

Spring 1-1-2015

New Tools for Investigating Student Learning in Upper-division Electrostatics

Bethany Rae Wilcox

University of Colorado Boulder, bethany.wilcox@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/phys_gradetds

 Part of the [Higher Education Commons](#), [Physics Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

Wilcox, Bethany Rae, "New Tools for Investigating Student Learning in Upper-division Electrostatics" (2015). *Physics Graduate Theses & Dissertations*. 135.

https://scholar.colorado.edu/phys_gradetds/135

This Dissertation is brought to you for free and open access by Physics at CU Scholar. It has been accepted for inclusion in Physics Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.

**NEW TOOLS FOR INVESTIGATING STUDENT LEARNING
in UPPER-DIVISION ELECTROSTATICS**

by

BETHANY R. WILCOX

B.A., University of Colorado, 2010

M.S., University of Colorado, 2012

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Department of Physics
2015

This thesis is entitled:
New Tools for Investigating Student Learning in Upper-division Electrostatics
written by Bethany R. Wilcox
has been approved for the Department of Physics

Steven Pollock

Noah Finkelstein

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol # 12-0011

Wilcox, Bethany R. (Ph.D., Physics)

New Tools for Investigating Student Learning in Upper-division Electrostatics

Thesis directed by Professor Steven J. Pollock

Student learning in upper-division physics courses is a growing area of research in the field of Physics Education. Developing effective new curricular materials and pedagogical techniques to improve student learning in upper-division courses requires knowledge of both what material students struggle with and what curricular approaches help to overcome these struggles. To facilitate the course transformation process for one specific content area – upper-division electrostatics – this thesis presents two new methodological tools: (1) an analytical framework designed to investigate students’ struggles with the advanced physics content and mathematically sophisticated tools/techniques required at the junior and senior level, and (2) a new multiple-response conceptual assessment designed to measure student learning and assess the effectiveness of different curricular approaches.

We first describe the development and theoretical grounding of a new analytical framework designed to characterize how students use mathematical tools and techniques during physics problem solving. We apply this framework to investigate student difficulties with three specific mathematical tools used in upper-division electrostatics: multivariable integration in the context of Coulomb’s law, the Dirac delta function in the context of expressing volume charge densities, and separation of variables as a technique to solve Laplace’s equation. We find a number of common themes in students’ difficulties around these mathematical tools including: recognizing when a particular mathematical tool is appropriate for a given physics problem, mapping between the specific physical context and the formal mathematical structures, and reflecting spontaneously on the solution to a physics problem to gain physical insight or ensure consistency with expected results.

We then describe the development of a novel, multiple-response version of an existing conceptual assessment in upper-division electrostatics courses. The goal of this new version is to provide an easily-graded electrostatics assessment that can potentially be implemented

to investigate student learning on a large scale. We show that student performance on the new multiple-response version exhibits a significant degree of consistency with performance on the free-response version, and that it continues to provide significant insight into student reasoning and student difficulties. Moreover, we demonstrate that the new assessment is both valid and reliable using data from upper-division physics students at multiple institutions. Overall, the work described in this thesis represents a significant contribution to the methodological tools available to researchers and instructors interested in improving student learning at the upper-division level.

ACKNOWLEDGMENTS

The research presented in this thesis has been supported by the National Science Foundation through a CCLI Grant DUE-1023028 and a Graduate Research Fellowship under Grant No. DGE-1144083. I would also like to acknowledge that the work in this dissertation appears in several articles published by the American Physical Society and American Institute of Physics. A sincere thank you to the E&M 1 students and instructors without whom this research would not have been possible.

I would also like to acknowledge the other two people who have contributed most significantly to the intellectual content of this thesis: my advisor Steve Pollock and my colleague Danny Caballero. Thank you both for supporting me throughout my graduate career and helping me to develop as a researcher, a teacher, a professional, and a person. Without fail, I have found our collaborations productive and our conversations both challenging and stimulating. I would also like to thank the other members of the PER group at CU for their ideas and support. You have all contributed to this thesis by creating the supportive and intellectually rich environment in which it was created.

I would also like to say thank you to my friends and family for everything they have done to help me discover and achieve my personal and professional goals. To my Mom, who has supported me in more ways that I can possibly express through every stage of my life. To my Dad, who has carefully fostered my scientific curiosity and intellectual development since I was kid. To my Sister, who is always there when I need advice or a sympathetic ear. To Alex, who agonized with me over endless graduate homework and has never failed to be there even when I didn't know I needed her. And to Ashley, who introduced me to the wonders of dance and always reminds me to take risks and explore the world outside my comfort zone.

Table of Contents

I. CHAPTER I: INTRODUCTION	1
A. Motivation	1
B. Overview of the Dissertation	3
C. Summary of Key Findings	5
II. CHAPTER II: DEVELOPING A FRAMEWORK	9
A. Introduction and Motivation	9
B. Existing Problem-Solving Strategies and Theoretical Frameworks	11
C. The ACER Framework	15
1. <i>Overview</i>	15
2. <i>Operationalizing the Framework</i>	20
III. CHAPTER III: APPLICATION OF THE FRAMEWORK	22
A. Context and Data sources	22
B. Multivariable Integration in the Context of Coulomb's Law	23
1. <i>Previous Research on Student Difficulties with Integration</i>	23
2. <i>Operationalizing ACER for Multivariable Integration</i>	25
3. <i>Interview and Exam Prompts</i>	29
4. <i>Student Difficulties with Multivariable Integration</i>	30
5. <i>Summary and Implications</i>	38
C. The Dirac Delta Function in the context of Electrostatics	39
1. <i>Previous Research on Student Difficulties with the Dirac Delta Function</i>	39
2. <i>Operationalizing ACER for the Dirac Delta Function</i>	40
3. <i>Interview and Exam Prompts</i>	44
4. <i>Student Difficulties with the Dirac Delta Function</i>	47
5. <i>Summary and Implications</i>	55

D. Separation of Variables in the context of Laplace’s Equation	57
1. <i>Previous Research on Student Difficulties with Separation of Variables</i>	57
2. <i>Operationalizing ACER for Separation of Variables</i>	58
3. <i>Interview and Exam Prompts</i>	62
4. <i>Student difficulties with Separation of Variables in Cartesian Coordinates</i>	66
5. <i>Student difficulties with Separation of Variables in Spherical Coordinates</i>	74
6. <i>Summary and Implications</i>	81
E. Synthesis and Discussion	84
IV. CHAPTER IV: DEVELOPING THE MULTIPLE-RESPONSE CUE	
DIAGNOSTIC	87
A. Introduction	87
1. <i>Background and Motivation</i>	87
2. <i>Previous Research on Multiple-choice testing</i>	91
3. <i>Testing Theory</i>	93
B. Creating a Multiple-response Version	95
1. <i>Adapting the Questions</i>	95
2. <i>Scoring</i>	98
C. Establishing Content Validity	100
1. <i>Expert Validation</i>	100
2. <i>Student Validation</i>	102
V. CHAPTER V: STATISTICAL VALIDATION OF THE CMR CUE	104
A. Comparing the Multiple-response and Free-response Versions	104
1. <i>Methods</i>	104
2. <i>Results: Statistical Comparison</i>	105
3. <i>Results: Reasoning Comparison</i>	110
B. Validating the CMR CUE	115

1. <i>Methods</i>	115
2. <i>Results: Statistical Validation</i>	116
C. Accessing Student Difficulties with the CMR CUE	123
D. Summary and Discussion	127
VI. CHAPTER VI: DISCUSSION & FUTURE WORK	130
References	135
Appendix A. Exam Questions Targeting Multivariable Integration	142
Appendix B. Interview Questions Targeting Multivariable Integration	144
Appendix C. Exam Questions Targeting the Dirac Delta Function	145
Appendix D. Interview Questions Targeting the Dirac delta function	146
Appendix E. Exam Questions Targeting Cartesian Separation of Variables	148
Appendix F. Exam Questions Targeting Spherical Separation of Variables	150
Appendix G. Interview Questions Targeting Separation of Variables	153
Appendix H. The Coupled Multiple-Response CUE Instrument	154
Appendix I. Detailed R2 Grading Scheme for the CMR CUE	175
Appendix J. Method Selections for Q1-7 on the CMR and FR CUE	182
Appendix K. Student Responses to All Questions on the CMR CUE	186

I. CHAPTER I: INTRODUCTION

A. Motivation

Historically, one of the major focuses of the Physics Education Research (PER) community has been on understanding student learning in order to improve the ways we teach undergraduate physics courses [1]. These efforts are often accompanied by the development and evaluation of new instructional materials and/or pedagogical techniques designed to improve student learning. For the introductory physics courses, a number of materials and techniques have been developed most of which focus on incorporating various degrees of interactive engagement into physics instruction. Examples of this work span the space from small changes to the traditional classroom structure (e.g., peer instruction through the use of concept tests in lecture [2] or incorporation of research-based tutorial worksheets into the traditional recitation environment [3]) to large scale overhauls of the classroom structure (e.g., workshop physics [4]). Significant increases in student learning gains on associated introductory conceptual assessments in courses that utilize one or more of these interactive materials and/or techniques represent one measure of the effectiveness of these course transformations [4–6].

Despite the success of course transformation efforts at the introductory level, upper-division physics courses are still largely taught using only conventional lecture techniques. However, some physics faculty at the University of Colorado Boulder (CU) have expressed dissatisfaction with student learning in upper-division courses. One example of this dissatisfaction came when the CU physics faculty voted to seek funding to support research-based transformations of their core upper-division courses [7]. Their concerns were driven in part by research findings that, for example, showed no gains in student performance on an introductory electricity and magnetism conceptual assessment after completing one semester of a conventionally taught junior-level electrostatics course [8].

Responding to this push from the faculty, researchers at CU set out to transform several

of their upper-division physics courses, including junior-level electrostatics [9, 10], electrodynamics [11], and quantum mechanics [12], as well as sophomore-level classical mechanics [13]. The course transformation process was guided by an iterative design process involving three key components [7]:

- (1) Establishment of consensus learning goals through faculty collaboration
- (2) Design of curricular materials and instructional techniques to help students achieve these learning goals
- (3) Assessment of student understanding to determine which learning goals are being met

For introductory courses, step (2) of the transformation process is facilitated by the existence of a large body of research aimed at identifying persistent student difficulties with various topics in physics (see Ref. [14, 15] for a review). Moreover, considerable work has been done to characterize student problem solving at the introductory level (see Ref. [16] for a review). However, research on student difficulties at the upper-division level has only recently gained traction in the Physics Education Research (PER) community (e.g., [17–24]). The relative lack of available information on the difficulties encountered by upper-division students hinders the process of developing effective curricular materials to improve student learning.

