory lecturer from the University of Cambridge stands correct with his assertion that "keeping track of minus signs, i's, and 2π's is the hardest thing in physics," if one considers the difference between a matrix and its transpose equivalent to a sign swap.

With the in- and out-strength, we can calculate the net-strength s_i^{net} which is simply the difference:

$$s_i^{\text{net}} = s_i^{\text{in}} - s_i^{\text{out}}. \tag{4.2}$$

For an illustration of these centrality measures, the node in Fig. 4.2 has an in-strength of eight, an out-strength of four, and a net-strength of four. In our networks, the total number of times a student gave help to other students over the course of the semester is that student's out-strength. The number of times they received help is their in-strength.

The in-disparity Y_i^{in} and out-disparity Y_i^{out} measure the non-uniformity of a node's inward and outward edges, respectively, and provide more information about the distribution of edges attached to a node. Nodes with large disparities tend to be connected to very few other nodes within the network, reflecting the large disparity between the few, present connections to other nodes and the many non-existent connections. A node's disparity tends to decrease the more connections the node has to other nodes, and is not defined for nodes without connections. The in- and out-disparities differ from s_i^{in} and s_i^{out}. For example, a node with a very strong connection to just one other node has large in- and out- strengths and large disparities. In contrast, a node with many weak connections to other nodes will also have large values for s_i^{in} and s_i^{out}, but small values for Y_i^{in} and Y_i^{out}. So, for example, a student who has a strong connection to a few other students will have large disparities and large in- and out-strengths. Alternatively, a student with many relatively weaker connections to other students will have smaller disparities, but could also have large in- and out-strengths depending on the sum of the strengths of the connections. The equations for the in- and out-disparities are a little complicated:

$$\begin{aligned} Y_i^{\text{in}} &= \frac{1}{(s_i^{\text{in}})^2}\sum_j (a_{ij}^T)^2 = \frac{\sum_j (a_{ij}^T)^2}{(\sum_j (a_{ij}^T))^2} \\ Y_i^{\text{out}} &= \frac{1}{(s_i^{\text{out}})^2}\sum_j (a_{ij})^2 = \frac{\sum_j (a_{ij})^2}{(\sum_j (a_{ij}))^2} \end{aligned} \tag{4.3}$$

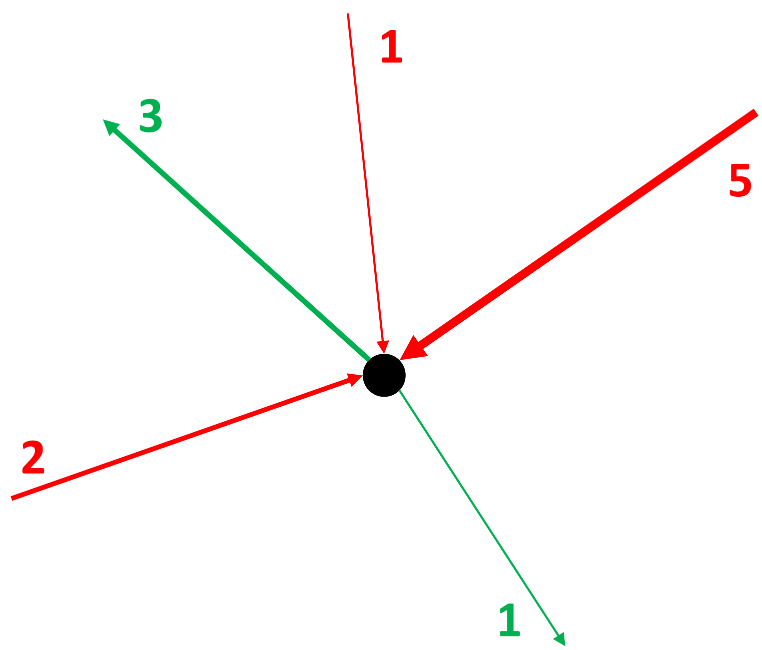
but basically they are the ratio of a sum of squares to a square of a sum. The in-disparity is the ratio of the sum of the squares of the weights of the inward links to the square of the sum of the inward links for a node. The out-disparity is the same as the in-disparity except for the outward edges. For an example, the node in Fig. 4.2 has an in-disparity of 30/64 and an out-disparity of 10/16.

The closeness c^C and harmonic c^H centralities measure how close a node is to all the other nodes within a network. The closeness centrality of node i is

$$c_i^C = \frac{n-1}{N-1}\frac{n-1}{\sum_{i\neq j} d_{ij}} \tag{4.4}$$

where N is the total number of nodes in the network, n is the number of nodes reachable (i.e., able to be

Figure 4.2: Sample Node Connected to a Larger Network



An isolated node with three inward edges and two outward edges. This node has in-strength $s_{\mathrm{i}} = 8$, out-strength $s_{\mathrm{o}} = 4$, and net-strength $s_{\mathrm{n}} = 4$.

reached by traversing one-way edges) from node i, and d_{ij} is the shortest distance between nodes i and j. If node j is not reachable from node i, then the d_{ij} term is not considered in the sum and the $(n-1)/(N-1)$ pre-factor scales the closeness centrality by the number of reachable nodes. The distance d_{ij} can be related to edge weights in a variety of ways. We discuss our definition of distance in Sec. 4.2.3 which addresses the fact that in our network strong connections mean short distances. This formula for the closeness centrality was proposed by Wasserman and Faust specifically to account for the case of networks in which some nodes are unreachable from others [37].

Another way to address the possibility of certain nodes being unreachable is to use the harmonic centrality:

$$c_i^H = \sum_{i \neq j} \frac{1}{d_{ij}}. \tag{4.5}$$

where d_{ij} is the shortest distance between node i and node j. If a node j is unreachable from node i, d_{ij} is effectively infinite and the term in the sum for this pair of nodes is zero. Within the context of social networks, the closeness and harmonic centralities capture the idea of 'degrees of separation'. If a student has many connections to their classmates, and if their connections also have many connections (and so on) the first student will have high closeness and harmonic centralities. Additionally, this means that for a student to be unreachable from another, they must share zero mutual collaborators at all levels (i.e., not only do they share no mutual collaborators, their mutual collaborators share no mutual collaborators, and so on).

With regards to directed networks, an important subtlety of the closeness and harmonic centralities is that they can be defined using either the shortest inward directed path or the shortest outward directed path. For example, if a node n has only outward directed edges, it is not reachable from other nodes in the network, but other nodes will be reachable from n. In this case using the inward shortest paths to compute c^C or c^H will result in centralities of zero, but using the outward paths will result in non-zero centralities. When discussing our results in later sections, we will refer to the closeness and harmonic centralities calculated using the inward shortest paths as c^{Ci} and c^{Hi}, respectively, and the quantities calculated using the outward distances as c^{Co} and c^{Ho}.

The last centrality measure we will consider is the betweenness centrality. For a node i, the betweenness centrality c^B is

$$c_i^B = \sum_{j,k \in V} \frac{\sigma(j,k|i)}{\sigma(j,k)} \tag{4.6}$$

where $\sigma(j,k)$ is the number of distinct shortest paths between nodes j and k, $\sigma(j,k|i)$ is the number of shortest paths between nodes j and k that pass through node i, V is the set of of nodes in the network and the sum runs over all pairs of nodes in the network (excluding pairs of nodes containing node i) [38]. Conceptually, the betweenness centrality quantifies the extent to which a node is a hub that provides connections between different regions within a network. So, a student who collaborates with two (or more) tight-knit groups which would otherwise be disconnected will have a large betweenness centrality.

These centrality measures can be broken down into two groups: local centrality measures which only consider a node i and the set of nodes directly connected to i, and global centrality measures which depend on the structure of the entire network. The in-strength, out-strength, net-strength, in-disparity, and out-disparity are all local centrality measures while the harmonic, closeness, and betweenness centralities are global measures.

The final network analysis concept relevant to this study (which was not considered in the previous study by Vargas *et al.* [10]) is reciprocity [35, 36]. Reciprocity r is only meaningful in directed graphs. In the case of an unweighted network, the reciprocity is

$$r = \frac{1}{m} \sum_{ij} a_{ij} a_{ji} = \frac{1}{m} \operatorname{Tr} \mathbf{A}^2 \tag{4.7}$$

where m is the total number of edges in the network. The reciprocity of a single node can also be defined by considering only the edges attached to the node. Furthermore, the reciprocity can be defined for networks with weighted edges as described by Squartini *et al.* [39]. Conceptually, reciprocity measures the extent to which the connections between nodes are bilateral. If a network has few pairs of bilateral edges between nodes then the reciprocity will be low. In our networks, reciprocity can arise from one of two cases: either from a single student report both getting help from and giving help to a second student, or from a pair of students both reporting getting (or giving) help to each other.

4.2.2 Context

CU Boulder: Thermal Physics

The two CU Boulder courses from which data was collected occurred during the fall 2020 and spring 2021 semesters. Both courses were an upper-division thermal physics course and taught by the same instructor and occurred in a hybrid format with an option to attend synchronous lectures either in-person or remotely and asynchronous lecture recordings available. Each course had a total of 12 weekly homework assignments, with collaboration data collected from all but the first assignment.

In the fall 2020 iteration, data on student collaboration was collected through a Qualtrics survey that students completed upon submission of their weekly homework assignments. In the spring 2021 semester, students reported their collaborators directly on their homework solutions. The fall course had three take-home midterm exams and a final with each of the four exams comprising 15% of a student's final grade. The spring course had only two take-home midterms and a final with each of the three exams comprising 20% of student's final grade. Students were allowed to submit revisions on both homework assignments and exams for the opportunity to earn back missing points.

Typically, on-sequence physics majors at CU Boulder take this thermal physics course during the fall of their senior year. This is reflected in the total enrollments of the two courses: 83 students took the course in the fall, and 55 took the course in the spring. The process of obtaining consent for collection of student data decreased the number of students from whom data was collected from 83 to 53 in the fall term and 55 to 48 in the spring term. We investigate the possible effects of this missing data in Sec. 5.4. CU Boulder is a large research institution with an undergraduate population of roughly 30,000 students with roughly 110 physics majors and 25 engineering physics majors per class year.

Colorado School of Mines: Math Methods

At Mines, collaboration data was collected from an intermediate-level Math Methods course during the fall 2020 semester, which was on-sequence with the normal curriculum at Mines. The course was fully remote and taught synchronously. Students reported their collaborators directly on their homework assignments, just as in the spring iteration of the CU Boulder thermal physics course.

Graded assignments in the course consisted of seven five-point homework assignments, a course project broken in two six 7.5-point assignments across the semester, and 15 points of participation (for a total of 95 possible points). The course had no exams. Mines students also had the opportunity to submit homework revisions to receive points back.

A total of 27 students enrolled in the course, and 23 students consented to the study. Mines is a medium research institution with roughly 5,000 undergraduates, and about 60 physics majors per class year.

4.2.3 Analysis

The first step in processing the collaboration data was to create the adjacency matrix which encodes all students' reports of getting and giving help across the semester. For all three courses, this adjacency matrix was built-up assignment by assignment. For each assignment, two separate matrices were created: one containing all reports of getting help (the "*got-help*" matrix) for the assignment, and one containing all reports of giving help (the "*helped*" matrix) for the assignment. Each of these matrices representing the collaboration on a single assignment contain only ones and zeros. The helped matrix was transposed so that the direction of edges would match that of the got-help matrix. The got-help and transposed-helped matrices were then combined using an element-wise logical OR[2] operation, which repeats the analysis performed in the prior study by Vargas *et al.*, to create one "combined" matrix for each homework assignment [10]. Finally, the individual combined matrices for each assignment were summed element-wise to create the "total" adjacency matrix A_{tot} representing all collaboration during the course. As a final step, the diagonal of the adjacency matrix was set to zero to ignore any cases of a student reporting themself.

This total adjacency matrix can be used to directly compute s_i^{in}, s_i^{out}, s_i^{net}, Y_i^{in}, and Y_i^{out} for each node according to the equations described in Sec 4.2.1 (and also in [10]). To compute the closeness, harmonic, and betweenness centralities, the matrix of distances d_{ij} must first be calculated. Entries in A_{tot} range from zero to eleven in the CU Boulder courses and from zero to seven in the Mines course, and large values in A_{tot} represent a large amount of collaboration between students. Large values in A_{tot} should correspond to

[2] Other methods for combining student reports were explored. These other methods are described in Appendix D, but did not result in particularly or meaningfully different results.

short distances in d_{ij}. To create a d_{ij} consistent with this relationship between collaboration and distance, we took the reciprocal of the elements of A_{tot}, unless the element was zero in which case it remained zero. In either case (in A_{tot} or d_{ij}), an edge weight of zero between two nodes means the two nodes are not directly connected. Since there was a finite number of homework assignments, the shortest distance between a pair of students is one over the total number of homework assignments.

The calculation of the closeness, harmonic, and betweenness centralities was done using built-in functions from the NetworkX python package. When calculating the closeness and harmonic centralities in a directed network, the NetworkX functions default to calculating c_i^{Ci} and c_i^{Hi}. To calculate c_i^{Co} and c_i^{Ho}, the directions of all edges in the adjacency matrix are swapped which is accomplished, in practice, by transposing the adjacency matrix.

The output from the functions which calculate the centrality measures are dictionaries with the node label as the key and the centrality measure of the node as the value. With the dictionaries, the Pearson correlations between the centrality measures and students' homework and exam grades were directly calculated. The students in the CU Boulder courses could submit revisions to exams to receive a percentage of points back, but we calculated correlations using the pre-revised exam scores. Correlations with homework scores were calculated using the post-revised grade as the pre-revised grade was never recorded for these courses.

Only reports from students consenting to the study were included in the construction of the social network. In some instances, though, a student would report collaboration with a student who had not consented to the study. In this case, a node and appropriate edges (depending on the reports from consenting students) for the non-consenting student would be added and included in the final network representing only information reported from the other, consenting student (i.e., collaboration reports submitted by non-consenting students as part of the normal coursework were not included in the data collection or the network). Thus, some non-consenting students appear in the final networks, the presence of which contributes to the centrality measures for other students in the network. However, the non-consenting students were not included in the calculations of correlations as data on their course scores was not collected as part of the study.

Table 4.1: Summary of Collected Collaboration Data

	Participating students	Total enrollment	Reports: getting help	Reports: giving help	Sum of network edge weights	Reports per student per assignment	Edge weight per student per assignment
CU FA20	53	83	777	691	1110	2.52	1.90
CU SP21	48	55	378	335	495	1.35	.94
Mines FA20	23	27	140	138	208	1.73	1.29

A summary of the results of the data collection on student collaboration. The final two columns directly compare the level of student collaboration since the courses at Mines and CU Boulder had a different number of homework assignments.

