Winter 2018

Probing the Nature of Student Reasoning Using Modified Chaining Tasks

Ryan Moyer
University of Maine

PROBING THE NATURE OF STUDENT REASONING USING MODIFIED

CHAINING TASKS

by

Ryan P Moyer

A Thesis Submitted in Partial Fulfillment
of the Requirements for a Degree with Honors
(Physics)

The Honors College

University of Maine

December 2018

Advisory Committee:
MacKenzie R. Stetzer, Associate Professor of Physics (Advisor)
John R. Thompson, Professor and Chair of Physics
Robert J. Lad, Professor of Physics
Michael C. Wittmann, Professor of Physics
François G. Amar, Professor of Chemistry and Dean, Honors College

ABSTRACT

Research-based materials developed by the physics education research community have helped improve student conceptual understanding in introductory physics courses. A growing body of work, however, suggests that poor student performance on certain physics tasks, even after research-based instruction, may result from the nature of student reasoning itself than from conceptual difficulties. Drawing upon dual-process theories of reasoning, it has been argued that some of the poor performance from the presence of salient distracting features (SDFs) in physics problems, which may cue an incorrect first-available mental model and effectively preclude the student from drawing upon relevant knowledge.

In this study, we explored the relationship between students' initial impressions of how to approach a given physics problem and their subsequent performance on the problem. We accomplished this by employing a novel two-stage methodology in which students were first given a problem, provided with reasoning elements, and asked to categorize these elements as being useful or not useful for solving the problem. Students were subsequently asked to use these elements to construct a reasoning chain in order to arrive at an answer. Three problems were administered to students in introductory calculus-based physics.

We found that there was a relationship between students' sorting of the elements and students' final answers. Specifically, students who initially rejected relevant reasoning elements in favor of elements related to a problem's SDF were more likely to settle upon an incorrect, SDF-cued answer than students who initially endorsed the relevant elements and rejected the SDF-related elements.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. MacKenzie Stetzer, for his endless patience with my poor timing, and without whom this project never would have developed beyond a vague pondering.

I would also like to thank Caleb Speirs, for designing the framework upon which I have built my project, and for teaching me what I know about statistical analysis.

This material is based upon work supported by the National Science Foundation under Grant Nos. DUE-1431857, DUE-1431541, DUE-1431940, DUE-1432765, DUE-1432052, and DRL-0962805.

CONTENTS

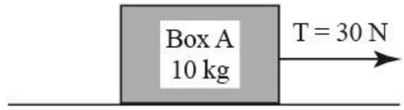
LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In physics education research, there has historically been a strong focus on conceptual understanding [3]. The results from more than thirty years of research have in turn led to the development of research-based instructional materials that have been shown to improve student conceptual understanding. However, it has been observed that student performance can vary on related questions that target student understanding of the same concept even after research-based instruction. As a specific example, researchers have administered a two-question sequence on the application of Newton's Second Law to situations involving static friction, as shown in Figures 1 and 2. Both questions require the same line of reasoning for the students to reach the correct conclusion; in particular, students need to recognize that, from Newton's Second Law, the net force on the boxes in both situations must be zero since they both remain at rest and therefore have no acceleration. In the first question, students were asked to determine whether the applied force was greater than, less than, or equal in magnitude to the frictional force [1]. 83% of students came to the correct conclusion that the forces were equal in magnitude.

Box A is initially at rest on a rough floor. A horizontal 30 N force is then applied to the box, as shown below. The box remains at rest.



Is the magnitude of the applied force *greater than, less than,* or *equal to* the magnitude of the force of friction? Explain.

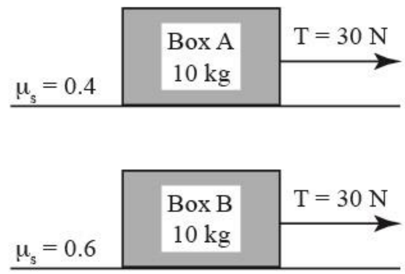
Figure 1: Screening question in the forces sequence from [1].

The key to the correct line of reasoning here is to note that the box remains at rest. If the box remains at rest, there must be no net force on the box, meaning the frictional force must be equal in magnitude to the applied force.

In the second question of the sequence, students were asked to compare the magnitudes of the frictional forces on two identical box in the same situation as the first question, differing only in the coefficients of friction between the box and the surfaces on which they are resting. Here the expected line of reasoning for the correct response remains unchanged; the boxes remain at rest means there must be no net force. Despite the fact that both questions rely on the same concept, Newton's Second Law in a static situation, only 65% of students correctly concluded that the frictional force is equal for both boxes. In contrast, those students who reached an incorrect answer, including roughly 1/5 of the students who applied the correct reasoning to the first problem, commonly reached the conclusion that the force of friction was greater on box B, often drawing upon the formula for maximum static friction, ($f_{s,max}=\mu N$, where μ is the coefficient of static friction and N is the normal force exerted on the box by the ground) in supporting their answers. This line of reasoning, which argued that the box with the larger coefficient of static friction would experience the larger frictional force, appeared to be connected to the inclusion of the two different coefficients of static friction.