Additionally, assessment of student learning gains in transformed introductory courses (step (3) of the transformation process) is made possible in part by the existence of validated conceptual assessments such as the Force Concept Inventory (FCI) [25] and Brief Electricity and Magnetism Assessment (BEMA) [26]. Scores on these assessments demonstrate that learning environments composed solely of traditional lecture (i.e., chalkboard lecture with occasional demonstrations and rhetorical questions) are only modestly successful at addressing known student difficulties [5]. Alternatively, learning environments which incorporate active learning techniques are significantly more successful at addressing student difficulties [17] and result in higher learning gains [5, 6]. As part of the course transformation efforts at CU, researchers recently developed a small number of conceptual assessments specifically

targeting upper-division physics content [11, 12, 27, 28]. These assessments have been used for multiple semesters at CU and other institutions. Moreover, the primarily open-ended format of these assessments has provided rich insight into student reasoning.

Another implicit goal of course transformation efforts at both the introductory and upper-division level is sustainability and wide-scale adoption [29]. Promoting the sustainability of materials and tools involved in investigating and improving student learning requires making explicit design choices that facilitate easy adoption of these tools by faculty across courses and institutions. For example, the open-ended format of the CU upper-division assessments, while providing valuable insight into student thinking, also makes them challenging and time consuming to grade and represents a significant barrier to their adoption on a large scale. If these assessments are to be used by a wide range of faculty, it may be necessary to adapt them to a more easily-graded format. Moreover, providing a scaffold that can be used to interpret and synthesize both the available research on student difficulties and personal observations may be critical part of encouraging faculty to pay explicit attention to student difficulties in the classroom.

B. Overview of the Dissertation

It is the intent of this thesis to facilitate the development and sustainability of transformed course materials and pedagogical techniques that improve student learning by addressing the following research questions:

- (1) *How do students use and reason with mathematical tools in upper-division problem solving?*
- (2) *What types of common difficulties do they encounter in the process of mathematical problem solving and can we organize these difficulties productively?*
- (3) *Is it possible to create a valid and reliable multiple-response version of a free-response, upper-division conceptual assessment that has the logistical advantages of a multiple-choice test while continuing to provide meaningful insights into student thinking?*

For the remainder of this thesis, we focus on one particular topical area: junior-level electrostatics. We have chosen to limit our focus to electrostatics for several reasons including the existence of a standard text [30] which is used by the majority of undergraduate universities across the country resulting in higher than average consistency in course content coverage. Additionally, researchers at CU have developed a suite of transformed curricular materials, including a validated assessment tool, [31] that have been used repeatedly both at CU and elsewhere.

This thesis presents the development and application of two methodological tools to address our research questions and facilitate the investigation of student learning in upper-division electrostatics courses. The first is an analytic framework for analyzing student difficulties with mathematical tools and techniques in physics. While the application of the framework described here is limited to topics and tools from electrostatics, its use could be extended to other areas of physics [32]. The second tool is an electrostatics conceptual assessment with a novel, multiple-response format that facilitates large-scale implementation. The topic coverage of this assessment targets electrostatics; however, the multiple-response format and the process of development and validation could potentially be replicated to produce conceptual assessments for other topical areas. While these two tools were developed and can be utilized independently, this thesis also highlights ways in which each informs and supports the other.

This dissertation is organized as follows. To address the challenge of characterizing student difficulties at the upper-division level, Chapter II introduces an analytical tool (the ACER framework) for investigating upper-division students' difficulties manipulating sophisticated mathematical tools in their physics courses. We summarize the existing problem-solving literature and define the underlying theoretical constructs that ground the framework. We also describe the development and validation of the framework, its general structure, and an overview of the process of operationalizing the framework for specific mathematical tools in the context of electrostatics.

Chapter III demonstrates the utility of the framework through analysis of student difficulties with three different mathematical tools: integration of continuous charge distributions (i.e., Coulomb's Law), delta functions as a method for mathematically describing volume charge distributions, and separation of variables as a technique for solving Laplace's equation. For each of these mathematical tools, we present the operationalized framework and common student difficulties identified via the framework.

Chapter IV shifts the focus to a second barrier to upper-division course transformation by describing the development and validation of a multiple-response adaptation of an existing conceptual assessment known as the Colorado Upper-division Electrostatics (CUE) diagnostic [27]. Beginning with a summary of the history, outcomes, and limitations of the original free-response instrument, we motivate the need for an easily graded version which can be utilized on a large scale. We then detail the process of crafting distractors and developing a grading scheme for the multiple-response CUE and discuss both expert and student validation.

Having established the content validity of the new instrument, Chapter V describes the statistical analysis of students' scores on the multiple-response CUE. Here, we use a variety of test statistics from Classic Test Theory to characterize the difficulty, discrimination, and consistency of scores on the new version. This chapter also reports both quantitative and qualitative findings from two direct comparisons of the free-response and multiple-response versions of the instrument. Finally, in Chapter VI, we briefly summarize the two methodological tools described in this thesis and suggest potential research directions for future work.

C. Summary of Key Findings

Here, we list the key research findings presented in this thesis loosely organized in the order that they appear in the text.

- **The ACER Framework** - ACER is a new analytical framework designed to help instructors and researchers characterize students' use of mathematical tools in upper-division physics problem solving. The framework assumes a resources perspective on the nature of knowledge, and was developed based on patterns identified in expert problem solving. The goal of ACER is to provide an organizing structure that identifies and focuses on important nodes in the complex and often iterative solutions characteristic of upper-division problem solving.
- **Student difficulties with multivariable integration** - Student difficulties with integration in the context of Coulomb's law centered primarily around constructing an integral expression for the potential that was consistent with the particular physical situation. This included expressing both the difference vector ($\vec{r} - \vec{r}'$) and the differential charge element (dq). Additional difficulties included activating the direct calculation of the potential via the integral form of Coulomb's law as the appropriate solution method and spontaneously reflecting on the final solution.
- **Student difficulties with the Dirac delta function** - Common student difficulties with the Dirac delta function in the context of expressing volume charge densities included activating the delta function as the appropriate mathematical tool, translating a verbal description of a charge distribution into a mathematical formula for the charge density, and recognizing that the delta function can have units. Our students also exhibited difficulties with the procedural aspects of integrating the delta function despite often recognizing that the delta function picks out the value of a function at a particular point.
- **Student difficulties with Cartesian separation of variables** - Student difficulties with separation of variables as a technique to solve Laplace's equation in Cartesian coordinates included recalling/justifying the separated form of the potential, applying the necessary logic to separate the PDE into ODEs, and recalling/justifying the need

for the infinite sum. Student also demonstrated a preference for formal integral procedures (i.e., Fourier's trick) over simpler algebraic manipulations (i.e., term matching) to solve for the final unknown constant(s).

- **Student difficulties with spherical separation of variables** - Common student difficulties with separation of variables as a techniques to solve Laplace's equation in spherical coordinates included activating separation of variables as the appropriate solution method, identifying all implicit boundary conditions, and spontaneously reflecting on the final solution. Unlike with Cartesian separation of variables, students demonstrated a strong preference for simpler algebraic manipulations (i.e., term matching) over formal integral procedures (i.e., Fourier's trick) to solve for the final unknown constant(s).
- **The multiple-response CUE assessment** - The coupled multiple-response (CMR) CUE diagnostic is a new assessment tool designed to help instructors and researchers assess student learning of electrostatics on a large scale. It utilizes a novel multiple-choice format that allows students to select multiple responses and awards points based on both the accuracy and consistency of their selections. We demonstrate that the CMR CUE is both valid and reliable according to multiple measures including expert and student review, item and test difficulty, item and test discrimination, and internal consistency.
- **Comparing the multiple-response and free-response formats** - Student performance on the new multiple-response version of the CUE is consistent with student performance on the free-response version. Student performance on the two formats was statistically similarly on multiple test statistics. Qualitative analysis of student responses to a subset of the questions on both tests show that while the majority of student responses to the free-response version included elements that matched the options on the CMR version, roughly a third also included elements that would not have

been captured by the CMR response options.

- **Investigating student difficulties with the CMR CUE** - The new multiple-response format preserves a significant fraction of the insight into student thinking afforded by the original free-response version. Students' responses to related questions or subparts demonstrate a high degree of consistency even when those responses are incorrect. Additionally, the multiple-response CUE can be used to quickly determine prevalence and persistence of known student difficulties within new student populations by examining patterns in students' responses.

II. CHAPTER II: DEVELOPING A FRAMEWORK

A. Introduction and Motivation

Upper-division physics content requires students to manipulate sophisticated mathematical tools (e.g. multivariable integration, approximation methods, special techniques for solving partial differential equations, etc.). Students are taught these tools in their mathematics courses and use them to solve numerous abstract mathematical exercises. Yet many students still struggle to apply mathematical tools appropriately to problems in physics. This is not necessarily surprising given that physicists use mathematics quite differently than mathematicians (e.g., to make inferences about physical systems) [33, 34]. However, persistent mathematical difficulties can undermine attempts to build on prior knowledge as our majors advance through the curriculum. Upper-division instructors face significant pressure to cover large amounts of new content, a task made more difficult by constantly having to review the relevant mathematical tools. It is often an explicit goal for advanced courses to develop students' ability to connect mathematical expressions to physics concepts. For example, consensus learning goals for upper-division courses at CU [35] include "Students should be able to translate a physical description of an upper-division physics problem to a mathematical equation necessary to solve it," and "to achieve physical insight through the mathematics of a problem." To improve student learning in advanced physics courses, we find it necessary to move away from merely noting students' conceptual difficulties towards systematically investigating how students integrate mathematics with their conceptual knowledge to solve complex physics problems.

In order to address the issues that arise when solving physics problems that rely on sophisticated mathematical tools, we must first understand how students access and coordinate their mathematical and conceptual resources. However, canonical problems in upper-division

The majority of Chapter II is taken from Ref. [32] on which BRW is first author. All text taken verbatim from this publication was written by BRW.

courses are often long and complex, and students' reasoning is similarly long and complex. Making sense of the difficulties that arise requires a well-articulated framework for analyzing students' synthesis of conceptual knowledge and mathematical tools. Here, we use the term framework to refer to a structure of guiding principles and assumptions about the underlying relationship between a physical concept and the mathematics necessary to describe it. At the upper-division level in particular, this relationship can be strongly dependent on the particular concept in question, suggesting that a useful framework needs to be adaptable to a wide variety of physical concepts and mathematical tools.

We first encountered the need for such a framework while investigating students' understanding of approximation methods (i.e., Taylor series) in a middle-division classical mechanics course [36] and with integration of continuous charge distributions (i.e., Coulomb's Law) in an upper-division electrostatics course [37]. Our initial analysis focused on identifying emergent themes in students' work. We quickly identified a multitude of common difficulties, but, beyond producing a laundry list of errors, we struggled to organize these difficulties in a productive way. This lack of coherence made it challenging to identify relationships between the difficulties and to produce actionable implications for instruction or further research.