4.3 Results

One of the main goals of this study was to compare the networks and correlations between student centrality and performance between courses taking place before the COVID-19 pandemic and courses occurring during the pandemic. An overview of the collected data in this study is provided in Table 4.1. In all three courses, students more often reported getting help than giving help, though the difference was smallest by far in the course at Mines. This observation cannot be compared to the prior study since only aggregated adjacency matrices remain from the prior study, but it suggests there may be a different conceptualization about the nature of giving and receiving help between students at Mines and CU Boulder.

Likewise, the frequency of reporting also cannot be directly compared between the pre-pandemic and pandemic-affected courses, but the density of the networks, in terms of the number of edges per student per assignment present in the full networks can be compared. In the courses analyzed in the prior study, the classical mechanics, electromagnetism, and quantum mechanics courses had 1.4, 1.5, and 1.7 edges per student per assignment. Except for the spring 2021 thermal physics course at CU Boulder, the level of collaboration occurring between students in the pandemic-affected courses is relatively similar to the pre-pandemic courses (see Table 4.1). In the fall 2020 course at CU Boulder and the course at Mines, students, on average, gave or received help from more than one other student on each assignment (see Table 4.1). This may be unexpected since the COVID-19 pandemic made face-to-face collaboration more difficult between students, but students' definition and threshold for reporting giving and receiving help may have changed due to the pandemic and online modes of interaction became more common.

The correlations between centrality measures and student performance in the three pandemic-affected courses are shown in Fig. 4.3. The results from the math methods course at Mines very closely matched the patterns seen in the pre-pandemic courses at Mines [10]. The fall 2020 course at Mines had strong correlations between performance and the in-strength, out-strength, closeness centrality, and harmonic centrality. Furthermore, the non-significant correlations between the net-strength and betweenness centrality also matched the pre-pandemic findings. The largest difference between the pre-pandemic and pandemic-affected courses at Mines is that during the COVID-19 pandemic, the in-disparity and out-disparity did not replicate statistically significant correlations with performance. Among the three pre-pandemic courses, though, the in- and out-disparities correlated negatively with performance in two of the courses, but not the third (Quantum Mechanics). The negative correlation suggested that students who collaborated with many other students (as opposed to only few other students) tended to get higher scores [10]. This replication of the patterns from the pre-pandemic courses indicates that the environment at Mines was largely able to preserve the connection between collaboration and performance during the pandemic despite the course being fully remote.

The fall 2020 iteration of the thermal physics course at CU Boulder showed some of the same patterns as the courses from Mines. The outward closeness centrality and outward harmonic centrality were positively correlated with performance on homework assignments and exams and at the $p < .05$ level with performance on homework assignments, and the out-strength (in addition to $c_i^{Co}$and c_i^{Ho}) was positively correlated with performance on homework assignments at the $p < .05$ level. Since outward directed edges from a node represent giving help to other students, the statistical significance of only the outward oriented centrality measures indicates that students who were a source of help close to many other students in the course tended to get higher grades. This could suggest that students doing well in the course were more able to provide help to others, or that providing help to others (but not receiving help) improves a student's performance.

The nuances of this distinction between the inward closeness/harmonic centralities and the outward closeness/harmonic centralities were not explored in the prior study. Both before and during the pandemic, in

Table 4.2: Weighted and Unweighted Reciprocities of Collaboration Networks.

Course	Weighted Reciprocity	Unweighted Reciprocity
Mines Fall 2020	.990	.978
CU Boulder Fall 2020	.861	.792
CU Boulder Spring 2021	.848	.914

The reciprocity for each of the courses in the present study. The reciprocity takes on values from 0 (for no reciprocity) to 1 (completely reciprocal). The method proposed by Squartini et. al was used to calculated the reciprocity taking edge weight into account [39]

the courses at Mines the correlations for the inward and outward centralities were nearly identical rendering a discussion of the distinction somewhat irrelevant. In both courses at CU Boulder, however, different correlations were observed which indicated that being a well-connected source of help was associated with high performance.

To investigate the different correlations between the inward-directed versus outward-directed closeness/harmonic centralities and grades observed in the CU Boulder courses (and lack of difference seen in the Mines courses, see Fig. 4.3) we looked at the reciprocity of each network. Reciprocity measures the degree to which a connection between nodes is bi-directional (i.e., that an edge from node i to node j has a matching edge from j to i). Using the method proposed by Squartini *et. al.* for calculating reciprocity in weighted networks [39], we found that all networks were highly reciprocal.

The reciprocities calculated in Table 4.2 are higher than typically seen in social networks [39, 35]. In the case of the Mines course, this finding appears consistent with the symmetry in the correlations between inward- versus outward-directed closeness and harmonic centralities and grades. Because, if for every outward directed edge there is a corresponding (equally weighted) inward edge, then the outward shortest distances from a node to all other nodes will be identical to the inward shortest distances. The reciprocities in the CU Boulder networks were lower than those for the Mines network, though still relatively large. Despite these large reciprocities, there is an asymmetry between the correlations with the inward versus outward centrality measures.

Such a result may seem counter intuitive, but it is not unexpected since reciprocity within a directed network, in general, is not related to the symmetry of its adjacency matrix [39, 40]. The larger asymmetry between students' reports of receiving versus giving help is a possible explanation for the asymmetry in the correlations, but is not, by itself, sufficient to account for the asymmetry as we will show in Chapter 5). This lack of consistency in the correlations at Mines and CU Boulder is possibly a sign of a cultural difference in how students define thresholds for giving and receiving help or in how students collaborate in the two institutions.

Correlations between a student's centrality and performance were almost non-existent in the spring 2021 iteration of the thermal physics course at CU Boulder. The only statistically significant correlation was between the students' net-strength and exam performance. This correlation was negative and, since the net-strength is the in-strength minus the out-strength, indicates that students who received more help than they gave tended to get lower scores on exams. This result appears consistent with and somewhat complementary to the results from the fall iteration of the course that students who provided help tended to perform better. Another feature of this course in the spring of 2021 was that it had the least level of student interaction in terms of edges and reports per student (see Table 4.1) and was the only course which contained multiple, disconnected components[3] (see Appendix C), which could be a consequence of the course occurring in the off-sequence semester, thus students in the spring semester course may be less likely to know each other and less likely to have taken prior courses together. Additionally, students may have been less engaged overall in the spring due to increased pandemic-related burnout, an effect which may be compounded by the large proportion of second semester seniors in the course who were approaching graduation. All these possible explanations are speculation; interviews with students in the course would provide more insight into these results.

Overall, these results show that the format of the course (whether in-person, remote, or hybrid), does not appear to dictate whether there will be a correlation between student collaboration and performance. The environment at Mines maintained the connection despite a fully remote instruction format during the COVID-19 pandemic, while the hybrid instruction format at CU Boulder was not able to consistently produce a connection despite having an option for in-person lectures during both semesters. To determine whether the correlation is just typically weaker at CU Boulder and whether the variation in results between the spring

[3] Components are collections of nodes completely disconnected from each other.

and fall semesters is due to on/off-sequence effects versus pandemic related burnout more research would need to be done after a return to normal instruction.

4.4 Reciprocity versus Symmetry

We decided to investigate the reciprocity within our networks as a means of trying to understand the observed symmetry (in the Mines network) and lack of symmetry (in the CU Boulder networks) between the inward and outward directed closeness and harmonic centralities. The first indication of this difference between the networks from Mines and CU Boulder was in the raw data of student reports. The students at Mines, with almost no exceptions, would report identical lists of helpers and helpees. That is, if student A reported getting help from students C, D, and G, then student A would also report helping students C, D, and G. Students at CU Boulder were much more likely to report different lists of helpers/helpees.

It is easy to see how symmetry in reporting will result in a symmetry between the inward and outward closeness/harmonic centralities. If for every edge out of a node there is a corresponding inward edge (starting and terminating on the appropriate, corresponding nodes), then all of the outward shortest distances will be identical to all the inward shortest distances. Effectively, this results in an undirected network in which there is no distinction (and in which it is impossible to define a difference) between inward and outward centrality measures. So, this symmetry in the Mines network (in the reports from students) results in reciprocity values that are nearly unity.

However, in general the symmetry in a network cannot be assumed to be related to the reciprocity in any simple way. As discussed in articles by Squartini *et al.* and Garlaschelli *et al.*, reciprocity and symmetry are distinct concepts in directed networks [39, 40]. Because of the case, discussed above, with the Mines network the strong caveat made by Squartini *et al.* that "the symmetry of weights (i.e. $w_{ij} = w_{ji}$) [in the adjacency matrix] is completely uninformative of the reciprocity structure" was initially quite confusing. Squartini *et al.* give two reasons for this claim which appear sound yet somewhat opaque to those[4] without much formal education in network analysis. In my attempts to understand this claim, I was able to create a simple counterexample which demonstrates one case where high symmetry does not produce

[4] i.e. me

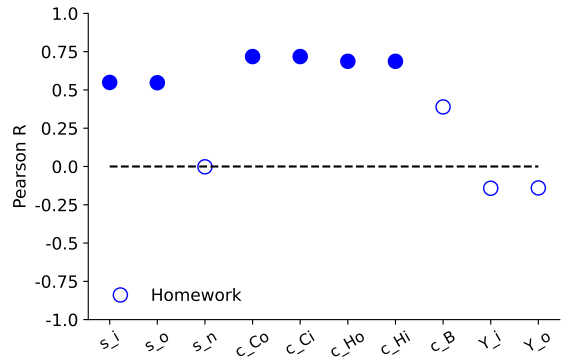
a large reciprocity.

If an adjacency matrix has many pairs of symmetric elements with small values, but a few non-symmetric elements with a large difference, overall there will be relatively large symmetry but low reciprocity when weight of the edges is accounted for. For an example consider the matrix below (note the italicized and bolded pairs of elements):

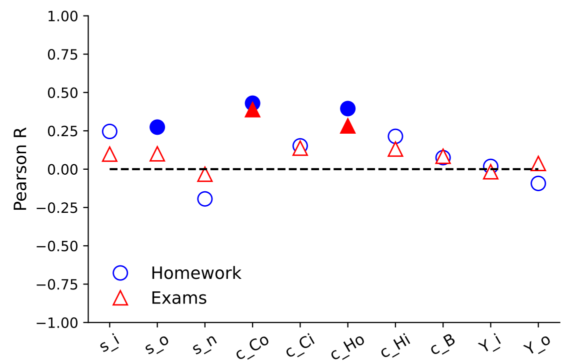
$$\begin{pmatrix} 0 & 1 & 2 & 1 & \mathit{21} \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & \mathbf{0} \\ 1 & 2 & 1 & 0 & 1 \\ \mathit{1} & 1 & \mathbf{18} & 1 & 0 \end{pmatrix}.$$

This counterexample, however, assumes a particular interpretation of the phrase "the symmetry of weights (i.e. $w_{ij} = w_{ji}$)" from the Squartini *et al.* article. 'Symmetry' does not appear to be a mathematically defined quantity within the discipline of network analysis, so there is some ambiguity in what Squartini *et al.* mean when talking about "the symmetry of weights." In constructing the counterexample, symmetry was taken to mean the number of pairs of diagonal entries in the adjacency matrix for which the equality $w_{ij} = w_{ji}$ holds. This does appear to be an appropriate interpretation since an alternative definition for symmetry which 'gives partial credit' to cross-diagonal pairs that are nearly, but not precisely equal would most likely, in fact, be the reciprocity. If this is the case, then reciprocity could be considered a generalization of symmetry that more finely accounts for the case where the pairs of elements (w_{ij}, w_{ji}) are nearly, but not precisely, equal.

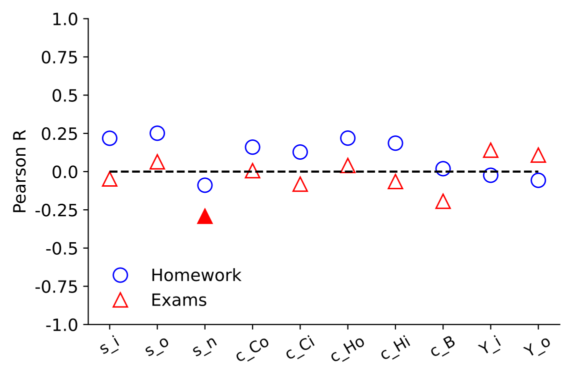
Figure 4.3: Correlations Between Centrality and Student Performance



(a) Fall 2020 Math Methods at Mines



(b) Fall 2020 Thermal Physics at CU Boulder



(c) Spring 2021 Thermal Physics at CU Boulder

Correlations significant at the $p < .05$ level are indicated with filled markers. The math methods course at Mines had no exams. The correlations between homework scores and closeness centrality in the math methods course at Mines and the fall 2020 iteration of the thermal physics course at CU Boulder satisfy the Bonferroni test, which increases the threshold for statistical significance when many tests for statistical significance are performed.

Chapter 5

Statistical Simulation of Collaboration Networks

5.1 Introduction

Networks are complex, non-linear objects. A small perturbation in a network could have anywhere from a negligible to a large affect on the network depending on the location and nature of the perturbation. As in any experimental study, our data collection is susceptible to random and systematic measurement errors. In particular, a cursory analysis of the consistency of student reporting in this study indicates that students may have different thresholds for what qualifies as giving or receiving help which can result in either missing or spurious edges. The effect of this inconsistency is partially addressed by combining student reports using a logical OR operation as discussed in Sec. 4.2.3, but as with any human subjects research, some level of human error is expected. Additionally, collaboration reports from students who did not consent to the study were not included in the construction of the social networks resulting in both missing edges and missing nodes.

To better understand the significance of our findings given the limitations of the data collection process, we conducted analyses of random networks and the effect that removing nodes from networks has on the correlations between centrality measures and performance. Due to the rather large proportion of missing nodes in the fall 2020 course at CU Boulder, most of our analyses focus particularly on understanding how the results from this semester's course may be affected by the missing nodes. As noted by a recent study on the effects of missing data on robustness of centrality measures, most applied network studies do little more than acknowledge that measurement errors may have occurred [41]; and in addition to better understanding the significance of our results, we hope to expand the knowledge of, and introduce a practice of, error analysis into social network research within the PER field. This work also appears in an article currently submitted for publication.