Suppose the coefficient of static friction between box A and the floor is 0.4, as shown below. The coefficient of static friction between box B and a different floor is 0.6, as shown below. $m_A = m_B = 10$ kg.

A horizontal 30 N force is applied to each box, and both boxes remain at rest.



Is the magnitude of the friction force exerted on box A *greater than, less than,* or *equal to* that exerted on box B? Explain.

Figure 2: Target question in the forces sequence from [1].

The published results from this sequence of questions suggest that something other than a lack of conceptual understanding is impacting students' reasoning abilities, as a significant percentage of students who gave a correct response on the first question adopted a new line of reasoning based on a different intuitive relationship. Other research using this two-question methodology has shown that roughly 50% of students who gave the correct response to the first question in the sequence gave the common incorrect response to the second question, with their reasoning cued by incorrect intuitive ideas about the physical scenarios (*e.g.* conservation of the voltage across two capacitors connected in series), despite having demonstrated an understanding of the relevant physical concepts in their responses to the screening question. [5]. In the case of the friction sequence shown in Figures 1 and 2, the addition of the coefficients of static friction appeared to inhibit students' ability to pursue the correct line of reasoning by perhaps cuing a more intuitive approach, where a larger number (the larger coefficient of static friction) should result in a larger number (the larger frictional force). We refer to such question components (*e.g.*, the coefficients of static friction in the friction sequence) that capture students' attention negatively affect student performances as salient distracting features, or SDFs.

Heckler [2] looks at the choice of which line of reasoning is selected through the lens of competing mental models. In the case of the friction sequence, the coefficients of friction cue a mental model that the brain processes more quickly and seems applicable to the problem on the surface. In order to determine why these salient distracting features can so significantly affect student performance on problems, we turn to cognitive science.

1.1 Dual-Process Theories of Reasoning and Decision-Making

A number of theories have been developed in cognitive science attempting to describe the process of individual reasoning and decision-making. Taken collectively, we refer to these theories as dual-process theories of reasoning [6,7]. These dual-process theories posit that human reasoning may be categorized by two processes: the heuristic (sometimes called process 1) and the analytic (or process 2). The heuristic process (process 1) occurs very quickly, often subconsciously. Essentially, the brain constructs a relevant mental model based on the individual's goals, expectations, previous experience, and situational cues, and this first-available mental model is often applied without any further interruptions. A common example of the heuristic process is the fact that you can quickly tell someone is angry when you first see that person, without thinking about the specific characteristics that "tell" you that is the case.

The analytic process (process 2) is much slower and more effortful than the heuristic process, which is why the heuristic process is responsible for most common, day-to-day decisions. The analytic process requires considerable mental effort, making it poorly suited for such common decisions. An example of analytical processing is solving a complex mathematical problem.

When the analytic process is first engaged, it begins by assessing whether the mental model generated by the heuristic process is a satisfactory model for the task at hand. Due to various biases people have regarding their own judgments (e.g. confirmation bias, a reluctance to expressly search for counterexamples, and a tendency to rationalize), it is likely that the original model will still be deemed satisfactory [6]. However, in the case in which the initial mental model is not deemed satisfactory, the mental model is modified or replaced before being reassessed. This pattern repeats until the mental model is considered a satisfactory response.

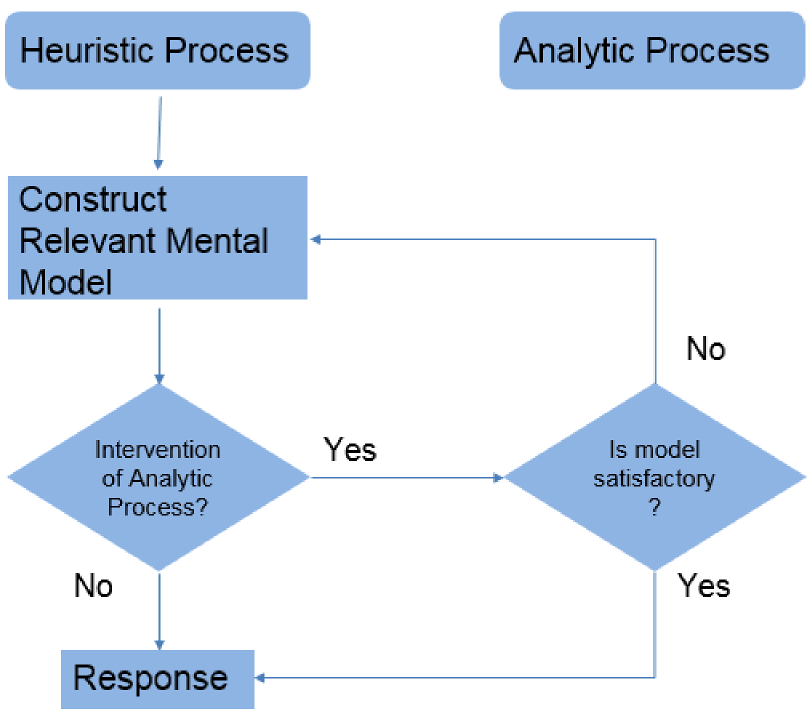


Figure 3: A Diagram, adapted from [6] Illustrating Evans' revised and extended heuristic-analytic theory.