To provide a suitable organizational structure, we developed a framework to address students' *activation* of mathematical tools, *construction* of mathematical models, *execution* of the mathematics, and *reflection* on the results (ACER). The ACER framework is a tool designed to aid both instructors and researchers in exploring when and how students employ particular mathematical tools to solve canonical problems from upper-division physics courses. Our goal is to provide a scaffold for describing student learning that is explicitly grounded in theories of learning but can still be leveraged by instructors who are not thoroughly versed in such theories.

B. Existing Problem-Solving Strategies and Theoretical Frameworks

While the investigations of the difficulties students encounter when utilizing mathematics in physics are diverse, we have identified two approaches to thinking about student problem solving that commonly appear in the literature. The first approach seeks to characterize expert physicists' use of mathematics. Such a characterization can help to produce instructional and analytical tools to align students' problem solving with that of experts'. The second approach focuses on describing students' in-the-moment reasoning, not just in terms of how it does not make sense to physicists, but in terms of how it does make sense to the students. Here, we review some of the previous research using these two approaches.

The first of these two approaches seeks to better understand the ideal landscape of the intersection between physics and mathematics. For example, Redish [33] developed an idealized model of how physicists use math to describe physical systems. He identifies four steps that guide this process (see Fig. 1):

- (1) *map the physical structures to mathematical ones*
- (2) *transform the initial mathematical structures*
- (3) *interpret the results in terms of the physical system*
- (4) *evaluate the validity of the results.*

This iterative model makes it clear that the source of students' difficulties may not be as simple as not knowing the necessary mathematical formalisms. While the intentionally broad nature of the model makes it widely applicable, we found it challenging to utilize it to identify

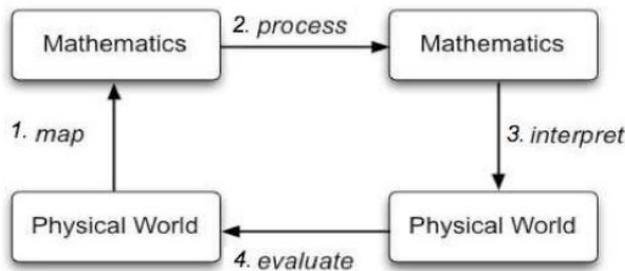


FIG. 1. An iterative representation of Redish's [33] idealized model for the use of math in science.

concrete, actionable implications for the instructor or researcher dealing with mathematical difficulties in the physics classroom.

It has been well documented that students do not approach physics problems in a manner consistent with Redish's model [33]. In fact, students often approach physics problems in a way that seems haphazard and inefficient to experts [38]. Some attempts have been made to address this at the introductory level by explicitly teaching students a problem-solving strategy that is more aligned with the expert approach. Wright and Williams [39] incorporated a problem-solving strategy (WISE) into their introductory physics course that involved four steps:

- (1) *What's happening?*
- (2) *Isolate the unknown*
- (3) *Substitute*
- (4) *Evaluation*

The WISE strategy was designed as a heuristic that physics students could use to become more efficient and accurate problem solvers.

Similarly, Heller *et al.* [40] developed a strategy to help their introductory students integrate the conceptual and procedural aspects of problem solving. This strategy included 5 steps:

- (1) *Visualize the problem*
- (2) *Physics description*
- (3) *Plan the solution*
- (4) *Execute the plan*
- (5) *Check and evaluate*

Docktor [41] modified and extended this strategy to develop a validated physics problem-solving assessment rubric. With the goal of providing consistent and reliable scores on problem-solving tasks, this rubric is scored based on five general processes: Useful Description, Physics Approach, Specific Application of Physics, Mathematical Procedures, and Logical Progression. Useful Description is the process of summarizing a problem statement by

assigning symbols and/or sketching. Physics Approach and Specific Application of Physics represent the process of selecting and linking the appropriate physics concepts to the specifics of the problem. Mathematical Procedures refers to the mathematical operations needed to produce a solution, and Logical Progression looks at the focus and consistency of the overall solution.

The strategies presented above suggest considerable agreement as to the general structure of expert problem solving as well as some indication that this structure can be used as a guide to assess student work at the introductory level. The prescriptive and linear nature of these problem-solving strategies lends itself well to the kinds of problems encountered in introductory physics. However, upper-division problems are more complex, often iterative, and less likely to respond to a prescriptive approach. Additionally, problem-solving strategies are intentionally independent of specific content so as to be generally applicable, and on their own offer limited insight into the nature of students' difficulties with specific mathematical tools.

The other approach to thinking about students' mathematical problem solving in physics focuses on explaining why students solve problems in a particular way by focusing on their in-the-moment reasoning. For example, Tuminaro [42] used videotaped problem-solving sessions with introductory students to develop a theoretical framework describing students' use of mathematics in physics. This model of student thinking blends three theoretical constructs: mathematical resources [43], epistemic games [44], and frames [45]. Mathematical resources are the abstract knowledge elements that are involved in mathematical thinking. Tuminaro [42] includes in the category of mathematical resources: a student's intuitive mathematics knowledge and sense of physical mechanism, their understanding of mathematical symbolism, and the strategies they use to extract information from equations. Epistemic games are coherent patterns of activities observed during problem solving. Each game is characterized by different sequences of moves and types of resources used by the student. The game that a student chooses to play is governed by the frame they are operating in,

which is determined by their tacit beliefs about and expectations of the kind of activity they are engaged in.

The framework presented by Tuminaro [42] was developed for introductory students and relies on students' explicit discussion of the details of their work. Upper-division students, on the other hand, tend to work more quickly and externalize fewer of their specific steps. To address this, Bing [46] leveraged the theoretical constructs of mathematical resources and epistemic framing to analyze upper-level students' use of mathematics. Epistemic framing is the students' unconscious answer to the question "What kind of activity is this?" Bing argues that a student's framing can be identified by examining the types of justifications and proof that they offer to support their mathematical claims, rather than the specific 'moves' they make. Bing identifies four epistemic frames used by upper-division students:

Invoking Authority: Students quote a rule or previously packaged result without offering further justification.

Physical Mapping: Students rely on the quality of fit between their mathematics and the physical situation at hand to justify their steps.

Calculation: Students rely on computational correctness to justify the validity of their result.

Math Consistency: Students establish relevant relationships and make inferences based on similarities in the underlying mathematical structures of different expressions.

There are several limitations to the theoretical frameworks from Tuminaro [42] and Bing [46]. To understand student work in terms of epistemic games or epistemic framing, one must have data on the students' real-time reasoning. This largely restricts the potential data sources to video and audio data, eliminating students' written work. Additionally, effective application of either framework requires considerable familiarity with the underlying theoretical constructs in PER. In practice this will prevent many instructors, particularly at the upper-division level, from productively utilizing the frameworks.

Describing experts' use of mathematics and utilizing theoretical models of students' in-the-moment reasoning are complementary aspects of understanding mathematical difficulties in physics. As described in the next section, the ACER framework leverages ideas from both in order to target students' use of mathematics in upper-division courses.

C. The ACER Framework

ACER is an analytical framework designed to guide and structure investigations of students' difficulties with the sophisticated mathematical tools used in their physics classes. When solving upper-division physics problems, students often make multiple mistakes or take unnecessary steps which must then be tracked through the solution. This undermines attempts to pinpoint the fundamental difficulties that cause the students to struggle or to identify relationships between these difficulties. The ACER framework provides an organizing structure that focuses on important nodes in students' solutions. This removes some of the "noise" in students' work that can obscure what is going on. This section provides a general overview of the framework and its development.

1. *Overview*

ACER was developed in conjunction with research into student learning of two topics in upper-division physics: Taylor series [36] and direct integration [37]. The results of applying the framework to specific topics will be discussed in detail in Chapter III; here, we present the general development and form of ACER. The ACER framework, like the frameworks presented by Tuminaro [42] and Bing [46], is fundamentally cognitive and assumes a resource view on the nature of knowledge [43].

In order to better understand students' difficulties, we performed a modified version of task analysis [47, 48] on canonical problems relating to each of the topics mentioned above. Task analysis is a method used to uncover the tacit knowledge used by experts when solving

complex problems. Our modified use of task analysis is described in greater detail in Sec. IIC2; however, the general process requires a content expert to work through the problem while documenting and reflecting on all elements of a complete solution. These elements are then discussed with several other content experts to reach consensus that all important aspects of the solution have been identified. After several iterations, we found that these various problem-specific elements could be organized into four components that appeared consistently in the solutions to a number of content-rich problems utilizing sophisticated mathematical tools. These four components are: *Activation of the tool*, *Construction of the model*, *Execution of the mathematics*, and *Reflection on the result*. Each component is described in greater detail below.

In order to solve the back-of-the-book or exam-type problems that ACER targets, one must determine which mathematical tool is appropriate (Activation) and construct a mathematical model by mapping the particular physical system onto appropriate mathematical tools (Construction). Once the mathematical model is complete, there is often a series of mathematical steps that must be executed in order to reduce the solution into a form that can be readily interpreted (Execution). This final solution must then be interpreted and checked to ensure that it is consistent with known or expected results (Reflection). The four general components are emergent from our own analysis of experts' problem solving and are consistent with previous literature on problem-solving strategies (see Sec. IIB). Though the framework suggests a certain logical flow, we are not suggesting that all experts or students solve problems in a clearly organized, linear fashion.

A convenient visualization of ACER is given in Fig. 2. The framework provides a researcher-guided outline that organizes key elements of a well-articulated, complete solution. The framework does not assign value by providing an ideal solution path towards which the students should strive. ACER is also not designed to be general enough to be applied to open-ended problems; however, its targeted focus means it can be operationalized for a variety of mathematical tools used in context-rich problems. Sec. IIC2 will discuss

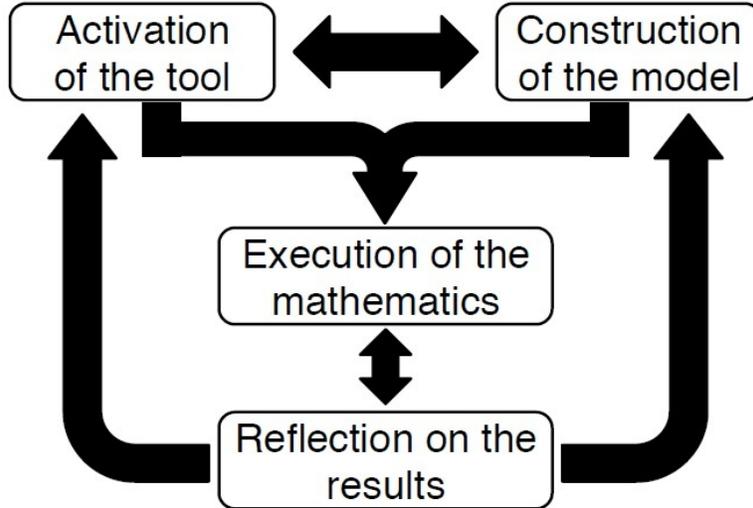


FIG. 2. A visual representation of the ACER framework. Arrows denote common feedback pathways between the components.

how ACER is operationalized for specific tools and topics. ACER is a tool for understanding and characterizing the difficulties seen in students’ work, but its structure is not meant to approximate students’ actual solutions. Instead, the general structure of ACER was developed to accommodate the complex and often iterative solution patterns characteristic of upper-division problems.