5.2 Models for Simulating Collaboration Networks

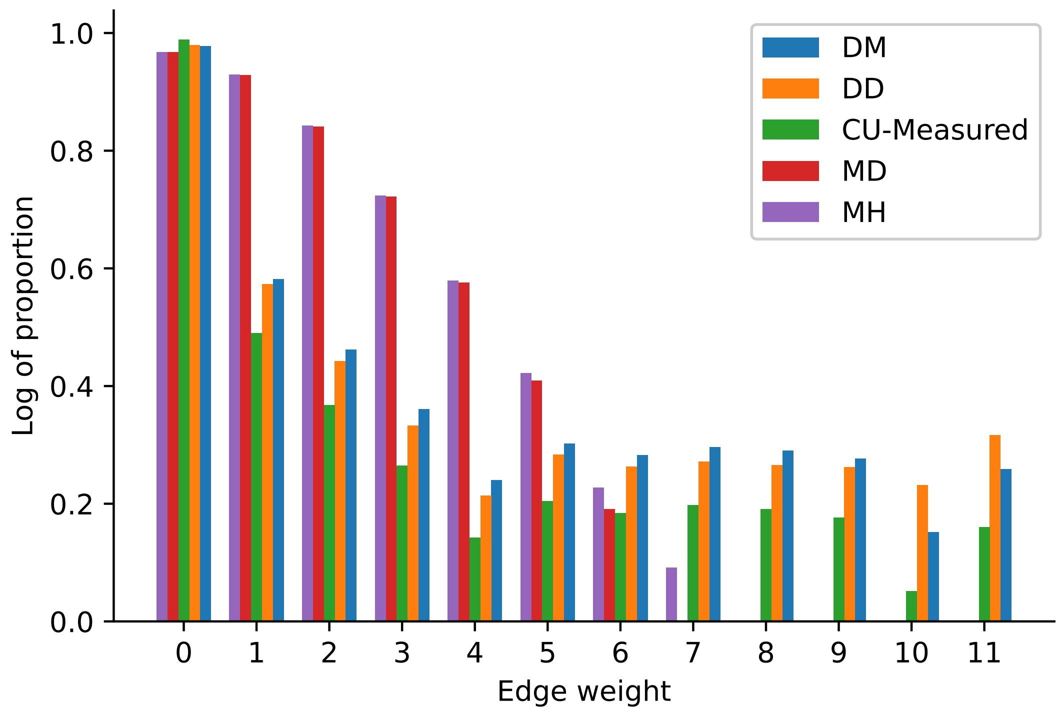
To provide a deeper perspective, beyond the simple application of a t-test, on the significance of our results, we developed four methods for constructing random networks. Real world networks, generally, are poorly modeled by random networks [42], so we attempted to create models which more closely matched the structure of the networks we measured. Simulating networks directly explores randomness in networks and helps to demonstrate how our results compare to ensembles of networks with similar characteristics, and demonstrate that the results we reported in Sec. 4.3 are unlikely to be due to random chance.

Data used to simulate the networks were taken from the fall 2020 CU Boulder thermal physics course as it was the course with the largest enrollment but the smallest fraction of participating students. While accumulating the network for this course from the 11 homework assignments, all reports of getting help were compiled into a "got-help" adjacency matrix. A second matrix containing all the reports of helping another student, the "helped" matrix, was also compiled. Elements in these matrices had values between 0 and 11 (inclusive), and the edge distribution represents the relative abundance of each of the possible edge strengths between nodes.

The most basic method for simulating social networks, which will be referred to as the Degree Match (DM) method, used the edge distributions from the CU Boulder fall 2020 course's got-help and helped matrices to construct simulated got-help and helped matrices which matched the distribution of edge weights seen in the fall 2020 CU Boulder thermal physics course. Then the simulated helped matrix was transposed for the sake of directly matching the analysis performed on the real networks. The two simulated adjacency matrices were then combined using an element-wise maximum function to mimic the effect of using the logical OR operation in the analysis of the real network data. As expected, the DM method succeeded in matching the distribution of edge weights seen in the measured networks (see Fig. 5.1).

The second method, which will be called the Multiple Homeworks (MH) method, uses the simple

Figure 5.1: Edge Distributions in Real and Simulated Courses



Distribution of edge weights in the fall 2020 CU Boulder thermal physics course and in simulated networks. The height of the bars is the log of the ratio of the number of edges of a particular weight to square of the number of nodes, i.e., the total number of ordered pairs of nodes. The column for zero weight represents the proportion of pairs of nodes that are not directly connected.

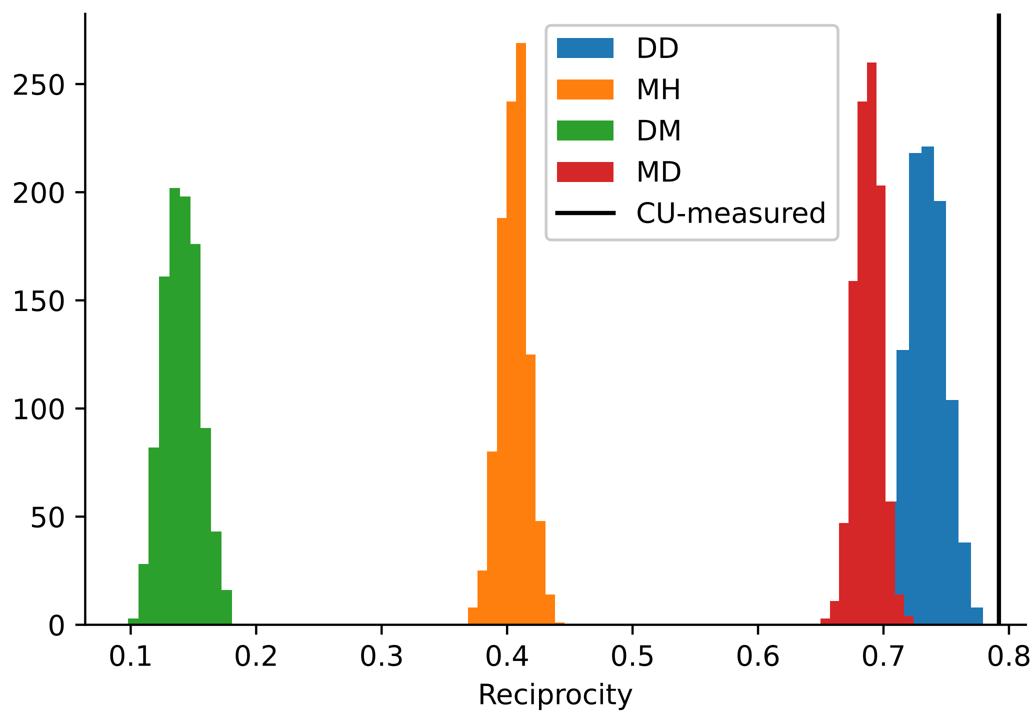
probability that a pair of students collaborated to create 11 pairs of got-helped and helped adjacency matrices, one pair for each homework assignment. Unlike the DM method, the matrices simulated in the MH method only contain ones and zeros which were pulled from a Bernoulli distribution. The helped matrices in each pair was transposed to match the original analysis process, then a logical OR operation was applied assignment-by-assignment and the total network is created by accumulating the combined helped/got-help adjacency matrices over all assignments. In further dissimilarity with the DM method, the MH method does not succeed in matching the distribution of edge weights seen in the real networks. As can be seen in Fig. 5.1, the MH method results in more low-weight connections and fewer high-weight connections between nodes than is seen in the real data.

In both of these methods above, the helped and got-help adjacency matrices were created from independent probability distributions. However, as noted in Sec. 4.3, our measured networks were all highly reciprocal. Upon analysis of the random networks generated with the MH and DM methods, we found they lacked the level of reciprocity seen in the real networks (see Fig. 5.2). This motivated the development of our third and fourth methods to make our simulations better match the real networks.

The Multiple Dependent (MD) method is identical to the MH method except it only randomly generates 11 got-help adjacency matrices (with the same probability of an edge as the MH method) to simulate the 11 homework assignments. Then, the helped adjacency matrices for each assignment are generated depending on the respective got-help adjacency matrix. If there was a got-help report between a pair of students (i.e., student A reports getting help from student B) there was a relatively large probability that there would also be a matching report of helping between the students in the helped matrix (specifically, student A would also report giving help to student B). If there was no report of getting help between a pair of students, there was still a small probability that a report of helping would exist. In a similar manner as the MH method, the MD method fails to match the edge distribution seen in the real networks (see Fig. 5.1).

Our final method, called the Direct Dependent (DD) method, comes the closest to matching the level of reciprocity seen in the real networks (see Fig. 5.2). The DD method creates a single got-help adjacency matrix representing reports of getting help from across all homework assignments using the same edge distribution as the DM method. However, similar to the MD method, the adjacency matrix representing all

Figure 5.2: Reciprocity in Simulated Networks



Distributions of reciprocities for each of the four random network methods developed in this study. For each method, 1000 random networks were created.

Table 5.1: Summary of Network Simulation Methods

	Simulates multiple assignments	Single matrix for all assignments
Independent sub-networks	Multiple HWs (MH)	Degree Match (DM)
Dependent sub-networks	Multiple Dependent (MD)	Direct Dependent (DD)

A summary of the key features of the four methods developed to simulate networks of student collaboration. The row for independent/dependent sub-networks refers to whether the got-help and helped matrices were created independently or whether the helped matrix depended on the got-help matrix.

reports of giving help across the course is created based on the got-help matrix. So, if there was a report of getting help between a pair of students (say student A reports getting help from student B), there was a high likelihood that there would be a reciprocal report of giving help between the pair (i.e., student A would also have given help to student B). Similar to the DM method, the DD method generated networks that better matched the edge distribution of the real networks. A summary of the key features of the methods are supplied in Tab. 5.1.

These second two methods required some tuning of the dependent probabilities, but we were successful in creating networks with the correct number of total students reports of getting and giving help we saw in the real networks. Furthermore, both methods greatly increased the reciprocity we saw in the simulated networks, nearly to the level seen in the real networks (see Fig. 5.2). Importantly, this finding indicates that the reciprocity seen in the real networks is not a consequence of combining student reports using a logical OR operation, but rather a real signal that student collaboration tends to be highly reciprocal since low reciprocity was seen in the DM and MH networks despite the use of the logical OR combination method.

5.3 Statistics of Simulation Models

An additional benefit of constructing the models for simulating course collaboration was a valuable statistics lesson. In the previous chapter, a t-test was used to test the statistical significance of the correlations between centrality and course performance, and these models provided an opportunity to validate the t-test as the appropriate statistical test to use in this situation.

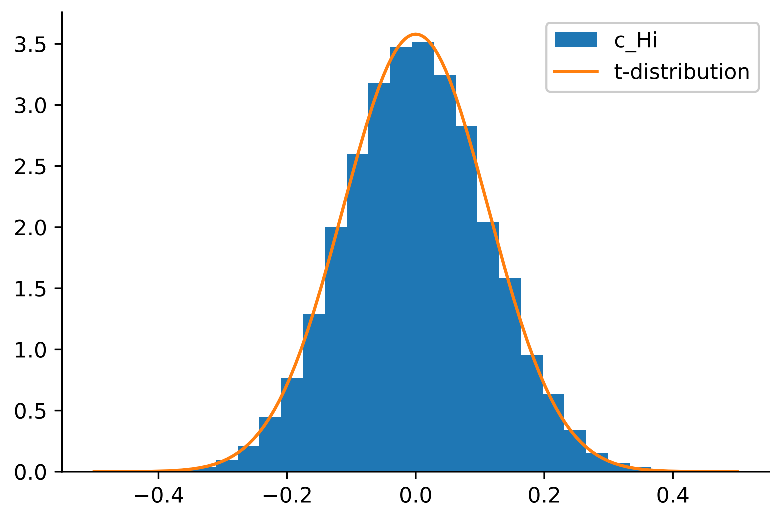
In summary, a t-test is based on a t-distribution. If many pairs of random lists of numbers are generated, the correlations between the pairs of lists are calculated (where one correlation is calculated between the two lists in each pair), and all correlations are placed into a histogram, then this histogram will follow a t-distribution: a probability density function (PDF) which is a function of the length of the random lists. Like many PDFs, a t-distribution resembles a normal (Gaussian) distribution and is centered at zero (see Fig. 5.3 for plots of t-distributions).

The histogram of the correlations from the random pairs of lists is called a sampling distribution. It shows the effective probability (on the $y-$axis) of the correlation (on the $x-$axis) between two random lists. As can be seen in Fig. 5.3, the t-/sampling distributions are peak around zero, meaning that correlations near zero are the most likely. However, there is a non-zero probability that, by random chance, a relatively larger correlation will appear. The purpose of a statistical test is to ask the question: how likely was the correlation calculated between collected data to be just by random chance? Typically a threshold is chosen, above which we say the correlation was unlikely to be due to random chance. This threshold is called a p-value.

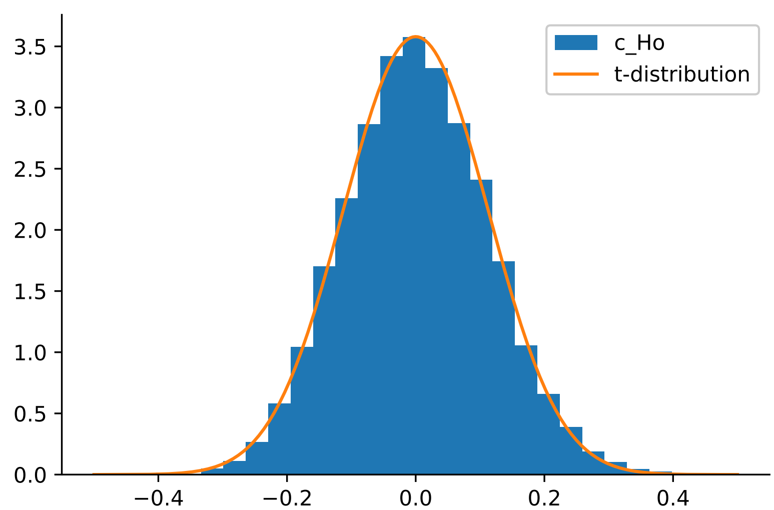
The standard statistical test for correlations is a t-test, however, the t-sampling distribution makes some assumptions about the statistics of the two lists between which the correlation is being calculated. As someone with little background in statistics, it was initially unclear to me whether the lists of student centrality and student grades (between which we were calculating correlations) would satisfy the assumptions necessary for the t-test to be an appropriate test of statistical significance of our calculated correlations. So, as an extra learning exercise, I created sampling distributions by calculating the correlation between centrality (in networks generated using the DD model described in the previous chapter) and student grades (from the Fall 2020 CU Boulder thermal physics course). As can be seen in Fig. 5.3, these more *a priori* sampling distributions closely match the t-distribution which means that the t-test was a valid test of statistical significance, regardless of whether our data satisfies the necessary assumptions.