Researchers in physics education have begun applying dual-process theories of reasoning to account for student performance on qualitative physics questions. In particular, Heckler has explored the role of salient distracting features and how to determine what information a student uses to solve a given problem, putting forward of model of competing relevant and irrelevant information. Kryjevskaia, Stetzer, and Groz observed, through sequences of screening questions to check conceptual understanding and target questions with SDFs, that "a significant fraction of students who applied correct and complete reasoning on the screening question(s) failed to do so on a target question that called for the same knowledge and skills [8]." They argued that, because such failure to apply proper reasoning arises from a failure to engage the analytic process based on heuristic-analytic theory, focused effort on improving student metacognition would improve ability to engage the analytic process and thus improve overall reasoning. As a specific example, if we reexamine the paired question sequence from Figures 1 and 2, one can recognize the general process leading to the incorrect conclusion. One sees the different μs for the two boxes. Through the heuristic process, this cues a mental model on the idea that bigger μ means bigger frictional force. Upon being asked to explain their reasoning, the analytic process is engaged. The students rationalize their current model based on the expression for maximum static friction or kinetic friction. The mental model is approved by the analytic process, and the student concludes that there is a greater frictional force on box B.

The primary goal of this investigation was to move closer to accessing students' heuristic thinking by identifying the features of the question and the reasoning elements that they thought would be useful. To this end, we administered a number of problems including an initial task in which students were first asked to sort provided reasoning elements (*e.g.* given statements about the physical situation as well as correct concepts and mathematical relationships) based on whether or not they thought the elements would be useful for solving the problem prior to formally solving it. Students then solved the problem in a reasoning chain construction task format, which is described in the next section.

1.2 Research Question and Methodology

Since our overarching goal was to investigate the relationship between students' first-available mental models (from the heuristic process) and their final answers, we wanted to document their initial perceptions of which reasoning elements, composed of true statements or physics concepts that one could apply to a given problem, would be useful in solving the problem. Doing so would enable us to examine the relationship between those perceptions and student performance on the problem. Thus, the research question driving the investigation was the following: *When first looking at a reasoning chain construction task, which elements do students identify as being useful (or not useful) in solving the physics question?*

Dual-process theories of reasoning would predict that incorrect first-available mental models, cued by the salient distracting features of the problem may prevent some students from exploring alternatives while actively solving the problem. Thus, one would predict that those students who exclusively attend to the salient distracting feature during the sorting task are more likely to arrive at an incorrect answer.

As the method to gather data, we administered several problems to students in introductory physics as part of online exam reviews, divided into two stages, using the

Qualtrics online survey platform [10]. In stage 1, students were first presented with the problem statement and all of the reasoning elements, and were asked to sort the reasoning elements into one of two groups: “Items you believe will be useful,” and “Items you believe will not be useful.” After confirming that they were satisfied with how they have sorted all the elements, students then proceeded to answer the problem through stage 2, the chaining task. The chaining task consists of a reasoning space and three different pools of "tiles" to place into the reasoning space: the reasoning elements (described previously), connecting words (such as, and, so, but, therefore, *etc.*), and the possible conclusions to the problem [9]. See Figure 4. By moving the reasoning elements into the reasoning space and incorporating connecting words as needed, students used the interface [10] to actively construct the lines of reasoning they employed to arrive at a conclusion.