Activation of the tool: A problem statement contains a number of explicit and/or implicit cues that prime or activate different resources (or networks of resources) associated with any number of mathematical tools [43]. These cues can include the goal of the problem (e.g., calculate the potential) as well as the language and symbols used. The resources that students activate depend on the individual student and their perception of the nature of the task (i.e., their epistemic framing [46]). Operationally, evidence of Activation can manifest in several different ways. For example, a student might explicitly state, “This problem requires a Taylor expansion,” or they might simply write down the general form for a Taylor expansion. Moreover, they might skip writing the general form of the expansion and jump straight to expressing each term in the expansion appropriate for this physical situation. Any of these would be considered evidence of the Activation of Taylor series.

Construction of the model: In physics, mathematics are often used to express a sim-

plified picture (i.e., a model) of a real system. These mathematical models are typically necessary to solve physics problems. Mathematical models are generally written in a remarkably compact form (e.g., $\Delta V = -\int \vec{E} \cdot d\vec{l}$) where each symbol has a specific physical meaning that may be context dependent. Different representations (e.g., diagrammatic or graphical) are sometimes necessary to construct or map the elements of the model [33]. Operationally, Construction encompasses elements of the problem-solving process related to customizing the general form of a mathematical tool for a specific physical situation. The output from this component should be in a form that requires only executing the necessary procedural mathematical steps to arrive at a final solution.

Execution of the mathematics: In order to arrive at a solution, it is usually necessary to transform the math structures produced in the construction component (e.g., unevaluated integrals) into mathematical expressions that can be more easily interpreted (e.g., evaluated integrals). Each mathematical tool requires specific background knowledge and base mathematical skills (e.g., how to take derivatives or integrals). The mathematical manipulations performed in this component are not necessarily context-free. When employing these base mathematical skills, an expert maintains an awareness of the physical meaning of each symbol in the expression (e.g., which symbols are constants when taking derivatives or integrals) [33]. Operationally, elements of the problem-solving process that fall into Execution are those that relate to the procedural mathematics of solving a physics problem. Generally, these are steps in the solution that could, in theory, be accomplished by a computer or person with no knowledge of the specific physical situation at hand.

Reflection on the result: Solutions to problems in upper-division physics usually result in expressions that are not merely superficial manipulations of formulas provided in the textbook or notes. Instead, they are new entities that offer meaningful insight into explaining or predicting the behavior of physical systems. Reflecting on these expressions is a crucial part of understanding the system and gaining confidence in the calculation performed (e.g., how do we know an expression is the correct one?). At the most basic level, reflection

involves checking expressions for errors (e.g., checking units) or comparing predictions to established or expected results (e.g., checking limiting behavior). This kind of reflection can help to identify mistakes that occurred in the other components of the framework. Operationally, evidence for Reflection requires explicit statements or mathematical checks related to confirming the validity of the final solution or intermediate steps. While Reflection may be occurring implicitly throughout the problem-solving process, only explicit checks can be unambiguously captured by the framework.

The theoretical constructs that ground the frameworks presented by Tuminaro [42] and Bing [46] are commensurate with the implicit theoretical constructs that ground ACER. For example, a problem solver accesses different, possibly overlapping, networks of resources depending on the component of the framework in which they are working. Similarly, certain epistemic frames would be more useful than others when operating in different components. For example, Invoking Authority (see Sec. IIB) can be a valuable frame while in the Activation component. Appealing to authority (e.g., taking a formula from the book or notes) is often a good way to identify which mathematical tool to use. However, in the Construction component, when trying to map that tool to a specific problem, relying on rules or previously derived results without further justification can easily sidetrack the unwary student. Alternatively, a Physical Mapping frame would likely be productive for both the Construction and Activation components where alignment between the mathematics and the physical situation is key.

While we acknowledge the value and importance of explicitly grounding ACER in theoretical constructs like resources and epistemic frames, we have intentionally avoided explicit identification of specific resources or frames as part of the framework. In this way, it is not necessary to have a strong background in theories of learning in order to utilize ACER. Moreover, while the ideas behind resources and epistemic framing strongly impacted its development, the ACER framework is likely compatible with a number of other theoretical constructs (e.g., schema [49]). Thus, the framework still represents a potentially useful tool

for instructors and researchers with a range of theoretical perspectives on learning.

The general components of the framework are also consistent with Redish’s idealized model for the way physicists utilize mathematics [33], as well as the steps in the problem-solving strategies presented for introductory physics [38–41]. Yet, ACER goes beyond these broad descriptions by providing a mechanism to target specific topics and mathematical tools. This mechanism is described in the following section.

2. *Operationalizing the Framework*

The utility of ACER as a framework for understanding students’ use of mathematics in physics comes when it is operationalized for a specific mathematical tool. Operationalization is the process by which a particular problem or set of problems that exploit the targeted tool are mapped onto the framework. This involves identifying important elements in each component that together result in what an expert/instructor would consider a complete and correct solution.

We used a modified form of task analysis to operationalize the framework. Formally, task analysis [47, 48] is accomplished by having a subject matter expert (SME) solve problems while explaining their steps and reasoning to a knowledge extraction expert (KEE) who keeps a record. This method for uncovering the tacit knowledge used by experts has been exploited to produce example solutions designed to improve students’ ability to solve novel problems [50].

Our modified task analysis does not include a KEE. This was done because such an expert was not readily available to us, nor did we want the need for a KEE to prevent other researchers or instructors from utilizing the framework. Instead, the SME works through the problems, documenting their reasoning and mapping the vital elements of their solution onto the components of ACER. This record is then shared with several other SMEs to ensure that all important aspects of the solution are accounted for. Additionally, these experts come to a consensus in classifying each element into a specific component (i.e., Activation,

Construction, Execution, or Reflection). These preliminary elements are then applied to student work and the operationalized framework is refined to accommodate patterns of student reasoning not present in the SMEs solutions.

Our motivation for removing the KEE was entirely practical in origin; however, not utilizing a KEE may have implications for the theoretical foundations of our modified task analysis. The KEE, as a content novice, helps to force the SME to fully and clearly justify their steps even when they include decisions based on procedural and declarative details the SME no longer thinks about [47]. Removing the KEE from the task analysis process makes it more difficult to ensure that the important elements identified in the solution are complete from the point of view of a novice as well as an SME. For this reason it is important that the operationalized ACER framework which is produced by the modified task analysis remains flexible to modification based on emergent analysis of student work.

The following chapter describes the operationalization of the framework for the use of three different mathematical tools and techniques in upper-division electrostatics: multi-variable integration in the context of Coulomb's law, the Dirac delta function as a tool to express volume charge densities, and separation of variables as a technique for solving Laplace's Equation.

III. CHAPTER III: APPLICATION OF THE FRAMEWORK

A. Context and Data sources

Data for these studies were largely collected from the first half of a two-semester Electricity and Magnetism sequence at the University of Colorado Boulder (CU). This course, called E&M 1, typically covers electrostatics and magnetostatics (chapters 1-6 of Griffiths [30]). The student population is composed of junior and senior-level physics, astrophysics, and engineering physics majors with a typical class size of 30-70 students (roughly 20% female). At CU, E&M 1 is often taught with varying degrees of interactivity through the use of research-based teaching practices including peer instruction using clickers [2] and optional out-of-class or in-class tutorials [10].

We collected data from up to three distinct sources for each of the investigations described in this chapter: student solutions to instructor designed questions on traditional midterm exams, responses to specific questions on the multiple-response Colorado Upper-division Electrostatics (CUE) Diagnostic [51], and think-aloud student interviews. Interviews were videotaped and students' written work was captured with embedded audio. Interviewees were paid volunteers who responded to an email request for research participants. All interviewees had successfully completed E&M 1 one to two semesters prior.

Students' exam solutions were analyzed by identifying each of the key elements from the operationalized framework (see Sec. III B 2) that appeared in the students' solutions. Each element was then further coded to identify the types of steps made by students. These codes represented emergent themes in the students' work and were not predetermined by the framework. This emergent coding helped to ensure that the expert-guided framework did not miss important but unanticipated aspects of student solutions. The interviews were similarly analyzed by classifying each of the student's major moves into one of the four components of the framework. Participants in all three studies demonstrated a wide range of abilities and received course scores ranging from A to D. In this way, exams provided

quantitative data identifying common difficulties and interviews offered deeper insight into the nature of those difficulties.

B. Multivariable Integration in the Context of Coulomb's Law

1. *Previous Research on Student Difficulties with Integration*

Integration is a ubiquitous mathematical tool in the undergraduate physics curriculum, and student difficulties with integration have been the focus of a number of investigations by both math and physics education researchers. Many of these studies have highlighted students' conceptual understanding of integrals. Orton [52], for example, interviewed 110 upper-level high school math students and pre-service math teachers' in Britain on 18 integration tasks. He documented a wide variety of issues, including difficulties recognizing that the exact value for the area under the curve could be achieved through the limit of a Riemann sum and justifying why particular steps were necessary when calculating integrals.

Thompson and Silverman [53] discussed student understanding of integrals as accumulation functions. They noted that integrals can be conceptualized both as the accumulation of infinitesimal amounts of multiplicative bits and as the area under the curve of a function over a particular interval. These two ideas can be thought of as equivalent only when a student has internalized the idea that the area under the curve can represent something other than area (e.g., the area under a force vs. distance curve represents work). Similarly, Bajracharya *et al.* [54] documented instances in which the area-under-the-curve concept interfered with students' interpretation of definite integrals with negative values; these students rejected the idea of 'negative area'. Nguyen and Rebello [55] also investigated students' understanding of the area-under-the-curve concept specifically in the context of introductory physics. They found that few students spontaneously connected the work done on a projectile to the area under the force curve. They also found that many students were not able to select the ap-

The majority of Section III B is taken from Ref. [32] on which BRW is first author. All text taken verbatim from this publication was written by BRW.

appropriate graph for which the area under the curve represented the value of a predetermined integral.