In hindsight, it seems that this may have been an unnecessarily cautious exercise since we were dealing with lists of data sufficiently large ($N \sim 50$) to satisfy the conditions for a t-test. However, this was a productive learning experience, and it provided an additional benefit. In the Fall 2020 course at CU, we noticed that the outward directed centralities correlated (significantly) with performance, but the inward

Figure 5.3: Sampling Distributions from Statistical Models



(a) Histogram of correlation between inward directed harmonic centrality in randomly generated networks and course grades



(b) Histogram of correlation between outward directed harmonic centrality in randomly generated networks and course grades directed centralities did not. One obvious possibility for this asymmetry was an asymmetry in how the CU students reported their collaborators. Generally, the CU students were more likely to report getting help than giving help. Naively, it could seem like this asymmetry in reporting could produce an asymmetry between the inward and outward centralties resulting in the asymmetry in the correlations.

In principle, however, this does not appear to be a sufficient explanation for the asymmetry in the correlations because, regardless of whether an edge was from a report of getting help or giving help, when looking at the network as a whole every edge contributes to the outward centrality of the node at which it terminates and contributes to the inward centrality of the node at which the edge originates. In the analysis of the sampling distributions generated from the collaboration models, sampling distributions for both the inward and outward centralities were created. Since the collaboration models were build to include the asymmetry in reporting, if the asymmetry in reporting was the cause for the asymmetry in the correlations, the asymmetry would appear in the sampling distributions in Fig 5.3. If the asymmetry in reporting could cause a difference between the inward and outward correlations, then the means of the sampling distributions in Fig 5.3 would not both be zero.

5.4 Effect of Removing Nodes

To investigate the effect of missing nodes within our networks, we simulated the removal of nodes from the Mines network and from networks generated using the DD method described in the previous section. This method was chosen since it roughly matched the distribution of edges seen in the real network, and came closest to matching the reciprocity of the real network. Nodes were missing from all networks, but were most prevalent in the fall 2020 course at CU Boulder which was missing nodes for 36% of the course's total enrolled population. In this analysis, we consider there to be a 'true,' 'complete' network which accurately and precisely represents student collaboration in a course. From the 'true' network we will remove nodes to make 'reduced' networks, then see how the correlation between centrality and grades is affected in the reduced networks.

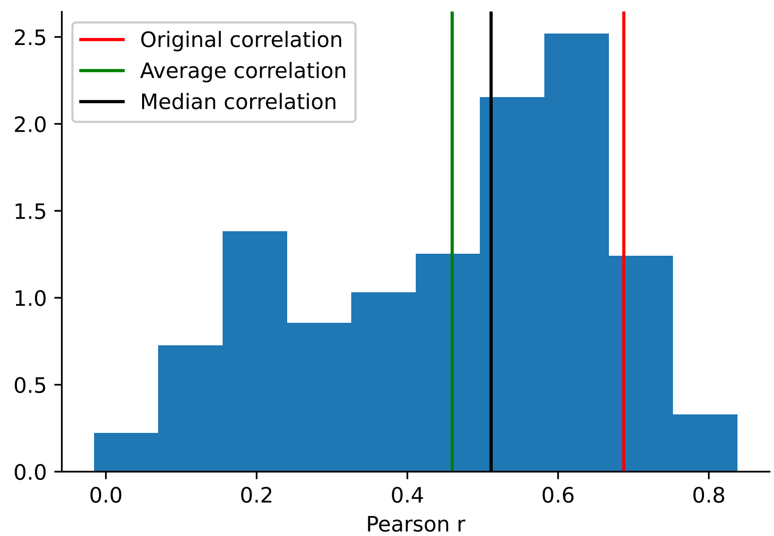
Five different methods for dropping students from the networks were tested. Each method used a different probability distribution to select students to drop from the network. Two probability distributions determining a node's likelihood to be dropped were created based on grades: one where high-performing students had a large probability of being dropped and another which gave low-performing students a higher likelihood of being removed. Two more distributions were created based on centrality: one in which students with high centrality were more likely to be dropped, and a second in which students with low centrality were more likely to be dropped. The final method removed students randomly, i.e., all students had the same probability of being removed.

Dropping nodes from the Mines network provides insight into the effect removing nodes from a 'complete' network which has a strong correlation between centrality and grades. For each of the five methods described above, 1000 sets of eight distinct nodes were selected for dropping from the network to match the proportion of missing nodes from the fall semester course at CU Boulder. The selected students were dropped to create reduced networks, and the centralities for the remaining nodes were recalculated. The new centralities were then correlated with the grades of the students remaining in the reduced network. The resulting distribution of correlations in the 1000 reduced networks are shown in Fig. 5.4 for the methods preferentially dropping students with low grades and low centrality. We chose to show these plots since we suspected that students with low grades or low centrality were less likely to participate in the study, report their collaborators, and more likely to be disconnected from the larger network.

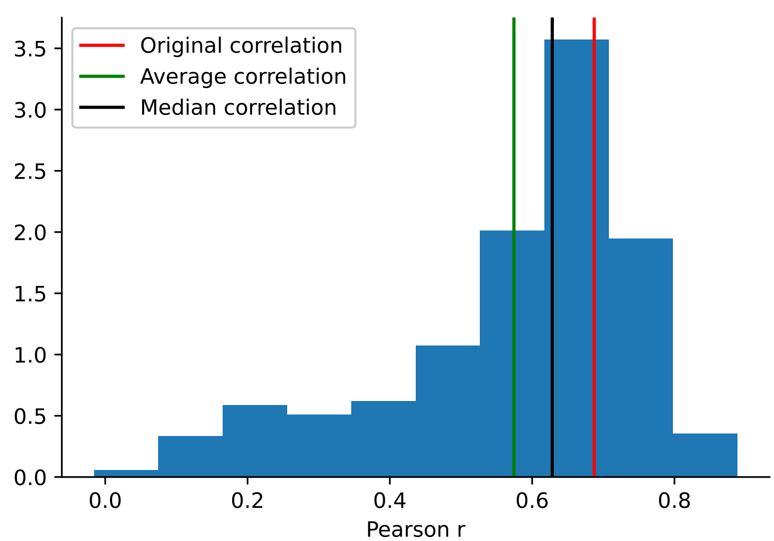
The red lines in each plot represent the correlation between grades and inward harmonic centrality measured in the original, unaltered Mines network. The green and black lines represent the mean and median correlation of the ensemble of reduced networks. The blue bars represent the histogram of correlations calculated from the networks in the ensemble.

As can be seen in the histograms in Fig. 5.4, very few reduced networks had a correlation larger than the correlation measured in the complete network (i.e., there are relatively fewer counts above the red lines in the histograms). Furthermore, the median correlation splits the distribution of correlations in half, so when removing nodes from the Mines network there is an equal probability of getting a correlation above or below the median. The median correlation among the reduced networks falls below the original correlation, indicating that removing nodes from a complete network with a strong correlation between grades and centrality tends to decrease the correlation measured in a reduced network. Removing nodes using the other

Figure 5.4: Histogram of Correlations in Reduced Networks



(a) Histogram of correlations between grades and inward harmonic centrality, c_i^{Hi}, in reduced networks created by preferentially dropping nodes with low grades from the complete Mines network.



(b) Histogram of correlations between grades and inward harmonic centrality, c_i^{Hi}, in reduced networks created by preferentially dropping nodes with low centrality from the complete Mines network.

three methods (high grades, high centrality, and randomly) all produced similar results as seen in Fig. 5.4.

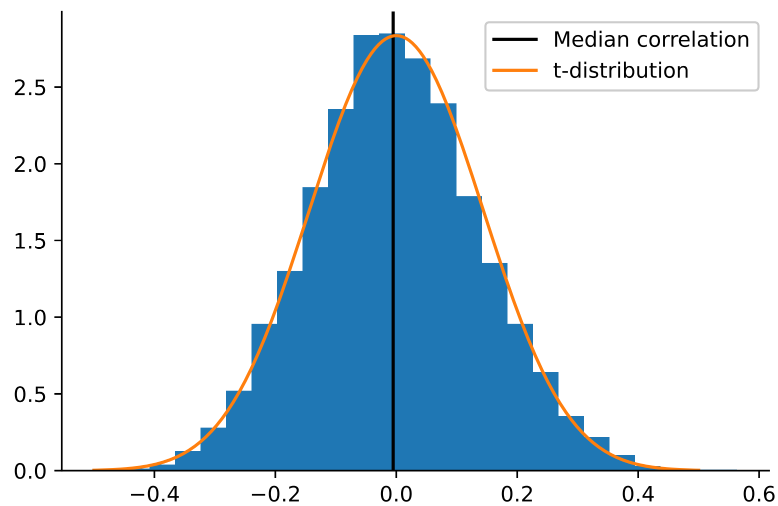
To extend the analysis of the effect of missing nodes, a similar dropping process was performed on random networks generated using the DD method described in the previous section. To create the probability distributions for dropping nodes by grades, the homework scores from the fall 2020 CU Boulder thermal physics course were applied randomly to the nodes of the simulated network. This process of randomly assigning grades to nodes resulted in complete networks that typically did not have significant correlations between grades and centrality (specifically, the distribution of correlations closely fit to a t-distribution centered at zero, as would be naively expected). So, while dropping nodes from the Mines network helped to show what happens when nodes are removed from networks with statistically significant correlations between grades and centrality, removing nodes from 'complete' networks with out a significant correlations helps show the likelihood of us finding a spurious correlation in the 'reduced' network.

Two approaches were taken to study the effect of missing nodes from these simulated networks. First, a single network was constructed (with 83 nodes), then 30 nodes were dropped, and correlations re-calculated to directly mimic the loss of nodes from the fall 2020 course at CU Boulder. This approach was applied iteratively and allowed for a large number of simulated networks to be analyzed. The distribution of resulting correlations in the reduced networks (after dropping the most and least central student) are shown in Fig. 5.5. The average and original correlations are not shown (as in Fig. 5.4) since they were all nearly equal to zero. The distribution of correlations in the reduced networks also fit very well to a t-distribution centered on zero, suggesting that dropping students does not tend to produce a net shift the correlation measured in the reduced networks.

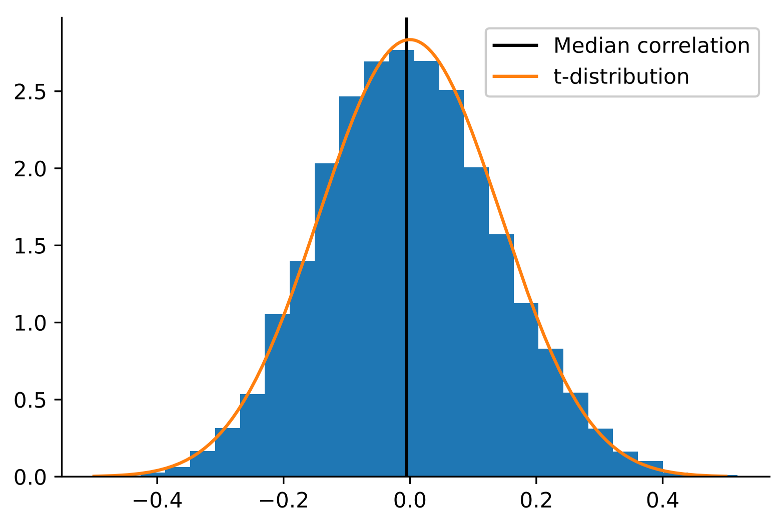
The second approach more closely matched the process applied to the Mines network. In this approach, a single network (of 83 nodes) was constructed, but instead of choosing only one set of nodes to drop, 500 sets of 30 nodes were produced and each set was dropped from the complete network to create 500 different reduced networks. New correlations were then calculated in all of the reduced networks before repeating the process with a new simulated complete network. This process repeated for 500 random 'complete' networks. This method allowed for a more detailed perspective on dropping nodes, but fewer networks could be analyzed with this process. All five methods for selecting nodes to drop were applied in this analysis and in each case the result was similar: dropping nodes tended to shift the correlation measured in the reduced network towards zero, and the proportion of complete networks with a non-significant correlation that became reduced network with a significant correlation when dropping nodes was less than three percent.

When taken together, the results from dropping nodes from the Mines and simulated networks suggests that the statistically significant correlations between homework scores and network centrality measured in the fall 2020 thermal physics course at CU Boulder likely are not spurious results caused by missing data.

Figure 5.5: Histogram of Correlations in Reduced, Simulated Networks



(a) Histogram of correlations between grades and inward harmonic centrality, c_i^{Hi}, in networks simulated with the DD method and 36% of nodes removed based on the nodes' centrality. Nodes with high centrality had greater probability to be dropped.



(b) Histogram of correlations between grades and inward harmonic centrality, c_i^{Hi}, in networks simulated with the DD method and 36% of nodes removed based on the nodes' centrality. Nodes with low centrality had greater probability to be dropped.

Chapter 6

Conclusion

6.1 Student Reasoning with Entropy

Conducting think-aloud interviews with graduate and undergraduate students to explore their conceptualizations and understanding of entropy allowed for a unique opportunity to examine and compare student reasoning across a wider cross-section of populations. The interview consisted of conceptual questions about ideal gases mixing across a semi-permeable partition, two solids exchanging energy, a system of a string waving in a water bath, and expansions of an ideal gas. Additionally, there was a final wrap-up question directly asking students about their conceptualizations of entropy. The interview questions were designed to elicit reasoning about entropy from both macroscopic and microscopic perspectives. Instructors are invited to use the string system as a tool for discussing entropy.

Two of the questions (the Blocks and Expansions) have appeared in previous studies investigating student reasoning about thermal physics concepts. Including these questions in our interviews and conducting interviews with both graduate and undergraduate students permits for a wide analysis across many perspectives by comparing between undergraduate reasoning between the current and previous research, and by comparing undergraduate to graduate student reasoning. The interviews were analyzed using the conceptual resources framework introduced by Hammer [19], and we corroborated resources found in prior studies and identified new resources used by students when reasoning about entropy.

In the interviews, both graduate and undergraduate students frequently related entropy to the multiplicity of microstates. This resource has also been observed in a prior study by Loverude [3] and is consistent with the results of another study which noted that students had a preference to think about entropy from a microscopic perspective [4]. In our study, we observed students thinking about microstates when reasoning about entropy in both the Blocks and Expansions questions which were intended to be addressed from a macroscopic perspective.

Throughout the interviews students also made a strong association between entropy and the temperature of a system. This resource generally helped students in the Blocks question when thinking about how the entropy of the separate objects changed as energy was exchanged, but it also caused some issues if students made incorrect assumptions about how the temperature of an object changed (such as in the Expansions question). We discuss situations in which this resource is valid as well as a specific situation in which this resource fails in Appendix B.