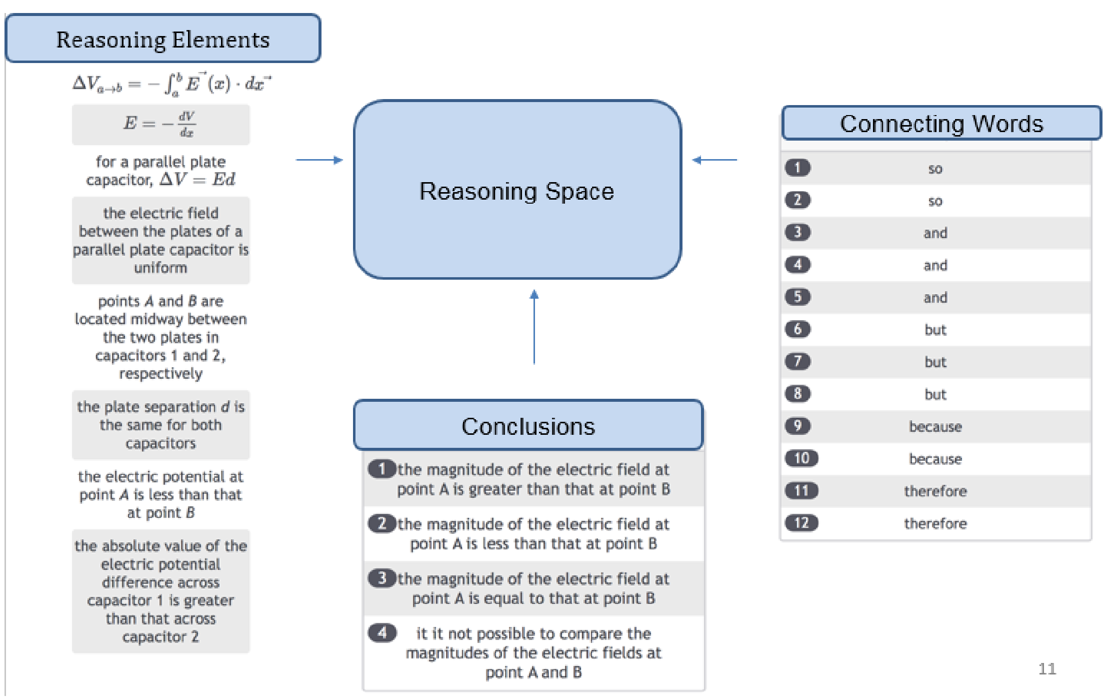


Figure 4: Example Chaining Task.

Research tasks were either drawn from the literature or developed by the research team such that there was a well-defined salient distracting feature (that would serve to cue a common incorrect response) as well as relevant reasoning elements. In this thesis, we discuss results from three different tasks: (1) the Capacitor Comparison Question, (2) Charge Ring Comparison Question: Different Total Charge, and (3) Charge Ring Comparison Question: Same Total Charge.

Research tasks were administered on exam review assignments for participation credit in the second semester of the introductory calculus-based physics sequence at the University of Maine. For each task, student data were analyzed on the basis of reasoning elements selected as useful or not useful in stage 1 (sorting task) and on the basis of the conclusions drawn at the end of stage 2 (reasoning chain construction task). To answer our research question, we focused on student responses that either endorsed the SDF-related reasoning element and rejected one or more of the relevant reasoning elements or that endorsed one or more of the relevant reasoning elements and rejected the SDF-related reasoning elements prior to reasoning chain construction. We then examined the performance of each group of students on the physics question in stage 2 and determined whether or not any differences between the two groups of responses were statistically significant.

RESEARCH TASKS AND RESULTS

In this section, we describe three research tasks and present the associated results from the data collected.

2.1 Capacitor Comparison Task

The capacitor comparison task was drawn from the literature and adapted slightly for use in this study. Note that it was selected due to the presence of a documented salient distracting feature that leads to a common incorrect response.

Problem Overview

In this problem (see Figure 3), originally developed and used by A.F. Heckler and reported in an article entitled “The Ubiquitous Patterns of Incorrect Answers to Science Questions: The Role of Automatic, Bottom-up Processes” [2], students are shown two capacitors and are provided with information about the electric potentials on each plate and in the middle of each capacitor. Students are asked to determine whether the magnitude of the electric field at point *A* in capacitor 1 is greater than, less than, or equal to that ~~of~~ at point *B* in capacitor 2.

Two capacitors, 1 and 2, of identical plate separation *d* are shown in the figure below. The plates of the capacitors are kept at the constant potentials shown below. Points *A* and *B* are located midway between the two plates in capacitors 1 and 2, respectively.

Is the magnitude of the electric field at point *A* *greater than, less than,* or *equal to* the magnitude of the electric field at point *B?*

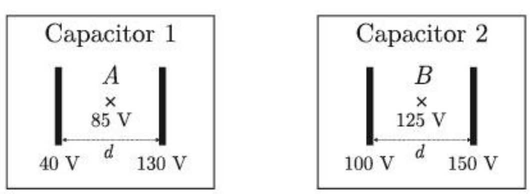


Figure 5: Capacitor Comparison Task.

The expected line of reasoning required to arrive at a correct conclusion (that the electric field at point *A* in capacitor 1 has a greater magnitude) is to recognize that the strength of the electric field in a parallel plate capacitor is related to the electric potential difference between the plates by $E = \frac{\Delta V}{d}$ and that both plates have the same spacing d. However, this question was particularly useful as a starting point for this investigation since the question has a known SDF – the electric potential at the midpoint of each capacitor. Thus, the SDF-associated line of reasoning is to associate the electric field at a point with the value of the electric potential at a point, thereby leading to the incorrect conclusion that the electric field at point *A* in capacitor 1 is smaller than that at point *B* in capacitor 2.

Table 1: Reasoning Elements In Capacitor Comparison Task Responses.