Other studies have focused on student understanding of differentials in the context of integration. Artigue *et al.* [56] conducted an interdisciplinary investigation of student ideas about what the differential represents and when it is useful. They found that, while students were able to recognize problems requiring the use of a differential (e.g., dz), they often could not clearly justify its use in terms of non-linear functional dependence. They also found two common extremes in student conceptions of the differential. Some students attributed no physical meaning to the differential except that it indicates the variable of integration, while others described it as a ‘little piece’ of something (e.g., a small piece of a wire or a small length). Hu and Rebello [57] also investigated student ideas about differentials through detailed analysis of group interviews with students. They identified 4 specific resources and 4 conceptual metaphors used by introductory students when working with differentials and suggest that some resources and metaphors may be more productive when solving certain types of physics problems.

Additional research around students’ use of integrals in physics has looked more broadly at how students solve problems involving integration. Hu and Rebello [58] utilized the idea of conceptual blending to understand how students combine their math and physics knowledge to set up integrals in physics. They found that their students used differing blends of three mental spaces relating to symbolic representation, math notation, and physics when setting up integrals with varying levels of success. Similarly, Khan *et al.* [59] identified different strategies that introductory students used when solving integrals in symbolic and graphical representations. They found that, while students employ a number of different strategies to solve integrals, after both first and second semester calculus-based physics, 40-70% of students did not recognize the need for an integral to find the quantity of interest.

The studies described above focus primarily on students conceptual difficulties with single-variable integration at the level required for introductory physics. In the study described in

the remainder of this section, we focus instead on multivariable integration in the context of the integral form of Coulomb’s law.

2. *Operationalizing ACER for Multivariable Integration*

Determining the electric potential or electric field from a continuous charge distribution using the integral form of Coulomb’s law is one of the first topics that upper-division students encounter in junior-level electrostatics. For the remainder of the paper, we use Coulomb’s Law to refer to the integral equation allowing for direct calculation of the electric field or potential from a continuous charge distribution.

$$E(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V \frac{dq}{|\vec{z}|^2} \hat{z} \quad (1)$$

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V \frac{dq}{|\vec{z}|} \quad (2)$$

Here, dq represents the differential charge element and \vec{z} is the difference vector $\vec{r} - \vec{r}'$ between the source and the observation location (i.e., Griffiths’ *script-r* [30]). In this case, the ‘tool’ we refer to is integration, and we describe its application to problems determining the potential or electric field from an arbitrary, static charge distribution via Coulomb’s Law. We will focus here only on charge distributions that cannot easily be dealt with using Gauss’s Law.

As an initial step in the process of operationalizing the ACER framework for multivariable integration in the context of Coulomb’s law, we generated and refined a network diagram illustrating the conceptual and procedural resources that experts use to solve these types of problems. Fig. 3 shows this network diagram organized according to the components of the ACER framework. The creation of a network diagram helps to fully externalize the important aspects of the problem-solving process and how the aspects are connected; however, the diagram in Fig. 3 is complex and can be difficult to parse. Therefore, the next

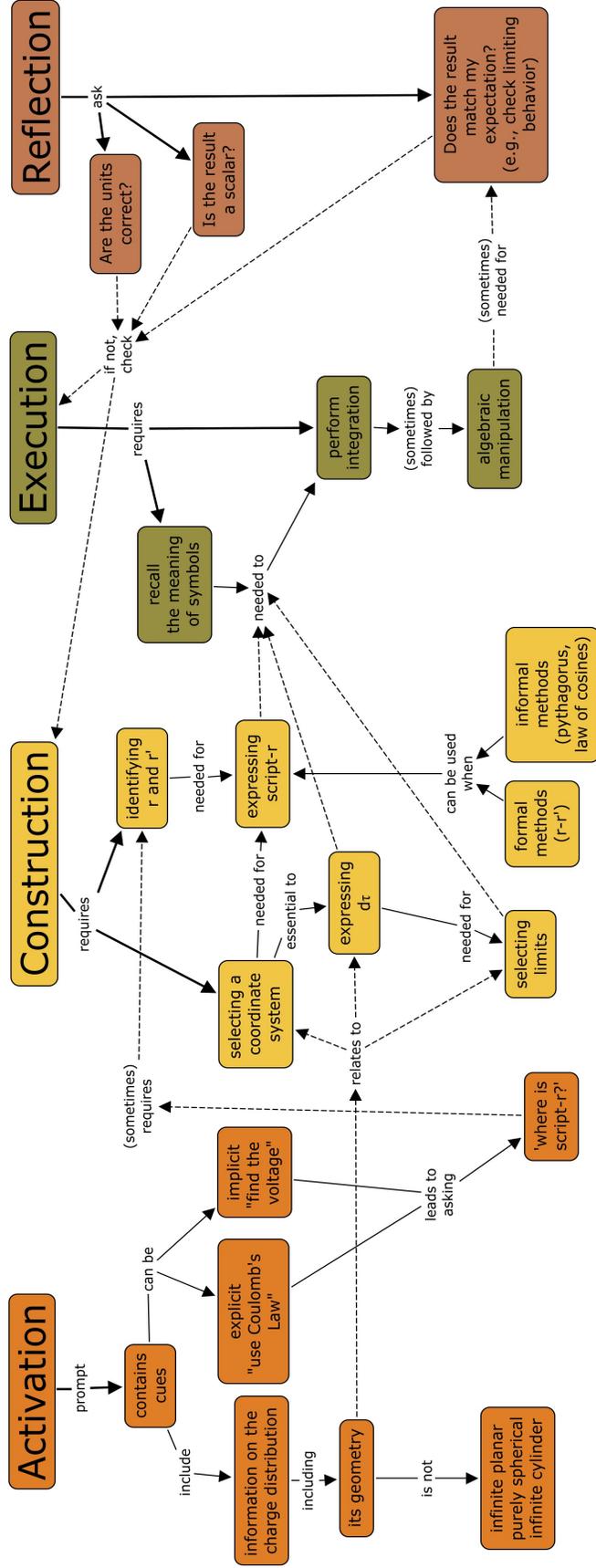


FIG. 3. Network diagram documenting and organizing the conceptual and procedural resources used by experts when solving physics problems requiring the using of the integral form of Coulomb's law. Each element is color coded to match the corresponding component of the ACER framework (color online), and dashed arrows mark connections between the components.

step in the operationalization process involves translating this diagram into a more concise list of the key elements of a complete and correct solution, referred to as the flattened framework. The flattened ACER framework for direct integration (i.e., Coulomb's law) is given below.

Element codes are for labeling purposes only and are not meant to suggest that there is a single correct order to the solution, nor are all elements always involved for any given problem. The alphanumeric code for each element of the framework provides topic, component, and element; for example, the third element in the Activation component for the use of integration is coded IA3. This convention will be used throughout the remainder of the thesis.

Flattened ACER framework

Activation of the tool: The first component of the framework involves the selection of a solution method. The modified task analysis identified four elements involved in the activation of resources identifying direct integration (i.e., Coulomb's Law) as the appropriate tool.

IA1: The problem asks for the potential or electric field.

IA2: The problem gives a charge distribution.

IA3: The charge distribution does not have appropriate symmetry to productively use Gauss's Law.

IA4: Direct calculation of the potential is more efficient than starting with the electric field.

Elements IA1–IA3 are cues typically present in the problem statement. Element IA4 is specific to problems asking for the electric potential and is included to account for the possibility of solving for potential by first calculating the electric field. This method is valid but often more difficult.

Construction of the model: Here, mathematical resources are used to map the specific physical situation onto the general mathematical expression for Coulomb's Law. The resulting integral expression should be in a form that could, in principle, be solved with no

knowledge of the physics of this specific problem. We identify four key elements that must be completed in this mapping.

IC1: Use the geometry of the charge distribution to select a coordinate system.

IC2: Express the differential charge element (dq) in the selected coordinates.

IC3: Select integration limits consistent with the differential charge element and the extent of the physical system.

IC4: Express the difference vector, \vec{z} , in the selected coordinates.

Elements IC2 and IC4 can be accomplished in multiple ways often involving several smaller steps. In order to express the differential charge element, the student must combine the charge density and differential to produce an expression with the dimensions of charge (e.g., $dq = \sigma dA$). Construction of the difference vector often includes a diagram that identifies vectors to the source point, \vec{r}' , and field point, \vec{r} .

Execution of the Mathematics: This component of the framework deals with the mathematics required to compute a final expression. In order to produce a formula describing the potential or electric field, it is necessary to:

IE1: Maintain an awareness of which variables are being integrated over (e.g., r' vs. r).

IE2: Execute (multivariable) integrals in the selected coordinate system.

IE3: Manipulate the resulting algebraic expressions into a form that can be readily interpreted.

Reflection on the result: The final component of the framework involves verifying that the expression is consistent with expectations. While many different techniques can be used to reflect on the result, these two checks are particularly common:

IR1: Verify that the units are correct.

IR2: Check the limiting behavior to ensure it is consistent with the total charge and geometry of the charge distribution.

Element IR2 is especially useful when the student already has some intuition for how the potential or electric field should behave in the limits. However, if they do not come in with this intuition, reflection on the results of this type of problem is a vital part of developing it.

In the next section, we will apply this operationalization of ACER to investigate student work on a canonical electrostatics problem (Fig. 4).

3. *Interview and Exam Prompts*

At CU, E&M 1 students are exposed to the Coulomb's Law integral for the electric field (Eqn. 1) before the analogous expression for the electric potential (Eqn. 2). However, the vector nature of the electric field makes Eqn. 1 significantly more challenging to calculate, and historically, instructors at CU tend to ask students to compute the potential on exams. The exam problem examined here asked students to calculate the electric potential along an axis of symmetry from a disk with charge density $\sigma(\phi)$ (Fig. 4). We selected this problem because it is a recognizable Coulomb's law question which requires integration and has been asked on the first midterm exam in E&M 1 for multiple semesters.