In the Partitioned Box and Strings questions, students commonly used the amount of mixing as a proxy to think about the entropy, specifically, students would commonly state that entropy would be maximized when things were the most mixed. 'Mixing' is not a well-defined idea within the context of thermal physics, so it is not very surprising that responses given by students invoking the idea of mixing were not consistent from student to student. In the undergraduate interviews in particular, it seemed like some students treated mixing as a process that generates extra entropy, more than is actually created by the equivalent process of allowing multiple substances to independently increase their volume.

Other interesting resources used by students included thinking of work as the opposite of entropy (or something that could decrease entropy) and entropy as a force or driver of processes. Somewhat surprisingly, students did not frequently mention 'disorder' when talking about entropy. Graduate students rarely invoked the term when reasoning through the conceptual problems (though brought it up more frequently on the final question of the interview). And though the undergraduates referenced the term more frequently, it was often accompanied with an acknowledgement that disorder was not the best way to think about entropy.

Specifically for thermal physics instructors, we also discussed some of what we see as broader implications and things to take away from this research project. This includes a discussion about disambiguating terms like 'system,' 'environment,' and 'universe' since students have been observed, both in this and prior studies [4, 15, 23], struggling to disentangle and distinguish these terms. A technical discussion of the concept was also included to help frame conversations about what entropy is and how it is different and similar to other physical quantities.

Future work on this topic could extend the research on student reasoning specifically about statistical mechanics concepts. As noted in the review by Dreyfus *et al.*, very little work has explored students' understanding of free energy, enthalpy, osmosis, diffusion, or how students understand the connections between probability and entropy [14]. This particular project was largely self-contained, but a final step could be producing and validating tutorial activities based on the string, particles, and water bath system. Tutorials based on this system could be developed that cover diffusion since the system was inspired from research on active diffusion.

6.2 Network Analysis

Performing social network analysis on self-reported student collaboration in physics courses affected by the COVID-19 pandemic showed that the correlation between a student's connectivity to their peers and their performance depended heavily on the context of the class. The 'on-sequence' iteration of the upper-division thermal physics course at CU Boulder (occurring in the fall semester of 2020) had statistically significant correlations between students' homework scores and their out-strength (a measure of how much help was given to others by a student) and outward-directed closeness and harmonic centralities (measures of a student's degree of separation from their peers). This indicates that students who not only gave help frequently, but were also a well-connected source of help tended to perform better in the course.

On the other hand, the subsequent iteration of the thermal physics course, occurring in the spring of 2021, had nearly no correlations between student performance and centrality. The only statistically significant correlation was a negative correlation between students' net-strength and their exams scores. Though this appears somewhat consistent with the results from the previous iteration of the course, since the net-strength is a measure of the difference between help received and help given, we are somewhat skeptical of this result since every other correlation calculated from this course was not statistically different from zero.

The results from an intermediate math methods course at Mines closely matched the results from the pre-pandemic study: there were statistically significant correlations between students' in-strength, out-strength, closeness centrality, and harmonic centrality which indicates that students who collaborate frequently and are closely connected to their peers tend to get higher grades on homework assignments. The consistency between the results found in this study and the pre-pandemic results from Vargas *et al.* show that the environment at Mines was able to preserve a connection between collaboration and course performance despite a fully remote course format. When contrasted with the lack of consistency in the results from hybrid courses at CU Boulder, this study demonstrates that course context is important for creating a connection between student collaboration and performance. For example, the on-sequence iteration of the thermal physics course at CU Boulder did show some correlation between collaboration and performance, but the off-sequence iteration showed nearly no correlation. Determining the relevant differences in contexts that results in significant correlations will require future research.

One of the primary limitations to this study is measurement error in the data collection process. This was specifically a concern for the fall 2020 course at CU Boulder which was missing roughly 36% of nodes representing student who did not consent to the data collection process. To address possible effects caused by missing data, we constructed several models for generating networks which resulted in one which was able to match the edge distribution and reciprocity seen in the measured networks. The simulated networks assumed that the level of collaboration in the missing part of the network matched the level seen in the measured networks. Simulations of random networks with this method corroborated the results of the t-test which established the statistical significance of our findings.

The development of the network simulation models also allowed for an investigation of the effect that removing nodes from larger network has on the correlations between centrality and course performance. 'Complete' networks of 83 nodes were created and sets of 30 nodes were selected for removal by various metrics. In all cases, the removal of nodes tended to shift correlations towards zero and resulted in a non-significant correlation (in the complete network) to shift to a significant correlation (in the reduced network) in less than three percent of trials. A similar approach to dropping nodes was applied to the measured network from the math methods course at Mines, which produced similar results: a tendency to decrease correlations. These results suggest that the statistically significant correlations measured in the fall 2020 course at CU Boulder reflect statistically significant correlations in the complete network.

Though prior research in education shows that collaborative interaction among students generally leads to better learning outcomes (at least in part due to developing a student's sense of belonging), lacking a correlation between student performance and centrality should not be considered a necessarily undesirable feature. In the case of commuting students, students who are working while attending school, recently transferred students, or other students less able to collaborate with their peers, an ideal course would overcome the obstacles faced by these students' disconnection from the course and result in learning outcomes not dependent on a students' ability to interact with their classmates. Further research on the utility of high centrality in the collaboration network of a course will benefit from a validated assessment of students' thermal physics understanding. This tool will help identify whether there are differences in learning gains between courses of well-connected versus more disconnected students.

Additionally, to explore the differences between the networks at Mines and CU Boulder (and to what extent the differences can be explained by cultural context), qualitative interviews with students to investigate their beliefs and conceptions of group work should be conducted. In particular, understanding students' thresholds for reporting helping or being helped by other students, the range of interactions that students' have with their collaborators, and whether or not students find collaboration to be helpful will provide deeper insight into our results. Students' qualitative responses may shed light on what aspects of a course help to create a correlation between collaboration and performance. Furthermore, more data collection on in-person courses at CU Boulder would establish whether collaboration is associated with performance during regular, in-person instruction, or whether the environment at CU Boulder generally tends to have a weaker connection between student collaboration and performance. Finally, it may be worthwhile to study whether there is a causal relationship between centrality and performance.

6.3 Final Thoughts

The work done on this thesis adds to the relatively small body of work examining upper-division thermal physics courses. Additionally it adds to the nearly non-existent body of research examining graduate student learning. In particular, this research contributes the string-and-water-bath system as new way of introducing students to and discussing the concept of entropy. Furthermore, we identified a number of conceptual resources commonly utilized by students. We observed a strong preference among students to conceptualize entropy from a microscopic perspective by frequently invoking the idea that entropy is related to the multiplicity of microstates. Other prominent conceptualizations of entropy included associations between entropy and temperature and between entropy and mixing.

This thesis also adds to a growing body of work that applies network analysis to study student learning and outcomes. We extended a study that analyzed the correlations between students' collaborations and grades at the Colorado School of Mines to explore courses at CU Boulder and provide a comparison between courses occurring before and during the COVID-19 pandemic. We found that the existence of correlations between collaboration and performance appeared to depend on the context of the course. Additionally, this study makes a unique contribution to the analysis of errors within social networks. In our specific context, we found that missing data tends to decrease correlations between centrality and student grades. Since this analysis was specifically tailored to address our study, we encourage others in the PER field to investigate possible effects of measurement errors in their research involving social networks.

Bibliography

[1] Daniel Schroeder. An Introduction to Thermal Physics. Addison-Wesley, New York, 2000.

[2] Katherine D. Rainey, Michael Vignal, and Bethany R. Wilcox. Designing upper-division thermal physics assessment items informed by faculty perspectives of key content coverage. Phys. Rev. Phys. Educ. Res., 16:020113, Aug 2020.

[3] Michael Loverude. Identifying student resources in reasoning about entropy and the approach to thermal equilibrium. Phys. Rev. ST Phys. Educ. Res., 11:020118, Sep 2015.

[4] Brandon R. Bucy, John R. Thompson, and Donald B. Mountcastle. What is entropy? advanced undergraduate performance comparing ideal gas processes. In AIP Conference Proceedings, volume 818, pages 81–84, 2006.

[5] Trevor I. Smith, Warren M. Christensen, Donald B. Mountcastle, and John R. Thompson. Identifying student difficulties with entropy, heat engines, and the carnot cycle. Phys. Rev. ST Phys. Educ. Res., 11:020116, Sep 2015.

[6] Risto Leinonen, Mervi A. Asikainen, and Pekka E. Hirvonen. Grasping the second law of thermodynamics at university: The consistency of macroscopic and microscopic explanations. Phys. Rev. ST Phys. Educ. Res., 11:020122, Sep 2015.

[7] Christian H. Kautz, Paula R. L. 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.

[8] L. H. Anderman and T. M. Freeman. Motivating Students, Improving Schools: The Legacy of Carol Midgley. Elsevier JAI, Amsterdam; London, 2004.

[9] Mary J Fischer. Settling into campus life: Differences by race/ethnicity in college involvement and outcomes. Journal of Higher Education, 78:125, 2007.

[10] David L. Vargas, Ariel M. Bridgeman, David R. Schmidt, Patrick B. Kohl, Bethany R. Wilcox, and Lincoln D. Carr. Correlation between student collaboration network centrality and academic performance. Phys. Rev. Phys. Educ. Res., 14:020112, Oct 2018.

[11] Nathan Crossette, Michael Vignal, and Bethany R. Wilcox. Investigating graduate student reasoning on a conceptual entropy questionnaire. Phys. Rev. Phys. Educ. Res., 17:020119, Sep 2021.

[12] Nathan Crossette, Michael Vignal, and Bethany Wilcox. Investigating how graduate students connect microstates and macrostates with entropy. In Physics Education Research Conference 2020, pages 98–103, 2020.

[13] Nathan Crossette, Michael Vignal, and Bethany Wilcox. Comparing undergraduate and graduate student reasoning on a conceptual entropy questionnaire. In Physics Education Research Conference 2021, pages 105–110, 2021.

[14] 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.

[15] Warren M. Christensen, David E. Meltzer, and C. A. 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.

[16] Christian H. Kautz, Paula R. L. 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.

[17] Michael Loverude. Investigating student understanding for a statistical analysis of two thermally interacting solids. In Physics Education Research Conference 2010, pages 213–216, 2010.

[18] Marsha E. Fonteyn, Benjamin Kuipers, and Susan J. Grobe. A description of think aloud method and protocol analysis. Qualitative Health Research, 3(4):430–441, 1993.

[19] David Hammer. Student resources for learning introductory physics. American Journal of Physics, 68(S1):S52–S59, 2000.

[20] Benjamin D. Geller, Benjamin W. Dreyfus, Julia Gouvea, Vashti Sawtelle, Chandra Turpen, and Edward F. Redish. Entropy and spontaneity in an introductory physics course for life science students. American Journal of Physics, 82(5):394–402, 2014.

[21] Brandon Bucy. PhD thesis, University of Maine, 2007.

[22] Paul W Irving and Eleanor C Sayre. Physics identity development: A snapshot of the stages of development of upper-level physics students. Journal of the Scholarship of Teaching and Learning, 13(4):68–84, Jun. 2013.

[23] E. Carson and J. Watson. Undergraduate students' understandings of entropy and gibbs free energy. University Chemistry Education, 6, 01 2002.

[24] Judith M. Bennett and Mustafa Sözbilir. A study of turkish chemistry undergraduates' understanding of entropy. Journal of Chemical Education, 84(7):1204, 2007.

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

[26] Catherine H. Crouch and Eric Mazur. Peer instruction: Ten years of experience and results. American Journal of Physics, 69(9):970–977, 2001.

[27] Justyna P. Zwolak, Remy Dou, Eric A. Williams, and Eric Brewe. Students' network integration as a predictor of persistence in introductory physics courses. Phys. Rev. Phys. Educ. Res., 13:010113, Mar 2017.

[28] Justyna P. Zwolak, Michael Zwolak, and Eric Brewe. Educational commitment and social networking: The power of informal networks. Phys. Rev. Phys. Educ. Res., 14:010131, May 2018.

[29] Remy Dou, Eric Brewe, Justyna P. Zwolak, Geoff Potvin, Eric A. Williams, and Laird H. Kramer. Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens. Phys. Rev. Phys. Educ. Res., 12:020124, Aug 2016.

[30] Eric Brewe, Laird Kramer, and Vashti Sawtelle. Investigating student communities with network analysis of interactions in a physics learning center. Phys. Rev. ST Phys. Educ. Res., 8:010101, Jan 2012.

[31] Eric Brewe, Jesper Bruun, and Ian G. Bearden. Using module analysis for multiple choice responses: A new method applied to force concept inventory data. Phys. Rev. Phys. Educ. Res., 12:020131, Sep 2016.

[32] James Wells, Rachel Henderson, Adrienne Traxler, Paul Miller, and John Stewart. Exploring the structure of misconceptions in the force and motion conceptual evaluation with modified module analysis. Phys. Rev. Phys. Educ. Res., 16:010121, Apr 2020.

[33] Christopher Wheatley, James Wells, Rachel Henderson, and John Stewart. Applying module analysis to the conceptual survey of electricity and magnetism. Phys. Rev. Phys. Educ. Res., 17:010102, Jan 2021.

[34] Eric A. Williams, Justyna P. Zwolak, Remy Dou, and Eric Brewe. Linking engagement and performance: The social network analysis perspective. Phys. Rev. Phys. Educ. Res., 15:020150, Dec 2019.

[35] M. Newman. Networks. Oxford University Press, New York, 2018.

[36] V. Latora, V Nicosia, and G. Russo. Complex Networks Principles, Methods and Applications. Cambridge University Press, 2017.

[37] Stanley Wasserman and Katherine Faust. Social network analysis: Methods and applications. Cambridge University Press, 1994.

[38] Ulrik Brandes. On variants of shortest-path betweenness centrality and their generic computation. Social Networks, 30(2):136–145, 2008.

[39] Tiziano Squartini, Francesco Picciolo, Franco Ruzzenenti, and Diego Garlaschelli. Reciprocity of weighted networks. Scientific Reports, 3(1):2729, Sep 2013.

[40] Diego Garlaschelli, Franco Ruzzenenti, and Riccardo Basosi. Complex networks and symmetry i: A review. Symmetry, 2(3):1683–1709, 2010.

[41] Christoph Martin and Peter Niemeyer. Influence of measurement errors on networks: Estimating the robustness of centrality measures. Network Science, 7(2):180, 2019.

[42] Mark Newman. The physics of networks. Physics Today, 61(11):33–38, 2008.