$\Delta V_{ab} = \int_a^b \vec{E}(x) \cdot \overrightarrow{dx}$
$E = \frac{dV}{dx}$
For a parallel plate capacitor $\Delta V = Ed$
Points *A* and *B* are midway between the plates for capacitors 1 and 2 respectively
The electric field between two plates of a parallel plate capacitor is uniform
The plate separation d is the same for both capacitors
The electric potential at point *A* is less than at point *B* **(SDF)**
The absolute value of the electric potential difference across capacitor 1 is greater than that across capacitor 2 **(Relevant)**

Results

This problem was administered in an online exam review to 245 students. From the overall student performance (Table 2), we can see there is roughly a 50-50 split between the correct and common incorrect responses. While Heckler did not report the prevalence of all responses in his study, our results are consistent with his reported "roughly 50% of students" correctly concluding that the magnitude of the electric field at point A is greater than the magnitude of the electric field at point B [2]. Based on this consistency, we are inclined to believe that both the inclusion of the sorting task and the use of the reasoning chain construction format did not significantly impact student performance on the problem.

Table 2: Student Performance on Capacitor Comparison Task in Chaining Format, N = 206.

Conclusion	**% of Total Responses**
$E_A < E_B$	47%
$E_A > E_B$ **(Correct)**	44%
$E_A = E_B$	9%

The results (Table 3) of the sorting task for the relevant reasoning element ($\Delta V_1 > \Delta V_2$) and SDF-associated reasoning element ($V_A < V_B$) show that for each element, approximately 50% of students said it was useful. Thus, students did not seem to demonstrate a strong initial preference for one of these elements over the other. We were, however, most interested in comparing the performance of those students who specifically sorted the relevant element as useful and the SDF-associated element as not useful to that of those who sorted the SDF-associated element as useful and the relevant element as not useful, as each group was, at least initially, strongly committed to focusing on a particular sub-set of problem features.

Table 3: Sorting Task Results for Capacitor Comparison Task.

Classification	Reasoning Element % of total responses (N=245)	
	$\Delta V_1 > \Delta V_2$	$V_A < V_B$
Useful	55%	47%
Not Useful	45%	53%

By looking at the extremes of the possible results from the sorting task (sorted both elements as useful, sorted both as not useful, or our results of interest where one element was sorted as useful and the other was sorted as not useful) we can see more clearly that those students who endorsed the SDF-related element and rejected the relevant element had much higher rates of reaching an incorrect conclusion on the problem than those who endorsed the relevant element while rejecting the SDF-related element.

Table 4: Examining the Effect of Endorsing the SDF and Rejecting Relevant Information or vice-versa on Student Performance for the Capacitor Comparison Task.

Conclusion	**SDF is useful, relevant is not** N = 42	**Relevant is useful, SDF is not** N = 56
Correct	24%	86%
Incorrect	76%	14%

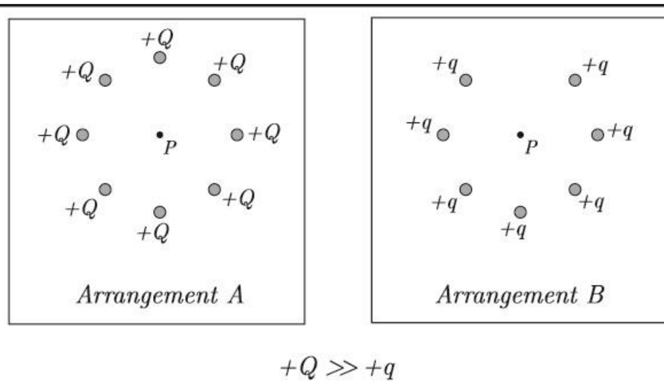
As a note, Table 4 reduces the categories from all possible conclusions down to correct and incorrect in order for the numbers to be large enough to perform a chi-squared test for statistical significance. This test gave us a χ^2 value of 38.074 and a P value less than 0.00001, showing that the two columns are so different that the difference cannot be attributed to random distribution and therefore can be associated with the differences in element sorting.

2.2 Charge Ring Comparison Task: Different Total Charge

In this section, we describe the first of two research tasks expressly designed for this project and report on the associated findings. Both tasks were administered as part of the same online exam review to a total 186 students, with 91 being randomly selected for this version and 95 for the alternative version described later. In order to ensure that the task was useful for the investigation, a salient distracting feature (as anticipated by researchers) was explicitly incorporated into the underlying physics problem.

Problem Overview

In this problem, shown in Figure 4, students are shown two arrangements of point charges in a ring about a point *P*, with all the point charges in a given ring being equal and are asked to determine whether the magnitude of the electric field is greater at the center of charge arrangement A or B.



Above are two different arrangements of point charges around point *P*. In each arrangement, the charges are fixed in place and are the same distance from point *P*. The charges in arrangement A ($+Q$) are all much greater in magnitude than those in arrangement B ($+q$).