Exams were collected from four semesters of the course ($N=172$), each taught by a different instructor. Two of these instructors were physics education researchers involved in developing the transformed materials and two were traditional research faculty. All four semesters

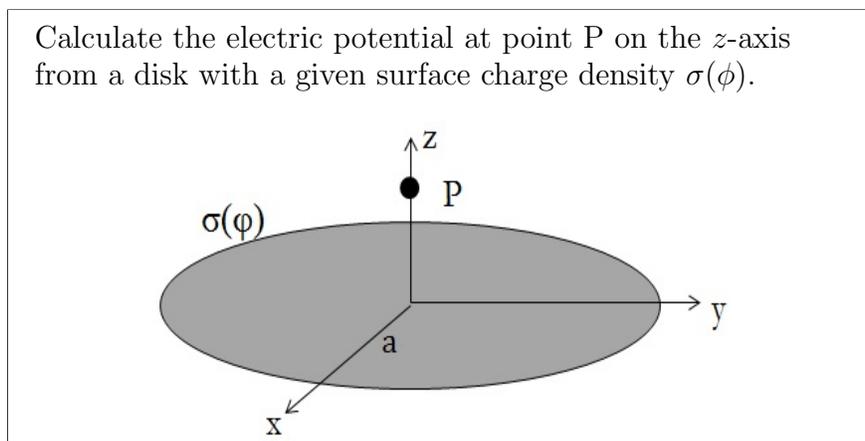


FIG. 4. An example of the canonical exam problem on continuous charge distributions.

utilized some or all of the available transformed materials. The exact details of the disk question, while similar, were not identical from semester to semester (see Appendix A for exact prompts). One of the PER faculty asked the students to sketch the charge distribution and then to calculate an expression for the potential on the z -axis (as in Fig. 4). The other PER faculty asked the students to calculate the total charge on the disk but only required them to set up the expression for the potential on the x -axis as the resulting integral cannot be solved easily by hand. Both non-PER faculty asked for the total charge on the disk first and then for the potential on the z -axis.

Interview data came from two sets of think-aloud interviews ($N=10$), performed approximately 1 year apart on different sets of students (see Appendix B for full interview protocols). The first set of interviews was structured to probe the preliminary difficulties identified in the student exams. The students were asked to calculate the potential from two parallel disks of charge by direct integration, and they were provided with a diagram of the charge distribution and Eqn. 1 and 2. In terms of the ACER framework, this prompt completely bypassed the Activation component. Also, while the first interview protocol offered important insight into how students spontaneously reflect (or not) on their solutions, it provided no explicit probe of the Reflection component. The second interview protocol specifically targeted Activation by asking students to find the potential along the z -axis outside a spherical shell with non-uniform charge density $\sigma(\theta)$ without providing a diagram or prompting them to solve the problem in any specific manner. An additional question targeted Reflection by asking students to determine which of three expressions could represent the potential from a static, localized charge distribution with total charge Q (see Fig. 5).

4. *Student Difficulties with Multivariable Integration*

Here, we present the identification and analysis of common student difficulties with Coulomb's Law integrals organized by component and element of the operationalized ACER framework (See Sec. III B 2).

Which of these expressions could represent the potential from a static, localized charge distribution with charge Q ? Here, d is a characteristic distance scale and a , b , and c are constants with undetermined units.

1. $\frac{a}{2\epsilon_0}(\sqrt{r^2 + d^2} - r)$
2. $\frac{b}{4\pi\epsilon_0} \frac{d}{r^2}$
3. $\frac{c}{2\epsilon_0} \frac{d}{r}$

FIG. 5. Three equations presented in the second interview set to target Reflection. Students must determine the units of a , b , and c . The prompt has been paraphrased; see Appendix B for full prompt.

Activation of the tool: Roughly three-quarters of our students (73%, $N=125$ of 172) correctly approached the exam question using Eqn. 2. The remaining students (27%, $N=47$ of 172) attempted to calculate the potential by determining \vec{E} , either by Gauss’s Law or Eqn. 1, and then taking the line integral (i.e., missing elements IA3 and IA4). Rather than stemming primarily from a failure to recall Eqn. 2, we argue below that this difficulty likely originated from a failure to reject these other methods.

Identifying evidence of Activation in the exam solutions was challenging because students did not typically write out their thought process as they began the problem. In particular, there was rarely explicit evidence that the students attended specifically to IA1 and IA2 (i.e., the prompt asked for potential and provided information on the charge distribution). However, we did not see students attempting to calculate quantities unrelated to the potential or attempting to utilize methods inconsistent with the information provided.

More easily identified was element IA3, which eliminates Gauss’s Law as a valid approach. Approximately a tenth of our students (11%, $N=19$ of 172) attempted to employ Gauss’ Law to solve for \vec{E} and then to calculate V by taking line integral. These students often justified their answers with comments such as, “Since we want the voltage at a point outside the disk, the E-field we use will appear to be that of a point charge at the origin.” This inappropriate use of Gauss’s Law is consistent with previous research at the junior-level [60].

Interestingly, none of the students in the single semester (N=25) that were asked to sketch the charge distribution rather than to calculate total charge attempted to use Gauss's Law. This suggests that calculation of the total charge may have activated resources associated with Gauss's Law.

The misapplication of Gauss's Law was also the primary issue observed in the interviews. Even when the students were explicitly prompted to use direct integration, one of five students still attempted to use Gauss's Law. Two students in the second set of interviews explicitly considered using Coulomb's Law but rejected it in favor of using Gauss's Law or the expression for E from a point charge. ACER states that there are a number of cues (elements IA1–IA3) embedded in the prompt of a physics problem that can guide a student to the appropriate solution method. For example, if the prompt provides a boundary condition rather than a charge distribution, this is likely to cue the student to use separation of variables or method of images. Elements IA1 and IA2 are identical for questions that can be solved by Gauss's Law and Coulomb's Law (i.e., it asks for V or E and provides $\rho(\vec{r}')$). However, our students tend to be more comfortable with Gauss's Law (i.e., their Gauss's Law resources are easily activated); therefore, they must first reject Gauss's Law as appropriate before they will attempt to use Coulomb's Law.

Even without Gauss's Law, it is still possible to solve for V by first calculating E using Eqn. 1, but this calculation requires considerably more work (element IA4). Indeed, of the students who attempted this method (15%, N=26 of 172) only a few (N=3) completed the exam problem successfully. One virtue of the electric potential in electrostatics is to allow for easier calculation of the electric field via $\vec{E} = -\vec{\nabla}V$. However, the students may have jumped to calculating V from E because they were exposed to \vec{E} first and resources associated with the electric field were more easily activated. This difficulty was not observed in the interviews.

Construction of the model: For Coulomb's Law integrals, the largest number of common student difficulties appeared in the Construction component, particularly when ex-

pressing the differential charge element and difference vector (elements IC2 and IC4). These difficulties cannot be explained purely by students failing to conceptualize the integral or lacking the mathematical skills to set up integrals over surfaces and perform vector subtractions. Rather, students had trouble keeping track of the relationships between various quantities as they adapted the deceptively simple general formula (Eqn. 2) to a specific physical system.

Almost all of the exams (97%, N=166 of 172) contained elements from the Construction component (i.e., the student did more than just write down the equation). Of these students, only two did not use the appropriate coordinates (cylindrical), indicating that students at this level are adept at selecting appropriate coordinate systems in highly symmetric problems (element IC1). Similarly, only one of the interview participants started with an inappropriate coordinate system, and this student eventually switched after attempting the problem in Cartesian coordinates. This finding is somewhat surprising given prior research indicating that even middle-division physics students often have a strong preference for Cartesian coordinates [61].

The remaining elements of Construction proved more challenging. Nearly half the students (42%, N=69 of 166) had difficulty expressing the differential charge element (element IC2) and some (14%, N=23 of 166) failed to provide limits of integration or gave limits that were inconsistent with their differential (element IC3). The most common errors made while expressing the differential charge element (dq) were (see Table I): performing the integration over a region of space with zero charge density, using a differential with the wrong units, and plugging in total charge instead of charge density.

Initially, we interpreted difficulties with dq as a failure to conceptualize Eqn. 2 as a sum over each little ‘bit’ of charge. Previous research on student difficulties with the concept of accumulation as it applies to definite integrals supports this interpretation [53]. However, the interviews suggest that the problem was more subtle than that. Even those interviewees who failed to produce an appropriate expression for dq made statements or gestures indicating

TABLE I. Difficulties expressing the differential charge element (dq). Percentages are of just the students who had difficulty with dq (42%, N=69 of 166). Codes are not exhaustive or exclusive but represent the most common themes, thus the total N in the table need not sum to 69.

Difficulty	N	Percent
Not integrating only over charges e.g., $dq = \sigma dr dz rd\phi$	37	54
Differential with the wrong units e.g., $dq = \sigma dr d\phi$	23	33
Total charge instead of charge density e.g., $dq = Q_{tot} dr rd\phi$	10	14

they understood the integral to be a sum over the charge distribution. Additionally, post-test data from the classical mechanics course at CU shows that more than 80% of our students can correctly determine the differential area element for a cylindrical shell one semester prior to taking E&M. Thus the problem appeared to be neither that the students were not conceptualizing the integral as a sum over the charges, nor that they could not construct a differential area element. Instead, the difficulties appeared when students were asked to apply these two ideas simultaneously to produce an expression for dq consistent with a specific charge distribution.

The magnitude of the difference vector, $|\vec{z}|$, must also be expressed such that it is consistent with the specific charge distribution (element IC4), and most students (86%, N=148 of 172) attempted to do so. About half of these (47%, N=69 of 148) were unable to produce a correct formula for $|\vec{z}|$. The most common errors included (see Table II): using a magnitude appropriate for a ring of charge, setting the magnitude equal to the distance to the source point (r'), setting the magnitude equal to the distance to the field point (r), and never expressing the magnitude in terms of given variables or quantities. It was often difficult to distinguish between the middle two difficulties because students' notation rarely distinguished clearly between the source and field variables; these issues are combined in Table II. The remaining students were distributed over a variety of distinct, but not widely-represented issues.

TABLE II. Difficulties expressing the magnitude of the difference vector (\vec{z}). Percentages are of just the students who had difficulties with \vec{z} (47% of 148, N=69). Codes are not exhaustive but represent the most common themes, thus the total N in the table need not sum to 69.

Difficulty	N	Percent
Ring of charge i.e., $ \vec{z} = \sqrt{a^2 + r'^2}$	27	39
Distance to source or field point i.e., $ \vec{z} = r$ or $ \vec{z} = r'$	17	25
No expression for $ \vec{z} $	8	12

Students' spontaneous use of diagrammatic representation may be an additional aspect of the Construction component. For example, drawing the vectors \vec{r} , \vec{r}' , and \vec{z} can be a helpful step towards a correct expression for $|\vec{z}|$. We found that about two-thirds our students (66%, N=98 of 148) drew one or more of these vectors on the exams; however, only half of these students (50%, N=49 of 98) made explicit use of this diagram in their solution. It may be that our students have seen enough of these types of problems to know that they should draw a diagram but have not internalized how to use it productively.

Six of the eight interview participants who used Coulomb's Law also spontaneously drew the difference vector, and a seventh drew the vector but did not explicitly identify it as \vec{z} . However, even those students who were able to articulate the difference vector as the distance between the source and field point struggled to produce a useful expression for it. Only one interview participant arrived at a correct expression for the difference vector while the others were either unable to express $|\vec{z}|$ or treated it as a single variable like r or r' . The greater degree of difficulty with \vec{z} observed in the interviews may be due to the time delay between the participants completing the course and sitting for the interview.