Appendix A

Entropy Interview Questionnaire

A.1 Introduction & Notes on Questionnaire

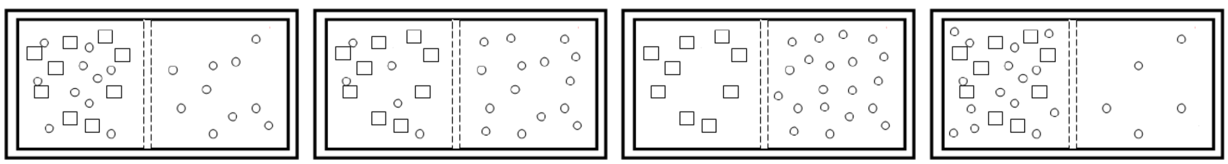
The questionnaire that students saw in interviews is in the following section. In the actual document seen by students in interviews, the margins were slightly smaller, the questions were single spaced, and twice as much space was given for students to show work, write answers, and physically separate the questions.

As an additional note, the wording of the third question (the strings question) was changed after the interviews with graduates students and before the interviews with undergraduates. In the undergraduate interviews, a reference to Brownian motion as the mechanism causing the string to move in the water bath and a statement to treat the water as a continuous medium were removed as the contradictory nature of the two statements sometimes lead to confusion among the graduate students. Also, a direction to treat the system as only two-dimensional was added for the undergraduate interviews. Both statements of the question appear in the following section and are labeled as "Graduates" and "Undergraduates" according to the group of students in which the question statement was used.

A.2 Questionnaire

1. Each of the four following images represent the same system: a box with two compartments separated by an immovable, semi-permeable partition through which only the circle particles can pass. The box is thermally isolated from the outer environment. The number of each particle species (squares and circles) is the same in each image. Both circle and square particles can be treated as ideal gasses.

Figure A.1: Partitioned box states



(a) 10 circles : 10 circles (b) 5 circles : 15 circles (c) 0 circles : 20 circles (d) 15 circles : 5 circles

a) Rate the states based on their entropy. Which has the highest?

b) Which of these pictures most closely represents the equilibrium state of the system? How will each state change with time?

c) For each of the four states above, which side (left or right) is at a higher pressure?

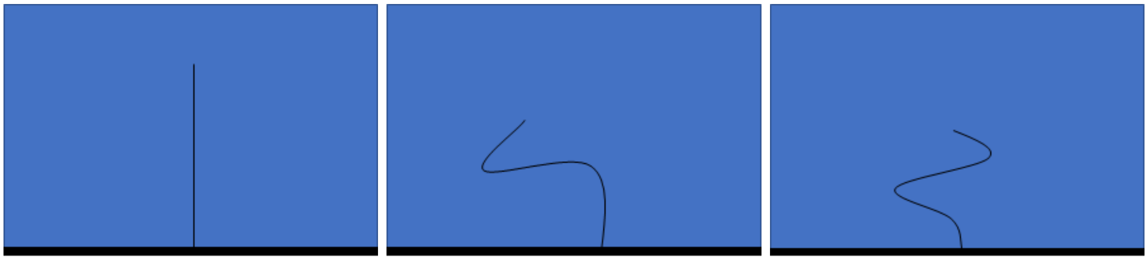
d) Is your answer to the previous question consistent to any claims made in part b about equilibrium? Is anything maximized, minimized, or equilibrated at equilibrium? Please elaborate.

2. Two solids (A and B) are placed in thermal contact and isolated from the rest of the universe. Before being placed in thermal contact, solid A had a higher temperature than solid B. Describe what will happen to the blocks. What can you say about the individual entropies of the solids before, and a long time after they come into thermal contact?

3. **(Graduates)** The figures below show snapshots of a thin string attached to a wall inside a bath of water. The density of the string is the same as that of water, so the gravitational force of the string is completely canceled by the buoyancy force from the water (i.e., don't worry about net external forces). Brownian motion of the water can change the "conformation" (meaning 'position' or 'arrangement') of the string, but for the purposes of the following questions, consider the water as a continuous medium (don't worry about the water being made up of individual molecules).

3. **(Undergraduates)** The figures below show snapshots of three "conformations" (meaning 'positions' or 'arrangements') of a thin string waving in 2-dimensions inside a bath of water (you can neglect the dimension that goes into/out of the page). The string is attached to the wall, and the density of the string is the same as that of water, so the gravitational force of the string is completely canceled by the buoyancy force from the water (so, don't worry about net external forces).

Figure A.2: String microstates



(a) (b) (c)

a) Based on your intuition, rank the probabilities of finding the string in each of the three conformations shown above.

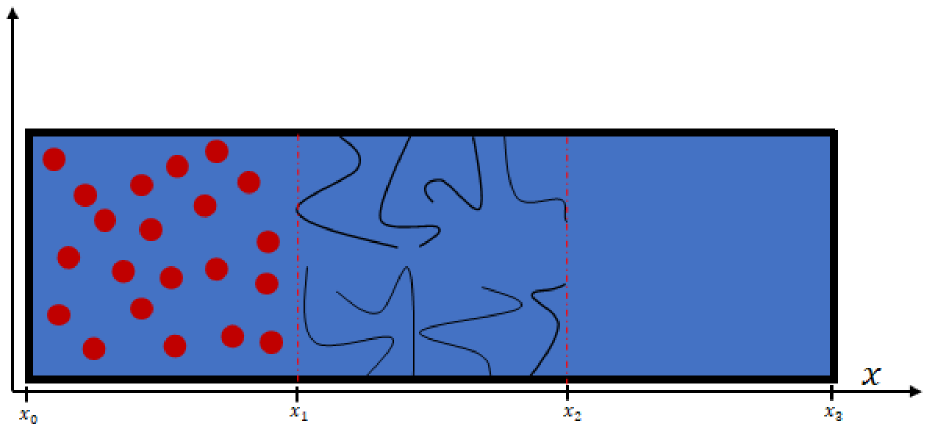
b) Is there a property of the string that can be used to define a set of distinct macrostates of the string? Are the conformations shown in the figure macrostates or microstates?

c) Based on your answer to part b, how would you rank the probabilities of finding the strand in each of the three conformations above?

d) Can you discuss what is meant by the "entropy" of the string, and how it relates to the possible conformations of the string?

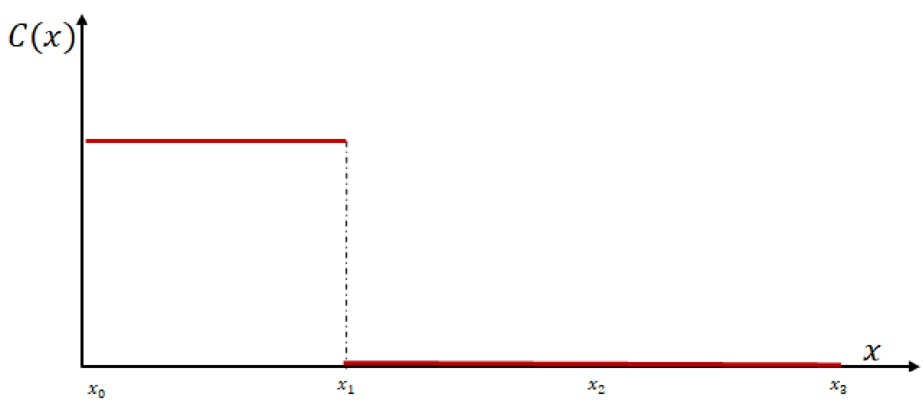
e) The figure below shows a channel with strings attached to a section of the walls of the channel.

Figure A.3: Channel with strings.



The channel is filled with water and the circles represent particles suspended in the water. The strings and circles can move in the 2D plane of the page. The third dimension going into and out of the page can be ignored. The strings and molecules do not interact through any long range attraction or repulsion, but can collide with one another. The molecules begin at a high concentration on the left, and the initial concentration profile, $C(x)$, is shown below.

Figure A.4: Channel concentration profile

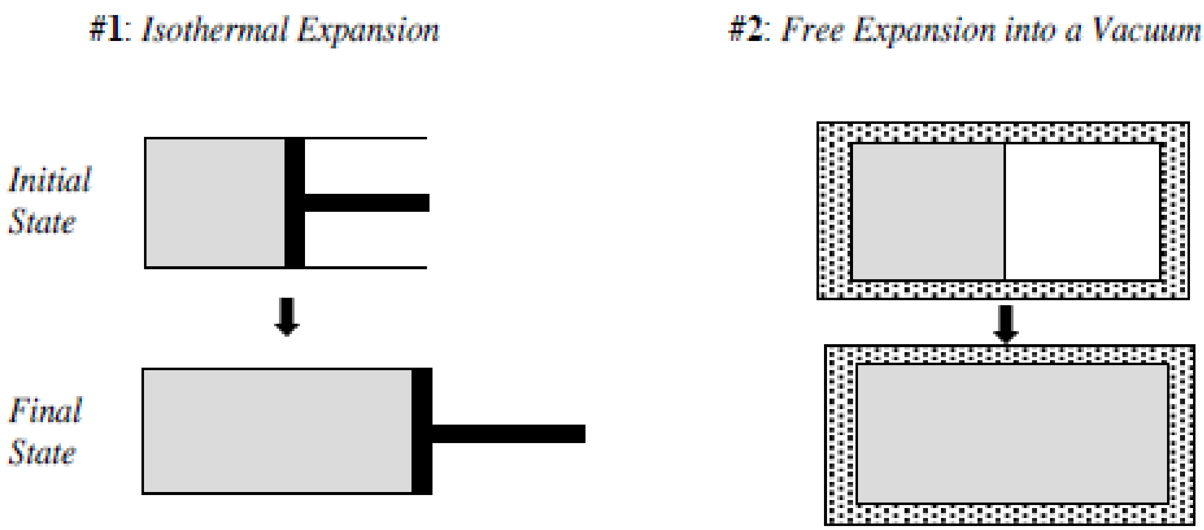


i) What happens to the number of possible conformations of the strings when the molecules enter the region with the strings?

ii) What will the concentration profile, as a function of the distance along the channel, of the molecules be after a long period of time? Explain your reasoning.

4. The figure below shows an expansion of two identical, ideal, monatomic gasses. Each gas begins at the same volume, pressure, and temperature, and expand to the same final volume. The first gas undergoes a quasistatic, isothermal expansion while in contact with a large reservoir. The second gas undergoes a free expansion into a vacuum in a container that is thermally isolated from the outer environment.

Figure A.5: Expansions



a) Is the change in entropy of gas #1 positive, negative, or zero? Explain your reasoning.

b) Is the change in entropy of gas #2 positive, negative, or zero? Explain your reasoning.

c) How does the change in entropy of gas #1 compare to the change in entropy of gas #2? Explain your reasoning.

d) How does the change in entropy of the environment outside of gas #1 compare with the change in entropy of the environment outside of gas #2? Explain your reasoning.

5. In your own words what is entropy? How do you, personally, conceptualize entropy?

Appendix B

Discussion on the relationship between temperature and entropy

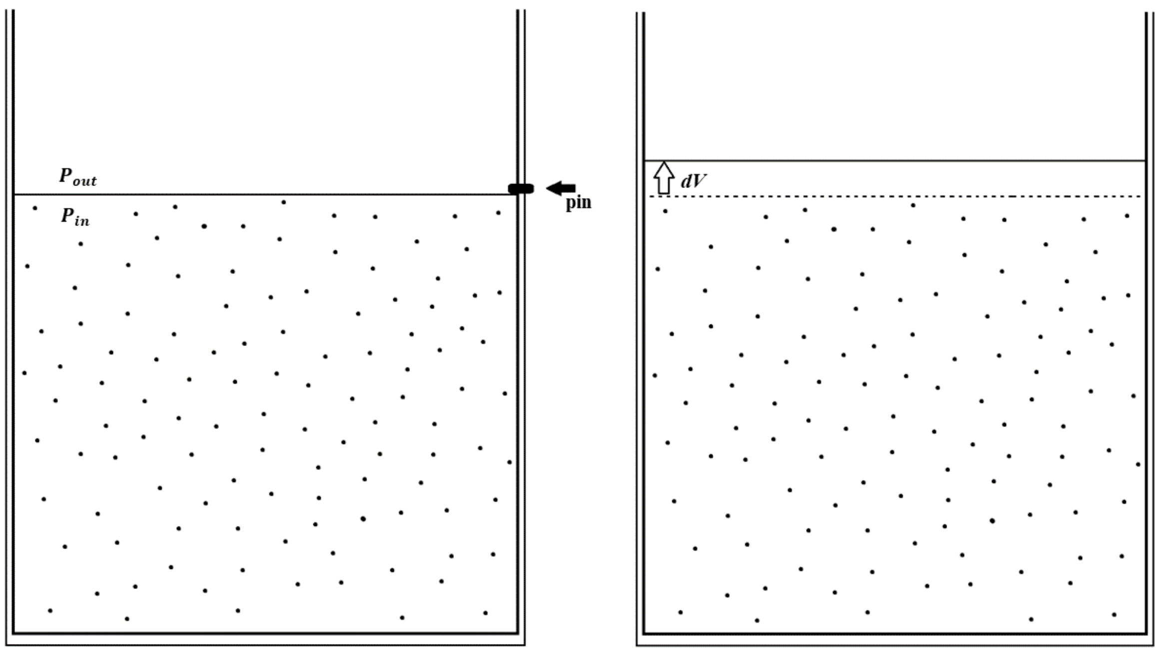
In our think-aloud interviews, we observed many students using a resource that associated increases in temperature with increases in entropy, as if entropy was a strictly monotonically increasing function of temperature. In many cases, such an intuition is correct and valid reasoning, yet no such general mathematical relationship exists between temperature and entropy in the context of thermal physics. The closest equations are the Sackur-Tetrode equation for ideal gasses and the high/low temperature limiting cases for the entropy of an Einstein solid, all of which relate entropy to *energy*, not temperature. However, energy and temperature are directly related for ideal gasses, and the energy as a function of temperature for Einstein solids takes the form:

$$U(T) = \varepsilon N \left(\frac{1}{2} + \frac{1}{e^{\varepsilon/kT} - 1} \right)$$

where N is the number of oscillators and ε is the harmonic oscillator energy level spacing. Since this is a monotonically increasing function of temperature and the entropy equations for Einstein solids are monotonically increasing functions of energy (within their realm of applicability), entropy of an Einstein solid (and an ideal gas) *must be* a monotonically increasing function of temperature *when all other thermodynamic variables are held constant.*

This is where the resource can break down. Thermodynamic systems, like ideal gases, typically have multiple thermodynamic variables that can change independently of each other, so an increase in entropy due to an increase in temperature can be offset by a change in some other thermodynamic variable. In the following proof, we will layout a process in which temperature decreases and entropy increases.