Is the magnitude of the electric field at point *P* for arrangement A *greater than*, *less than*, or *equal to* that for arrangement B?

Figure 6: Charge Ring Comparison Task: Different Total Charge.

The expected line of reasoning for the correct conclusion is to note that the symmetry of arrangement A means the electric field contributions from each pair of charges will cancel out due to superposition, meaning that at point *P* in arrangement A the electric field will be 0. As this symmetry argument cannot be applied to arrangement B, there will be some electric field, so the magnitude of the electric field at point *P* is less in arrangement A than in arrangement B. The SDF in this problem is the charge comparison statement indicating that Q>>q. The line of reasoning associated with the SDF-related element (Q>>q), is similar to that given on the friction sequence target question (Figure 2) in that larger charges will result in a larger field.

Table 5: Reasoning Elements In Charge Ring Comparison Task: Different Total Charge

The net electric field at a point is the vector sum of the electric fields due to all charges
The electric field from identical charges on opposite sides of point *P* cancel
The electric field due to a positive charge is directed radially outward from the charge
Each identical point charge contributes an electric field at point *P* of the same magnitude
The charges in Arrangement A have greater magnitude than those in Arrangement B (+Q>>+q) **(SDF)**
The charges in Arrangement A are distributed symmetrically about point *P* **(Relevant)**
There are more charges in Arrangement A than in Arrangement B
The charges in Arrangement B are not distributed symmetrically about point *P* **(Relevant)**
$\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}\hat{r}$

<u>Results</u>

Looking at the overall student performance on the reasoning chain construction task, roughly half of students arrive at the correct conclusion, similar to the overall performance on the capacitor comparison task. For the incorrect responses, however, there is more of a distribution of responses, with 23% of students arriving at the common incorrect conclusion compared to 47% for the capacitor comparison task.

Table 6: Student Performance on Charge Ring Comparison Task: Different Total Charge in Chaining Format, N = 91.

Conclusion	**% of Responses**
$E_A > E_B$	23%
$E_A < E_B$ **(correct)**	55%
$E_A = E_B$	4%

Note that the percentages in Table 6 above do not sum to 100% due to an excluded subset of students. These excluded students either did not include a conclusion tile in their reasoning chain construction task response or they concluded that the answer could not be determined.

The sorting results for the SDF element in this task (charges in A are larger) are comparable to those for the capacitor comparison task, with roughly half of the students sorting it as useful. There is a distinction to be seen in the sorting results of the relevant elements (A is symmetrical and/or B is not), though. We see here that 73% of students sorted at least one of them as useful, whereas only 55% of students did so in the capacitor comparison task. That being said, this discrepancy may be due to the number of relevant elements. In the charge ring tasks, students were considered to be sorting relevant information as useful if they sorted either or both of the relevant elements as useful, whereas in the capacitor comparison task there was only one relevant element.

Table 7: Sorting Task Results for Charge Ring Comparison Task: Different Total Charge.

Classification	Reasoning Element % of total responses (N=91)	
	A is symmetrical and/or B is not	Charges in A are larger
Useful	73%	44%
Not Useful	27%	56%

From Table 8, we can again see that those students who sorted relevant elements as useful and the SDF-associated element as not useful had higher rates of reaching the correct conclusion than those who sorted the SDF-assoiated element as useful and relevant elements as not useful. The fact that those students who sorted the SDF-associated element as useful and relevant elements as not useful had equal rates of reaching correct and incorrect conclusions, compared to 76% of students in that subset reaching incorrect conclusions in the capacitor comparison task, can likely be attributed to the low N value.

Table 8: Examining the Effect of Endorsing the SDF and Rejecting Relevant Information or vice-versa on Student Performance for Charge Ring Comparison Task: Different Total Charge.

Conclusion	**SDF is useful, relevant is not** N = 14	**Relevant is useful, SDF is not** N = 40
Correct	50%	73%
Incorrect	50%	27%

We used a χ^2 test to check statistical significance of these results. While we cannot claim statistical significance due to the number of responses being too small for a χ^2 test to be valid, the test yielded a χ^2 of 4.668 and a P of 0.0307, which would normally indicate statistical significance. In order to gauge the extent to which a small fluctuation would impact these results, we arbitrarily shifted two student responses from the pool of "SDF is useful, relevant is not useful, correct," to "SDF is useful, relevant is not, incorrect," in order to satisfy the requirements of the test, and found that it still showed statistical significance.

2.3 Charge Ring Comparison Task: Same Total Charge

This problem is an alternate version of the previous Charge Ring Task. However, in this version, the total charge in each arrangement is the same, which serves as a new salient distracting feature, cueing the idea that the magnitudes of the electric fields in both arrangements will also be the same.

Problem Overview

In this problem, as in the Charge Ring Comparison Task: Different Total Charge, students are shown two arrangements of point charges arrayed around point of interest *P*, with the point charges on opposite sides of *P* being equal, and are asked to determine whether the magnitude of the electric field is greater at the center of arrangement A or B.