Using Griffith's "script-r" notation, rather than $\vec{r} - \vec{r}'$, has a number of advantages including making Coulomb's law for continuous charge distributions look very similar to Coulomb's law for a point charge. However, it may be that this notation also encourages students to look at \vec{z} as a separate entity that they must remember rather than a quantity they construct. In fact, most students made comments in the interviews about not remembering the formula

for \vec{z} or which direction it pointed, and few even attempted to use the source and field point vectors to answer these questions. Only three of the eight interviewees spontaneously drew \vec{r} and \vec{r}' , suggesting that the “script-r” notation obscured the importance of these two vectors. Failure to properly distinguish between z , r , and r' often resulted in improper cancellations in the Execution component.

Execution of the mathematics: Given the high pressure and individual nature of both exams and interviews, we expected that many students would make mathematical errors particularly with element IE3. Yet our data offer no evidence that mathematical errors either with integrals or algebraic manipulations (elements IE2 or IE3) were specific to solving Coulomb’s Law problems nor that they represented the primary barrier to student success on these problems. More than half the student exam solutions (60%, N=103 of 172) contained elements from the Execution component. The significant reduction in number was due primarily to the one of the four classes (N=55) that was only asked to set up the integral for V. Additionally, not all students progressed far enough in their solutions to actually evaluate integrals.

We were not able to produce a quantitative measure of student difficulties with element IE1 from the exams because the majority of students did not consistently distinguish between source and field variables (i.e., r vs. r'). However, of the four interview participants who made a distinction between the source and field point, none consistently used the primed notation. Two of these students ended up integrating over the r variable as if it were r' .

Overall, half the students’ exams containing elements of Execution (51%, N=53 of 103) made various mathematical errors while solving integrals or simplifying their expression algebraically (elements IE2 and IE3). Roughly half of the students with mathematical errors (49%, N=26 of 53) made only slight mathematical errors, such as dropping a factor of two or plugging in limits incorrectly. The remaining students (51%, N=27 of 53) made various significant mathematical errors, such as pulling integration variables outside of integrals or not completing one or more integrals. Only a small fraction of exam students (8%, N=8

of 103) had difficulties *only* with elements in the Execution component (i.e., no mistakes in Activation or Construction). Similar trends were observed with the seven interview participants who attempted to complete one or more calculations. Four students made significant mathematical errors, two made only slight mathematical errors, and one made no errors.

Reflection on the result: In many cases, mistakes in the Construction or Execution component resulted in expressions for the potential which had the wrong units and/or limiting behavior (elements IR1 and IR2). While our students were able to identify these checks as valuable when explicitly prompted, we found that they rarely spontaneously check these properties to gain confidence in their solutions.

Only a small number of students (8%, N=13 of 172) made explicit attempts to check their work on exams and almost exclusively by checking limiting behavior. While it is possible that a greater number of students did perform one or more checks (i.e., elements IR1 and IR2) but simply did not write them out, the interviews suggest this is less likely. When they had not been prompted to check or reflect on their solutions, half of the interview participants made no attempt to do so. Two of the remaining students only made superficial comments about being uncertain if their solution was correct. One stated that her answer did not make sense but was not able to leverage this realization to correct her earlier work. The final two students both mentioned checking the units of their solutions, though not recalling the units of ϵ_o prevented one of them from actually doing so.

The second set of interviews explicitly targeted Reflection by directly asking the students to determine if three formulas (Fig. 5) could represent the potential from a static, localized charge distribution with positive total charge Q . All five students suggested checking the units of these expressions, yet all but one had difficulty doing so because they did not recall the units of ϵ_o . This may be part of why units checks were not more common in the exam solutions as well. Eventually, all the students were able to execute a units check once shown a method for getting around the units of ϵ_o by considering the formula for the potential of a point charge. Additionally, all five students suggested checking that in the limit as $r \rightarrow \infty$

the potential went to zero. Only two students spontaneously argued that V would need to fall off as $\frac{1}{r}$. The other three made this argument when their attention was specifically drawn to the fact that the charge distributions was localized and had positive total charge.

One of the three expressions for V required an appropriate Taylor expansion in order to determine its behavior at large r (i.e., expression 1 of Fig. 5). Only one of the five students recognized the need for an expansion without prompting. Another three argued that the expression clearly did not fall off like a point charge. However, when directed to Taylor expand, all three were able to manipulate the expression in order to isolate the small quantity and determine the leading term in the series.

5. *Summary and Implications*

We found that our junior-level students tended to encounter two broad difficulties which inhibited them from successfully solving for the potential from a continuous charge distribution using Coulomb's Law. First, students struggled to activate direct integration via Coulomb's Law as the appropriate solution method. In particular, some students tried to calculate the potential by first calculating the electric field by Gauss's Law or Coulomb's Law. For instructors, this suggests that presentation of Eqn. 2 should be accompanied by explicit emphasis on *when* and *why* Gauss's Law cannot be used as well as the utility of calculating the electric potential rather than the electric field. The latter should be aimed at helping students to develop strong connections between the conceptual idea of the potential and various mathematical formula which allow them to calculate $V(\vec{r})$. Second, students had difficulty coordinating their mathematical and physical resources to construct an integral expression for the potential which was consistent with the particular physical situation, specifically when expressing the differential charge element, dq , and difference vector, \vec{z} . Instructors may be able to help by highlighting the relationships between these quantities to encourage students to view Eqn. 2 a as coherent whole rather than a conglomeration of disconnected pieces. We also found that while our juniors were capable of correct and meaningful reflection

when explicitly prompted, very few executed these reflections spontaneously. We consider the ability to translate between physical and mathematical descriptions of a problem and to meaningfully reflect on or interpret the results as two defining characteristics of a physicist, yet these are areas where our students struggled most when manipulating Coulomb’s Law integrals.

Our investigation of student difficulties was highly focused on the use of multivariable integration in a specific physics context; as such, our findings do not always directly connect to findings from previous research (Sec. III B 1). For example, much of the existing literature focuses on the area-under-the-curve concept [52–55]. However, we argue that the area-under-the-curve concept may not be a productive way of conceptualizing multivariable integrals, particularly 3D integrals. Alternatively, conceptualizing multivariable integrals as the accumulation of infinitesimal amounts of multiplicative bits, as discussed by Thompson and Silverman [53], is likely to be a much more productive strategy particularly in terms of Construction. Similarly, the finding from Artigue *et al.* [56] that some students attributed no physical meaning to the differential except to indicate the variable of integration offers a potential explanation for our finding that some of our students did not construct expressions for the differential charge element that were consistent with it representing a little bit of charge. Moreover, the conceptual blending framework discussed by Hu and Rebello [58] could offer a valuable analytical tool for identifying productive and unproductive strategies used by students when operating the Construction component.

C. The Dirac Delta Function in the context of Electrostatics

1. *Previous Research on Student Difficulties with the Dirac Delta Function*

Another mathematical tool that appears repeatedly throughout the undergraduate physics curriculum is the Dirac delta function (hereafter referred to as simply, the δ -function).

The majority of Section III C is taken from Ref. [62] on which BRW is first author. All text taken verbatim from this publication was written by BRW.

The δ -function represents an interesting topic to investigate student difficulties for several reasons. Despite often being discussed as a purely abstract, mathematical construct, it is rarely used in purely mathematical contexts. In fact, physicists often introduce the δ -function into a problem in order to describe or model a concrete physical system. Additionally, it is our experience that the δ -function is often perceived by experts as trivial to manipulate, and is often introduced to simplify the mathematics of a problem. However, we still observe common and persistent student difficulties using the δ -function. Moreover, while the δ -function has appeared in literature around student difficulties with Fourier transforms and measurement in quantum mechanics [63–66], we are unaware of any existing literature directly targeting student difficulties with the δ -function.

2. Operationalizing ACER for the Dirac Delta Function

A physics student may be exposed to the δ -function at several different points in the undergraduate curriculum. These include its use as a tool to express volume mass or charge densities, to describe potential energies or wavefunctions in quantum mechanics, in the context of Fourier transforms, and, for more advanced situations, in the context of Green’s functions. In mathematics courses, the δ -function is rigorously defined as a distribution [67], and is typically not seen by students until senior or masters level analysis courses. Such courses are not often taken by physics majors. Here, we will focus almost exclusively on the δ -function as a tool to describe volume charge densities in the context of upper-division electrostatics as this is one of the earliest and perhaps simplest uses. Despite the general sense of many faculty at CU that students will be familiar with the δ -function by the time they reach the junior-level, we find that for some (arguably many) students junior-level electrostatics will be their first exposure to the δ -function.

To guide our data collection and analysis, we operationalized the ACER framework for problems involving the use of δ -functions to express the volume charge (or mass) densities of 1, 2, and 3D charge distributions. For example, the volume charge density of a line charge

running parallel to the z -axis and passing through the point $(1,2,0)$ can be expressed as $\rho(\vec{r}) = \lambda\delta(x - 1)\delta(y - 2)$, where λ is a *unitful constant* representing the charge per unit length. Expressing volume charge densities in this way is often necessary when working with the differential forms of Maxwell's Equations and can facilitate working with the integral forms of both Coulomb's Law and the Biot-Savart law. Given its many possible uses, we do not claim that the research presented here will span the space of all possible difficulties with the Dirac δ -function; however, it will provide a sampling of the kinds of challenges that students face when manipulating δ -functions.

Our network diagram expressing the conceptual and procedural resources used by experts when solving these types of δ -functions questions is given in Fig. 6. The elements of the flattened ACER framework extracted from this network diagram are detailed below. Recall, element codes are for labeling purposes only and are not meant to suggest a particular order, nor are all elements always necessary for every problem.

Flattened ACER framework

Activation of the tool: The first component of the framework involves identifying δ -functions as the appropriate mathematical tool. We identified two elements in the form of cues present in a prompt that are likely to activate resources associated with δ -functions.

DA1: The question provides an expression for volume charge density in terms of δ -functions

DA2: The question asks for an expression for the *volume* charge density of a charge distribution that includes point, line, or surface charges

We include element DA1 because, in electrostatics, δ -functions are often provided explicitly in the problem statement, effectively short-circuiting Activation.

Construction of the model: Elements in this component are involved in mapping the mathematical expression for the charge density to a verbal or pictorial representation of the charge distribution or vice versa.

DC1: Relate the shape of the charge distribution to the coordinate system and number of δ -functions

DC2: Relate the location of the charges with the argument(s) of the δ -function(s)

DC3: Establish the need for and/or physical meaning of the *unitful constant* in front of the δ -function

For problems that also require integration of the δ -function (e.g., to find total charge from $\rho(\vec{r})$) there are additional elements in construction related to setting up this integral.