Figure B.1: A Non-Quasistatic, Spontaneous Expansion of a Gas



A hot gas at a higher pressure than the outer surroundings inside a piston. A pin holds the piston in place. When the pin is removed, the gas begins a spontaneous expansion. An infinitesimal change in volume is shown on the right half of the figure.

Proof. Consider a gas inside a sealed, thermally insulated piston like the system depicted in Fig. B.1. The pressure inside the piston is much greater than the pressure outside the piston ($P_{in} >>> P_{out}$). Initially, the top of the piston is held in place by a pin. We will consider two processes. First, the pin will be removed and the gas will be allowed to expand quasi-statically from the initial volume, V_i, to a final volume of $V_i + dV$. In the second, the pin will be removed and the gas will expand non-quasistatically from V_i to the same final volume of $V_i + dV$.

In the quasi-static expansion, we can invoke the thermodynamic identity to consider the change in entropy:

$$
\begin{aligned}
dU &= TdS - PdV \\
dS &= \frac{dU}{T} + \frac{P}{T}dV
\end{aligned}
\tag{B.1}
$$

In this form, we have the change in entropy in terms of the change in energy dU and the change in volume, dV. We want to consider the relationship between changes in entropy and changes in temperature, so we need to replace the dU term with a dT term. From the equipartition theorem ($U = \frac{3}{2}Nk_bT$), we have a direct relationship between the energy and temperature of the gas. This will let us to change variables by replacing the dU's with dT's:

$$dU = \frac{3}{2}Nk_b dT \tag{B.2}$$

Substituting this in to Eq. B.1 and cleaning up:

$$dS = \frac{3}{2}\frac{Nk_b}{T}dT + \frac{P}{T}dV \tag{B.3}$$

Since the gas is insulated, we know that the heat flow, Q, is zero, so only work being done on the piston will change the energy of the gas. We can use this to relate dU, and by extension dT, to the change in volume dV:

$$\begin{aligned} dU &= \cancelto{0}{Q} + W \\ dU &= -PdV \\ \frac{3}{2}Nk_b dT &= -PdV \end{aligned} \tag{B.4}$$

Use this relationship to substitute into the first term of Eq. B.3 and we will see that the change in entropy of the gas is zero:

$$\begin{aligned} dS &= -\frac{P}{T}dV + \frac{P}{T}dV \\ dS &= 0 \end{aligned} \tag{B.5}$$

This result should not be a surprise, since all quasi-static, adiabatic processes are isentropic. In this case, temperature decreased, but entropy remained constant which *is* consistent with a monotonic relationship between temperature and entropy. Now let's consider the non-quasistatic expansion.

We have chosen a situation in which $P_{in} >>> P_{out}$, so that when the pin is removed the top of the piston will accelerate quickly (to a speed faster than the speed of sound inside the piston) and create a layer of pressure, just beneath the top of the piston, lower than the average pressure of the gas. This region is depicted on the right side, between the dashed and solid lines, in Fig. B.1.

This region of lower pressure will cause the gas to do *less* work than the PdV from the quasi-static case. This can be understood microscopically as fewer gas particles colliding with the piston and, overall, losing less energy than in the quasi-static case.

In the two processes, the gas expands to the same final volume, but ends up with a higher internal energy (and therefore, temperature) in the non-quasistatic case. We can use the Sackur-Tetrode equation to compare the final entropies of the gas in the two processes.

$$\frac{S}{Nk_b} = \ln\left[\frac{V}{N}\left(\frac{4\pi m}{3h^3}\frac{U}{N}\right)^{3/2}\right] + 5/2 \tag{B.6}$$

Since the gas undergoing the non-quasistatic process ends at the same volume as the gas in the quasistatic process, but at a higher energy, we can see that $S_{f,nqs}$ is greater than $S_{f,qs}$. Also, the initial entropies in both cases are identical, because the two processes start in the same initial state. Therefore, ΔS_{nqs} must be greater than ΔS_{qs}, and since ΔS_{qs} was zero, ΔS_{nqs} must be positive.

Therefore, we have that in the non-quasistatic case temperature *decreases* and entropy *increases.*

■

This result can be understood qualitatively as well. When $P_{in} >> P_{out}$, the expansion of the gas is a spontaneous, non-reversible process meaning entropy must increase. The gas loses energy, therefore its temperature decreases. The increase in entropy due to the volume expanding, then, must be larger than the decrease in entropy due to the temperature decrease.

Somewhat ironically, while disproving a strictly monotonic relationship between temperature and entropy, the intuition that higher temperature means more entropy is vindicated to some extent. In the comparison of the final entropies of the gasses in the two processes, we found that the higher temperature gas (in the non-quasistatic expansion) had the larger entropy. This demonstrates that the association can be useful and productive, when it is applied correctly. The resource is valid when temperature is the *only* difference between two systems.

We hope that this discussion may prove to be useful to instructors and students looking to better teach and understand the subtleties of the connection between temperature, energy, and entropy. This example could serve as the basis for homework questions or lecture material. For an additional exercise, one can also show that in a non-quasistatic compression of a gas (the inverse of non-quasistatic expansion), entropy can *decrease* while temperature *increases.*

Appendix C

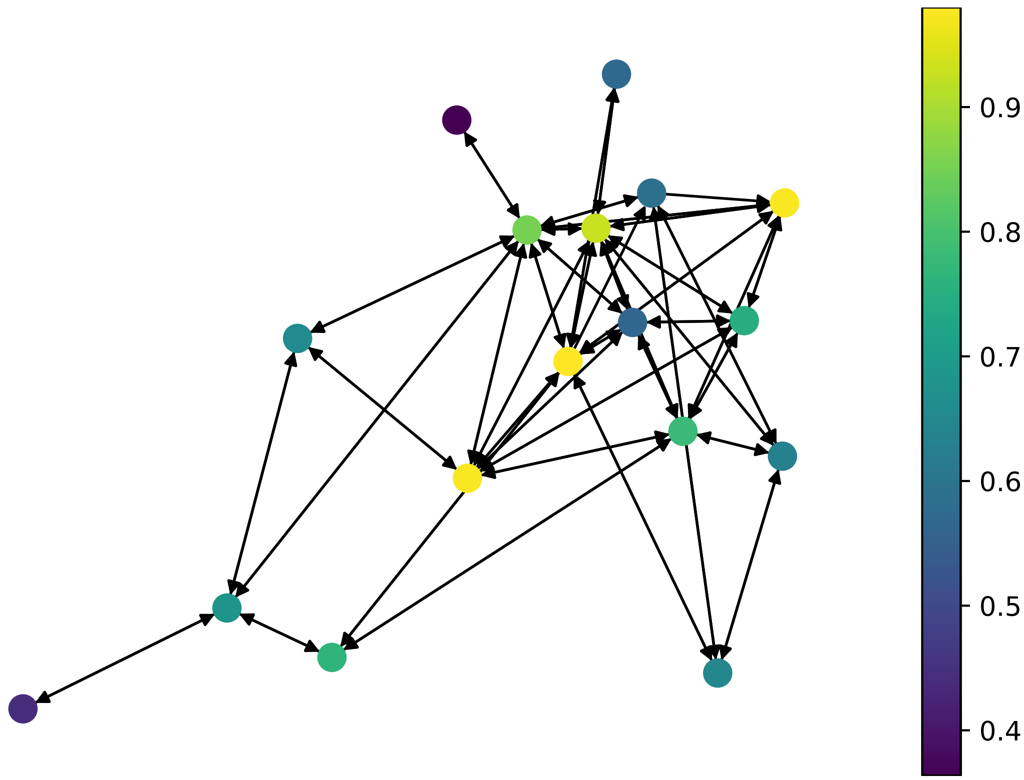
Additional Details on Social Network Data

This appendix contains tables (Tables C.1, C.2, and C.3) of centrality measures for the students in each of the three courses and graphical drawings of the collaboration networks (Figs. C.1, C.2, and C.3).

Table C.1: Network Data from Fall 2020 Math Methods Course at Mines

Student ID	s_i^{in}	s_i^{out}	s_i^{net}	c_i^{Co}	c_i^{Ci}	c_i^{Ho}	c_i^{Hi}	c^B	Y_i^{in}	Y_i^{out}
100000619	0	0	0	0	0	0	0	0	undef.	undef.
100001315	0	0	0	0	0	0	0	0	undef.	undef.
100001665	18	18	0	0.975	0.975	29.78	29.78	74.333	0.142	0.142
100001746	3	3	0	0.598	0.598	16.579	16.579	0	0.556	0.556
100002042	14	13	1	0.975	0.975	30.127	30.127	16	0.224	0.254
100002267	22	22	0	0.795	0.795	30.183	30.183	72.667	0.153	0.153
100004563	19	19	0	0.932	0.932	30.395	30.395	43	0.147	0.147
100004701	7	7	0	0.595	0.595	16.487	16.487	0	0.184	0.184
100004784	2	2	0	0.48	0.48	13.652	13.652	0	1	1
100005133	16	16	0	0.86	0.86	25.533	25.533	93.667	0.141	0.141
100005203	7	7	0	0.28	0.28	7	7	0	1	1
100005332	10	10	0	0.763	0.763	23.925	23.925	0	0.28	0.28
100005404	11	11	0	0.616	0.616	22.692	22.692	0	0.455	0.455
100005775	7	7	0	0.778	0.778	24.142	24.142	24	0.388	0.388
100006325	1	1	0	0.404	0.404	10.554	10.554	0	1	1
100006345	0	0	0	0	0	0	0	0	undef.	undef.
100006960	0	0	0	0	0	0	0	0	undef.	undef.
100007150	9	9	0	0.701	0.701	21.769	21.769	34	0.309	0.309
100007614	0	0	0	0	0	0	0	0	undef.	undef.
100008165	12	12	0	0.671	0.671	25.277	25.277	30	0.431	0.431
100008311	8	8	0	0.652	0.652	20.452	20.452	8	0.281	0.281
100008829	6	6	0	0.68	0.68	19.417	19.417	0	0.333	0.333
100008902	18	19	-1	0.98	0.98	30.515	30.515	79.333	0.16	0.147

Figure C.1: Mines Collaboration Network

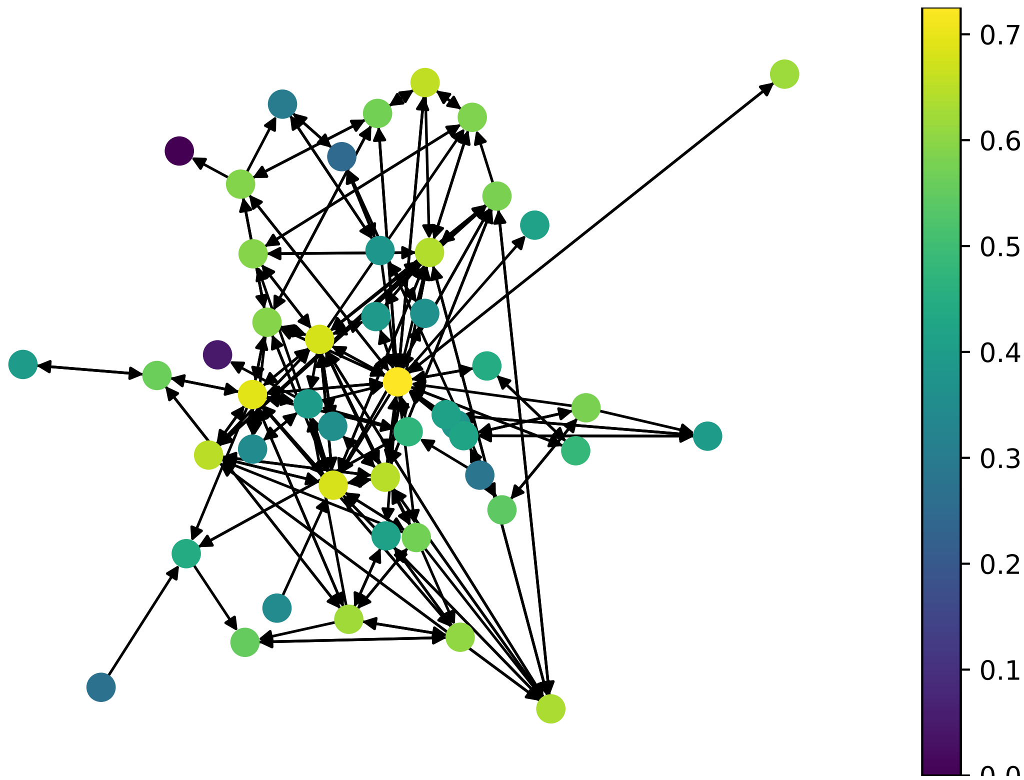


This network also contains six disconnected nodes, representing students for whom no collaboration was reported, which are not shown in this image. Nodes are colored based on their outward directed closeness centrality.