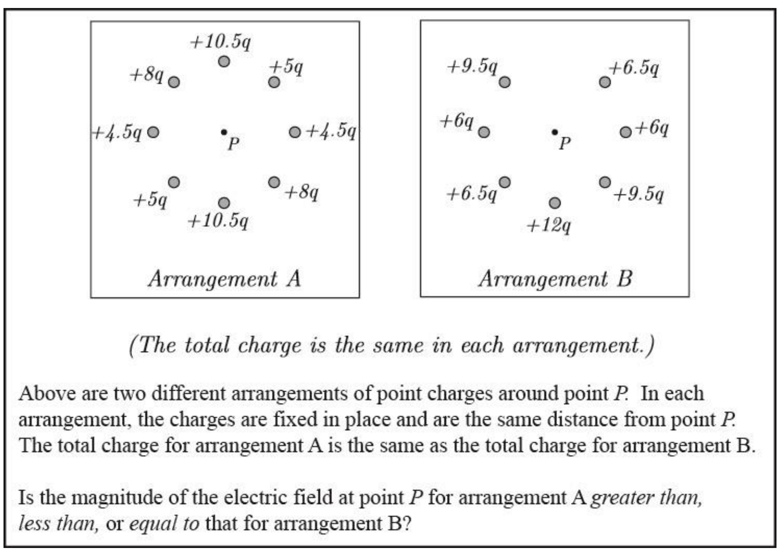


Figure 7: Charge Ring Comparison Task: Same Total Charge.

The expected correct approach is to use symmetry to show that the electric field contributions in arrangement A all cancel, while they do not for arrangement B. The SDF-associated line of reasoning focuses on the fact that the total charge is equal in both arrangements, leading students to conclude that the electric field is the same in both arrangements.

Students were provided with the reasoning elements shown in Table 9 below, including those elements given for the same total charge version of the task as well as elements unique to this version (signified by italics).

Table 9: Reasoning Elements In Charge Ring Comparison Task: Same Total Charge. Elements unique to this version of the Charge Ring Comparison Task are italicized.

The net electric field at a point is the vector sum of the electric fields due to all charges
The electric field from identical charges on opposite sides of point *P* cancel
The electric field due to a positive charge is directed radially outward from the charge
Each identical point charge contributes an electric field at point *P* of the same magnitude
The charges in Arrangement A are distributed symmetrically about point *P* **(Relevant)**
There are more charges in Arrangement A than in Arrangement B
The charges in Arrangement B are not distributed symmetrically about point *P* **(Relevant)**
$\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}\hat{r}$
The magnitude of all charges in Arrangement A are not the same
The magnitude of all charges in Arrangement B are not the same
The total charge is the same in each arrangement ***(SDF)***

Results

As with the other two tasks, roughly half of the students reached the correct conclusion in the reasoning chain construction task. Perhaps unsurprisingly, changing the salient distracting feature has changed the common incorrect response from $E_A > E_B$ in the different total charge version to $E_A = E_B$ in the same total charge version.

Table 10: Student Performance on Charge Ring Comparison Task: Same Total Charge in Chaining Format, N = 95.

Conclusion	% of Total Responses
$E_A > E_B$	8%
$E_A < E_B$ **(correct)**	57%
$E_A = E_B$	20%

As with the different total charge version, Table 11 shows that roughly half of the students sorted the SDF-associated element as useful, while a greater portion of students sorted relevant elements as useful rather than not useful. Again, this is possibly explained by the presence of multiple relevant reasoning elements in the task

Table 11: Sorting Task Results for Charge Ring Comparison Task: Same Total Charge

Classification	Reasoning Element % of total responses (N=91)	
	A is symmetrical or B is not	Total charge is the same
Useful	65%	49%
Not Useful	35%	51%

Looking at the extreme groups from the sorting task in Table 12, we see a continuation of the trend observed in the capacitor comparison task and the other charge ring comparison task, where those students who sorted the relevant elements as useful and the SDF-associated element as not useful are markedly more likely to reach the correct conclusion than those who sorted the SDF-associated element as useful and relevant elements as not useful

Table 12: Examining the Effect of Endorsing the SDF and Rejecting Relevant Information or vice-versa on Student Performance for Charge Ring Comparison Task: Same Total Charge.

Conclusion	**SDF is useful, relevant is not** N = 23	**Relevant is useful, SDF is not** N = 38
Correct	30%	76%
Incorrect	70%	24%

The χ^2 test for these data gives a χ^2 value of 12.47 and a P value of 0.0004, showing that the columns are different beyond what can be accounted for by random distribution. This statistically significant difference in performance suggests that there is, in fact, a relationship between the aspects of the problem context the students endorse and reject and the likelihood of arriving at a correct conclusion.