Table C.2: Network Data from Fall 2020 Thermal Physics Course at CU Boulder

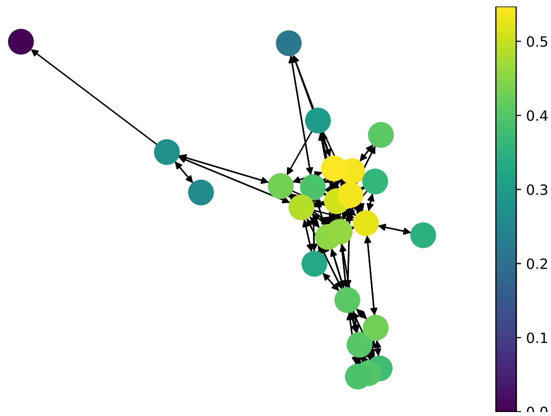
Student ID	s_i^{in}	s_i^{out}	s_i^{net}	c_i^{Co}	c_i^{Ci}	c_i^{Ho}	c_i^{Hi}	c^B	Y_i^{in}	Y_i^{out}
1011000080	21	20	1	0.593	0.604	67.475	70.543	114.833	0.147	0.175
1041052959	28	31	-3	0.694	0.754	90.052	92.336	1636.167	0.12	0.105
1041407859	15	5	10	0.398	0.679	38.5	75.991	340	0.138	0.28
1051126960	22	12	10	0.586	0.679	67.458	91.685	172	0.198	0.264
1051182994	13	13	0	0.482	0.608	56.307	70.911	121	0.538	0.527
1051335967	2	2	0	0.398	0.445	41.177	45.213	0	1	1
1051343194	2	3	-1	0.379	0.394	36.561	38.123	0	1	0.556
1051351288	8	5	3	0.248	0.297	25.569	35.906	20	0.438	0.44
1051413178	34	33	1	0.604	0.674	87.181	92.859	444	0.194	0.207
1051421707	61	51	10	0.634	0.622	107.443	116.448	0	0.139	0.148
1051427494	73	84	-11	0.679	0.629	130.161	124.453	231.5	0.122	0.103
1051433470	0	0	0	0	0	0	0	0	undef.	undef.
1051717375	1	0	1	0	0.348	0	32.127	0	1	undef.
1051872475	2	2	0	0.366	0.413	33.776	38.153	0	0.5	0.5
1061127931	3	3	0	0.389	0.422	36.887	40.374	0	0.333	0.333
1061421508	4	4	0	0.047	0.379	4	37.071	67	0.625	1
1061458033	64	63	1	0.643	0.617	109.874	111.842	243.5	0.107	0.105
1061682802	57	50	7	0.725	0.809	93.213	102.287	2255.167	0.055	0.069
1061850613	1	1	0	0.187	0.252	17.013	23.146	0	1	1
1061960044	5	7	-2	0.381	0.312	37.209	30.986	216	0.28	0.224
1061973208	34	33	1	0.564	0.624	88.283	92.512	240	0.272	0.286
1071222904	14	19	-5	0.572	0.582	70.404	77.993	20	0.255	0.136
1071261664	15	14	1	0.396	0.103	48.984	16.693	11	0.476	0.541
1071411427	8	7	1	0.302	0.384	31.025	39.928	235.5	0.25	0.306
1071432095	36	33	3	0.58	0.589	84.153	91.238	0	0.17	0.157
1071643222	60	61	-1	0.651	0.62	116.681	113.368	11	0.14	0.142
1071668161	13	12	1	0.444	0.495	54.027	58.014	57	0.728	0.847
1071674263	0	1	-1	0.273	0	25.383	0	0	undef.	1
1071692195	11	10	1	0.446	0.56	51.639	63.941	0	0.702	0.82
1071739846	4	2	2	0.347	0.417	32.66	39.221	0	0.25	0.5
1071780268	23	15	8	0.593	0.611	80.447	91.861	26	0.27	0.351
1071884161	15	16	-1	0.472	0.524	57.856	61.632	232	0.564	0.508
1071956551	66	67	-1	0.648	0.626	121.932	119.585	63.5	0.132	0.133
1071968359	12	12	0	0.618	0.693	74.347	82.205	121	0.5	0.5
1081058456	7	10	-3	0.589	0.611	61.759	66.012	171	0.388	0.24
1081172960	8	9	-1	0.422	0.116	46.755	11.907	18	0.469	0.383
1081189052	17	17	0	0.57	0.631	63.91	73.964	173.167	0.253	0.218
1081307369	13	10	3	0.554	0.625	72.815	80.39	0	0.728	1
1081345838	65	79	-14	0.681	0.665	127.853	117.678	950.833	0.114	0.104
1081401350	7	5	2	0.416	0.524	42.738	55.823	0	0.306	0.28
1081501160	8	5	3	0.418	0.544	43.895	58.177	63	0.469	1
1081519226	22	23	-1	0.656	0.742	85.681	96.932	323	0.335	0.308
1081618052	0	0	0	0	0	0	0	0	undef.	undef.
1081804766	0	0	0	0	0	0	0	0	undef.	undef.
1081950749	0	0	0	0	0	0	0	0	undef.	undef.
1091031818	0	1	-1	0.283	0	26.362	0	0	undef.	1
1091086505	0	0	0	0	0	0	0	0	undef.	undef.
1091278703	30	35	-5	0.62	0.685	92.796	94.849	455	0.211	0.182
1091524010	0	1	-1	0.345	0	32.891	0	0	undef.	1
1091558015	9	10	-1	0.414	0.087	48.605	14.556	53	1	0.82
1091862920	12	17	-5	0.545	0.078	66.365	15.065	20	0.847	0.509
1091873336	13	16	-3	0.578	0.089	68.608	15.661	100.5	0.74	0.508
1108651380	3	3	0	0.362	0.375	35.181	37.097	0	0.333	0.333

Figure C.2: Fall 2020 CU Boulder Collaboration Network

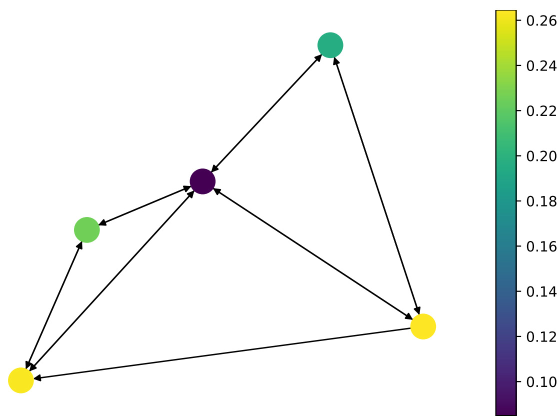


This network also contains six completely disconnected nodes, representing students for whom no collaboration was reported, which are not shown in this image. Nodes are colored based on their outward directed closeness centrality.

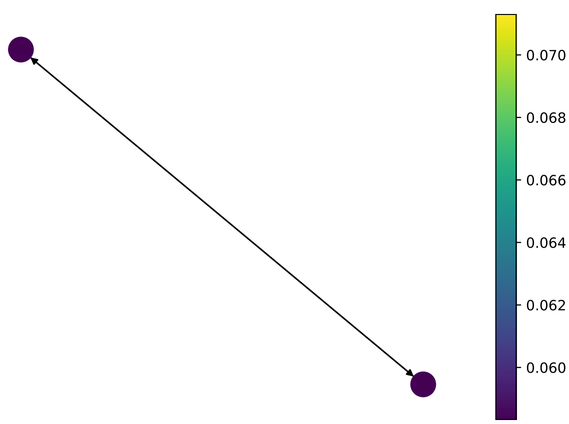
Figure C.3: Network Components of Spring 2021 CU Boulder Collaboration Network



(a) First Component of Spring 2021 CU Boulder Collaboration Network



(b) Second Component of Spring 2021 CU Boulder Collaboration Network



(c) Third Component of Spring 2021 CU Boulder Collaboration Network

Nodes are colored based on their outward directed closeness centrality. In addition to the students represented in these three components, there were 16 completely disconnected students who do not appear in any of the three components above.

Table C.3: Network Data from Spring 2021 Thermal Physics Course at CU Boulder

Student ID	s_i^{in}	s_i^{out}	s_i^{net}	c_i^{Co}	c_i^{Ci}	c_i^{Ho}	c_i^{Hi}	c^B	Y_i^{in}	Y_i^{out}
1021212261	20	28	-8	0.403	0.34	43.164	34.532	13.333	0.2	0.194
1041206667	26	18	8	0.527	0.525	36.391	41.34	182.667	0.121	0.154
1061698495	18	17	1	0.376	0.367	35.063	35.795	0	0.241	0.239
1061960044	1	2	-1	0.065	0.034	3.556	1.857	0	1	1
1071068671	0	0	0	0	0	0	0	0	undef.	undef.
1071104911	0	0	0	0	0	0	0	0	undef.	undef.
1071123445	0	0	0	0	0	0	0	0	undef.	undef.
1071222745	23	24	-1	0.4	0.385	42.415	41.343	33.833	0.217	0.219
1071299926	0	0	0	0	0	0	0	0	undef.	undef.
1071345547	19	10	9	0.391	0.534	24.5	38.517	69.167	0.152	0.16
1071411427	8	8	0	0.065	0.111	8	8	2	0.625	0.781
1071594646	19	27	-8	0.509	0.451	44.809	35.117	37.85	0.169	0.16
1071781558	16	15	1	0.428	0.422	33.299	33.943	109.667	0.172	0.182
1071880003	0	0	0	0	0	0	0	0	undef.	undef.
1071999988	0	0	0	0	0	0	0	0	undef.	undef.
1081001198	1	2	-1	0.35	0.254	21.282	14.808	0	1	1
1081037249	34	19	15	0.408	0.424	30.959	43.951	99	0.135	0.108
1081107980	0	0	0	0	0	0	0	0	undef.	undef.
1081317356	3	1	2	0.22	0.301	12.588	17.228	0	0.333	1
1081708268	15	8	7	0.196	0.225	17.384	20.583	0	0.476	0.781
1081785008	10	10	0	0.431	0.441	27.881	28.512	69.333	0.2	0.2
1081812866	11	11	0	0.226	0.213	21.556	20.792	0	0.835	0.835
1081855331	0	0	0	0	0	0	0	0	undef.	undef.
1081935365	0	0	0	0	0	0	0	0	undef.	undef.
1081945064	16	14	2	0.453	0.441	30.652	29.868	9.083	0.117	0.143
1081948367	22	26	-4	0.265	0.247	28.857	27.425	11	0.417	0.334
1081950749	0	0	0	0	0	0	0	0	undef.	undef.
1091067536	28	28	0	0.391	0.373	43.726	45.245	48.5	0.25	0.191
1091155322	4	5	-1	0.3	0.28	17.343	16.55	26	0.25	0.2
1091200214	34	30	4	0.538	0.538	42.862	48.333	140.85	0.137	0.111
1091241929	20	23	-3	0.263	0.247	28.909	25.909	11	0.395	0.372
1091276144	0	0	0	0	0	0	0	0	undef.	undef.
1091298002	0	0	0	0	0	0	0	0	undef.	undef.
1091315402	0	0	0	0	0	0	0	0	undef.	undef.
1091392244	7	6	1	0.408	0.426	28.288	31.074	0	0.51	0.5
1091402021	13	17	-4	0.274	0.271	28.747	25.533	75	0.728	0.481
1091463212	4	0	4	0	0.255	0	19.513	0	1	0
1091519009	24	31	-7	0.539	0.476	48.481	42.085	55.483	0.194	0.165
1091598128	10	15	-5	0.459	0.38	31.348	23.735	3	0.14	0.164
1091612333	0	0	0	0	0	0	0	0	undef.	undef.
1091769499	19	21	-2	0.547	0.502	39.414	37.293	70.583	0.152	0.143
1091799827	9	7	2	0.331	0.373	19.899	23.839	0	0.16	0.184
1091812334	18	19	-1	0.487	0.508	33.586	34.97	149.033	0.117	0.108
1091857421	11	11	0	0.261	0.258	26.944	24.751	0	1	1
1091895554	0	0	0	0	0	0	0	0	undef.	undef.
1091978729	2	3	-1	0.085	0.08	4.65	4.411	0	0.5	0.333
1091989257	7	8	-1	0.36	0.352	23.241	22.997	0	0.224	0.188
1108290413	0	0	0	0	0	0	0	0	undef.	undef.

Appendix D

Results of Testing Methods for Combining Student Reports

In the early stages of the networks project, a variety of methods for combining students' reports were tested. In the original study, students reports of getting and giving help were combined homework assignment-by-assignment using a logical OR operation. Four other methods for combining reports were implemented in this study to explore whether they resulted in different correlations between centrality and grades. The first method used a logical AND operation. In this case, in order for a reported interaction to make it into the adjacency matrix, both students had to have consistently reported the interaction. This is a much stronger condition and resulted in much sparser matrices. Two more methods (represented in the final two columns in Tables D.1, D.2, and D.3) looked simply at the total matrices generated by only looking at the reports of getting help (the *gothelp* matrix) and giving help (the *helped* matrix), respectively with out doing any combination. The final method (referred to as 'Total' in the the fourth column of Tables D.1, D.2, and D.3) directly summed the *gothelp* and *helped* (after the appropriate transpositions were made). This method is similar to using the logical OR, except it gives a boost to the edge weight when students both students report an interaction. As can be seen in Tables D.1, D.2, and D.3, each of the five methods generally gave consistent results, which provides some level of comfort that the results we reported in Chapter 4 were dependent on the method used to combine students' reports.

Table D.1: Correlations from Different Report Combination Methods: Mines Fall 2020

Centrality Measure	OR	AND	Total	Gothelp	Helped
s_i^{in}	0.549*	0.460*	0.545*	0.455*	0.533*
s_i^{out}	0.547*	0.496*	0.554*	0.540*	0.465*
s_i^{net}	-0.002	-0.165	-0.144	-0.003	-0.015
c_i^{Co}	0.718*	0.542*	0.712*	0.624*	0.662*
c_i^{Ho}	0.687*	0.522	0.671*	0.580*	0.631*
c^B	0.389	0.445	0.406	0.373	0.386
Y_i^{in}	-0.142	-.688*	-0.164	-0.124	-0.110
Y_i^{out}	-0.14	-.708*	-0.17	-0.106	-0.132

Correlations between centrality and homework grades from networks generated with different methods for combining student reports of getting and giving help. Statistically significant correlations (at $p < .05$) are indicated with an asterisk.

Table D.2: Correlations from Different Report Combination Methods: CU Boulder Fall 2020

Centrality Measure	OR	AND	Total	Gothelp	Helped
s_i^{in}	0.246	0.237	0.250	0.228	0.260
s_i^{out}	0.274*	0.240	0.269*	0.268*	0.254
s_i^{net}	-0.194	-0.033	-0.183	-0.022	-0.082
c_i^{Co}	0.430*	0.300*	0.408*	0.391*	0.241
c_i^{Ho}	0.395*	0.240	0.361*	0.356*	0.253
c^B	0.075	0.189	0.089	0.101	0.142
Y_i^{in}	0.018	-.285	0.012	0.011	-0.131
Y_i^{out}	-0.093	-0.17	-0.076	0.026	-0.177

Correlations between centrality and homework grades from networks generated with different methods for combining student reports of getting and giving help. Statistically significant correlations (at $p < .05$) are indicated with an asterisk.

Table D.3: Correlations from Different Report Combination Methods: CU Boulder Spring 2021

Centrality Measure	OR	AND	Total	Gothelp	Helped
s_i^{in}	0.218	0.214	0.225	0.237	0.201
s_i^{out}	0.251	0.316*	0.281	0.275	0.276
s_i^{net}	-0.089	-0.223	-0.157	-0.081	-0.151
c_i^{Co}	0.160	0.276	0.181	0.183	0.205
c_i^{Ho}	0.219	0.288*	0.250	0.235	0.237
c^B	0.019	0.121	0.072	0.085	0.098
Y_i^{in}	-0.024	0.074	-0.013	0.158	-0.051
Y_i^{out}	-0.057	-0.143	-0.041	-0.264	0.035

Correlations between centrality and homework grades from networks generated with different methods for combining student reports of getting and giving help. Statistically significant correlations (at $p < .05$) are indicated with an asterisk.