DISCUSSION

In this investigation, we attempted to move closer to probing students' first-available mental models by asking them to identify specific reasoning elements (including features of the problem) that they felt were useful or not useful prior to solving a problem. By examining the relationship between student performance on the problem itself and the reasoning elements they simultaneously endorsed and rejected, we found that students who initially rejected relevant reasoning elements in favor of SDF-related elements were more likely to reach an incorrect, SDF-cued conclusion than students who initially endorsed relevant elements and rejected SDF-related elements. The findings from this investigation further highlight the utility of dual-process theories of reasoning and decision-making in accounting for seemingly disparate student performance on physics problems by providing additional evidence that first-available mental models generated by the heuristic process and cued by salient distracting features can inhibit students' reasoning processes and lead them to incorrect conclusions. This would support the established idea of confirmation bias, stating that most reasoners are poor at generating alternative possibilities and instead tend to rationalize their first-available mental models.

CONCLUSION

The main goal of this research was attempting to get closer to the heuristic process by asking students about what physical concepts and problem features they initially think will be useful in problem solving prior to being asked to solve the problem. To accomplish this we administered a number of questions to students on online exam reviews in two parts: a sorting task in which students were asked to sort reasoning elements by whether the students thought they would be useful to solving the problem, followed by a reasoning chain construction task in which students were asked to use a number of reasoning elements to illustrate their line of reasoning. This method seems to have been successful at providing insight into students' initial thoughts on the problems, as we can somewhat reliably reconstruct the mental models seen in the chaining task or reject possible mental models based on what reasoning elements a student endorsed or rejected in the sorting task. The importance of this insight lies in enabling both researchers and instructors to identify the patterns in reasoning that would indicate students are failing to engage the analytical process in a productive manner, allowing them to modify instruction accordingly, possibly through the development and implementation of research that explicitly attend to the dual-process nature of student reasoning in physics.

BIBLIOGRAPHY

[1] M. Kryjevskaia, M. R. Stetzer, and T. K. Le, "Failure to Engage: Examining the Impact of Metacognitive Interventions on Persistent Intuitive Reasoning Approaches," *2014 PERC Proceedings* [Minneapolis, MN, July 30-31, 2014], edited by P. V. Engelhardt, A. D. Churukian, and D. L. Jones, doi:10.1119/perc.2014.pr.032.

[2] A. F. Heckler, "The Ubiquitous Patterns of Incorrect Answers to Science Questions: The Role of Automatic, Bottom-up Processes," *Psychology of Learning and Motivation*, 2011, pp. 227–267., doi:10.1016/b978-0-12-387691-1.00008-9.

[3] L. C. McDermott, "Research on Conceptual Understanding in Mechanics," *Physics Today*, vol. 37, no. 7, 1984, 24–32., doi:10.1063/1.2916318.

[4] L. C. McDermott, E. F. Redish, "Resource Letter: PER-1: Physics Education Research," *Am. J. Phys*. 67, 755 (1999)

[5] M. Kryjevskaia and M. R. Stetzer "Examining inconsistencies in student reasoning approaches," *2012 PERC Proceedings* [Philadelphia, PA, August 1-2, 2012], edited by P. V. Engelhardt, A. D. Churukian, and N. S. Rebello [AIP Conf. Proc. 1513, 226-229 (2013)], doi:10.1063/1.4789693.

[6] J. S. Evans "The heuristic-analytic theory of reasoning: Extension and evaluation," *Psychonomic Bulletin & Review* 13(3), 378-395, doi: 10.3758/bf03193858

[7] D. Kahneman, *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux (2015)

[8] M. Kryjevskaia, M. R. Stetzer, N. Grosz "Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics," *Phys. Rev. ST Phys. Educ. Res*, vol 10, no. 2, 2014, doi: 10.1103/PhysRevSTPER.10.020109

[9] J. C. Speirs, W. N. Ferm Jr., M. R. Stetzer, and B. A. Lindsey, "Probing Student Ability to Construct Reasoning Chains: A New Methodology," *2016 PERC Proceedings* [Sacramento, CA, July 20-21, 2016], edited by D. L. Jones, L. Ding, and A. Traxler, doi:10.1119/perc.2016.pr.077.

[10] The Leading Research & Experience Software. (n.d.). Retrieved from http://www.qualtrics.com/

AUTHOR'S BIOGRAPHY

Ryan Moyer was born May 28, 1996, in Virginia Beach, VA, to mother Stephanie Moyer and father Danny Moyer, and has lived in Freeport, ME since the age of 3. Ryan is currently pursuing a Bachelor of Science in physics from the University of Maine. Ryan is a member of the Sigma Pi Sigma physics honor society and Pi Mu Epsilon math honor societies. He has been involved in the University of Maine's chapter of the society of physics students since his sophomore year of college, and has participated in a number of outreach events for physics education.

Upon completion of his Bachelor's degree, Ryan intends to continue on to grad school to attain a Master's of Science in Teaching degree from the University of Maine, specializing in physics education and education research.