DC4: Express a differential volume element consistent with the geometry of the charge distribution

DC5: Select limits of integration consistent with the differential volume element and region of interest

Alternatively, for sufficiently simple charge distributions, one can bypass setting up and computing integrals (elements DC4 & DC5) by using the physical meaning of the unitful constant to simply state the total charge (e.g., $Q(\text{uniform, spherical shell}) = \sigma * 4\pi R^2$).

Execution of the mathematics: This component of the framework deals with elements involved in executing the mathematical operations related to the δ -function. Because this component deals with actually performing mathematical operations, these elements are specific to problems requiring integration of the δ -function.

DE1: Execute (multivariable) integrals that include one or more (potentially multidimensional) δ -functions

When the results of the integrals in DE1 must be simplified for interpretation, Execution would include a second element relating to algebraic manipulation; however, none of the integrals included in this study elicited or required significant algebraic manipulation.

Reflection on the result: The final component includes elements related to checking and interpreting aspects of the solution, including intermediate steps and the final result. While many different techniques can be used to reflect on a physics problem, the following three are particularly common when dealing with δ -functions.

DR1: Check/determine the units of all relevant quantities (e.g., Q , ρ , the unitful constant)

DR2: Check that the physical meaning of the unitful constant is consistent with its units and the units of all other quantities

DR3: Verify that the value of the charge in a region is consistent with expectations

While the first two elements are similar, we consider element DR2 to be a distinct and potentially more sophisticated reflection task in that it is seeking consistency between the student's physical interpretation of the unitful constant and its units. The distinction between elements DR1 and DR2 was motivated in part by preliminary analysis of student work, which demonstrated that the link between the units of the constant and its physical interpretation may not be made spontaneously by many students (see Sec. III C 4).

In the next section, we will apply this operationalization of ACER to investigate student work on canonical δ -function questions.

3. Interview and Exam Prompts

Midterm exam data were collected from 7 semesters of the E&M 1 course (N=372) taught by 5 different instructors. Of these, four were traditional research faculty, and one was a physics education researcher. Two of these instructors, including the PER faculty member (SJP), taught the course twice during data collection. Questions on the exams were developed solely by the instructor for that semester; however, in all cases, the question provided a mathematical expression for a charge density and asked for a description and/or sketch of the charge distribution (e.g., Fig. 7; see Appendix C for exact prompts). In one case, the students were also asked to calculate the integral of the given δ -function expression. This type of canonical exam question will be referred to as an A1-type prompt as it corresponds to element DA1 of the operationalized ACER framework.

Student responses to the multiple-response CUE [51] provided an additional data source. The multiple-response CUE is a research-based, end-of-semester conceptual assessment that will be discussed in greater detail in Chapters IV-V. Only one question on the CUE deals with δ -functions, and it is of the same general type as the questions asked on the exams

A1-type Prompt:

Sketch the charge distribution:

$$\rho(x, y, z) = c\delta(x - 1)$$

Describe the distribution in words too.

What are the units of the constant, c ?

FIG. 7. An example of a canonical exam question involving the Dirac δ -function. This prompt corresponds with element DA1 of the operationalized ACER framework (Sec. III C 2).

(A1-type). However, rather than an open-ended prompt, the CUE asks a two-part, multiple-choice question (see Fig. 8). CUE data from CU were collected from four semesters (N=145) of the E&M 1 course; two of these were courses for which we also have exam data. In addition to the CU data, we also collected multiple-response CUE data from 9 courses at 7 external institutions (N=161). These institutions range from small liberal arts colleges to large research universities.

Think-aloud interviews (N=11) were conducted in two sets performed roughly a year apart on different sets of students. Interview protocols for both sets of interviews were designed, in part, to probe more deeply into difficulties identified in the exam solutions, and thus included one or more exam style (i.e., A1-type) questions (see Appendix D for full interview protocols). However, the operationalization of the ACER framework for δ -functions (Sec. III C 2) showed that these standard exam questions do not capture all aspects of problem-solving with the δ -function. In particular, this type of question bypasses Activation at anything more than the most surface level assessment of whether a student recognizes the δ -function. To probe Activation more deeply, we began the first set of think-aloud interviews with a question that provided a description of the charge distribution and asked for a mathematical expression for the charge density (see Fig. 9). This interview prompt will be referred to as an A2-type prompt as it corresponds to element DA2 of the operationalized framework. The A2-type prompt also provides a different perspective on Construction than the A1-type exam prompt by requiring students to generate a multidimensional δ -function expression rather

Q8 - A **mass** density is given by $\rho(\vec{r}) = m_1\delta^3(\vec{r} - \vec{r}_1) + m_2\delta^3(\vec{r} - \vec{r}_2)$, where m_1 and m_2 are constants.

- i. What is the value of $\int_{all\ space} \rho(\vec{r})d\tau$?
 - a) $m_1\delta^3(\vec{r} - \vec{r}_1) + m_2\delta^3(\vec{r} - \vec{r}_2)$
 - b) $m_1\delta^3(\vec{r}_1) + m_2\delta^3(\vec{r}_2)$
 - c) $m_1 + m_2$
 - d) $m_1\vec{r}_1 + m_2\vec{r}_2$
 - e) $m_1r_1 + m_2r_2$
 - f) $m_1r_1^2 + m_2r_2^2$

- ii. What physical situation does this mass density represent?
 - a) Two solid spheres of radius r_1 and r_2
 - b) Two concentric spherical shells of radius r_1 and r_2
 - c) Two nested cylindrical shells with radius r_1 and r_2 (i.e., a coaxial cable)
 - d) Two point masses located at \vec{r}_1 and \vec{r}_2
 - e) Two spheres of non-zero radius located at \vec{r}_1 and \vec{r}_2 (i.e., a dumbbell)

FIG. 8. The multiple-response CUE question related to the δ -function. This prompt is of the same general type as the A1-type prompt used on the midterm exams.

A2-type Prompt:

What is the volume charge density of an infinitely long, linear charge distribution running parallel to the z-axis and passing through the point (1,2,0)? Be sure to define any new symbols you introduce. Sketch this charge distribution.

FIG. 9. An example of the interview question used to specifically target spontaneous activation of the δ -function. This prompt corresponds to element DA2 of the operationalized ACER framework (Sec. III C 2).

than just interpreting it. However, participants in the first interview set often failed to activate the δ -function in response to the A2-type prompt and thus never moved on to deal with Construction. In order to target the Construction component more clearly, the second

set of interviews directly prompted students to use δ -functions to express the charge density.

Additionally, only one of the seven exam questions asked students to integrate an expression containing a δ -function. To further probe students ability to integrate the δ -function (element DE1), both interview protocols included questions that prompted students to calculate the total charge within a finite region of space. The second set of interviews also ended by asking students to perform a set of context-free integrations of various δ -function expressions (Fig. 10) in order to more clearly investigate the Execution component. These particular integral expressions were each designed to target a specific difficulty we anticipated students might have with the procedural aspects of integrating the δ -function. While all integrals were presented without a physical context, in the case of integral *d*) students were also asked if they could come up with a physical situation in which they might set up this integral.

4. *Student Difficulties with the Dirac Delta Function*

This section presents the identification and analysis of common student difficulties with the Dirac δ -function organized by component and element of the operationalized ACER framework (Sec. III C 2).

Perform the following integrations:

a) $\int_{-\infty}^{\infty} \delta(x) dx$

b) $\int_{-\infty}^{\infty} x\delta(x) dx$

c) $\int_0^{10} [a\delta(x-1) + b\delta(x+2) + c\delta(x-3)] dx$

d) $\iiint a\delta(r-r')r^2 \sin(\theta) dr d\phi d\theta$

FIG. 10. Context-free integral problems used in the second set of interviews (N=6) in order to directly target the Execution component of the framework.

Activation of the tool: Elements DA1 and DA2 of the framework represent two different types of prompts that can cue students' to activate resources related to the δ -function. In the case of A2-type prompts, the student must first recognize that the δ -function is the appropriate mathematical tool before they can correctly answer the question. However, for A1-type questions, the δ -function is given as part of the prompt, effectively short-circuiting Activation and providing little information about students ability to recognize *when* the δ -function is appropriate.

None of the instructor-written exams included A2-type questions. Instead, this element was specifically targeted during the first of the two interview sets. When presented with the A2-type prompt in Fig. 9, two of five interviewees spontaneously suggested using δ -functions. The remaining three participants all expressed confusion at being asked to provide a volume charge density for a line of charge. Two of these students attempted to reconcile this by defining an arbitrary cylindrical volume, V , around the line charge and using $\rho = Q/V$. This strategy, while incorrect, represents a reasonable attempt to make sense of the problem in lieu of the δ -function. Later in the interview, all five of these students were presented with the δ -functions expression for a linear charge density and all but one correctly interpreted this expression as describing a line charge. This result suggests that even immediately after completing a junior-level electrostatics course, many students may have difficulty recognizing *when* the δ -function is the appropriate mathematical tool even when they are able to provide a correct physical interpretation of it.

Three participants in the second interview set had just completed one semester of upper-division quantum mechanics at the time of the interview. To investigate the context-dependent nature of student Activation of the δ -function, we began these students' interviews by asking for a mathematical expression for the potential of a finite square well in the limit that the well became *very* narrow and *very* deep. Two of these three participants spontaneously suggested the δ -function as the appropriate tool. Both of these students explicitly focused on the use of the words "*very* narrow and *very* deep" just before suggesting the

δ -function. The third student initially attempted to write the potential as a piecewise function but brought up the δ -function when explicitly told there was a more compact way to represent the potential without using a piecewise notation. Successful activation of the δ -function in the quantum case seemed to be linked to the high degree of similarity between the visual representation of a deep, narrow potential well and the commonly used graphical representation of $\delta(x)$.

Construction of the model: The Construction component deals with mapping between the physics and mathematics of a problem. In the case of δ -functions, this mapping can be done in two directions: from mathematics to physical description, or from physical description to mathematics. The A1-type prompts (Fig. 7) used on exams, the CUE, and in interviews investigated students' ability to translate a mathematical expression for the charge density into a physical description of the charge distribution. As part of this process, students needed to connect the coordinate system and number of δ -functions in the given expression to the shape of the charge distribution (element DC1). For example, the expression $\rho(x, y, z) = c\delta(x - 1)$ contains one Cartesian δ -function and thus represents an infinite plane of charge. Roughly a quarter of students (23%, N=87 of 372) had an incorrect shape on exams. The most common error was misidentifying volume charge densities that included 1 or 2 δ -functions as point charges (62%, N=54 of 87).

On the CUE diagnostic administered at the end of the semester, the fraction of students selecting the incorrect shape was roughly a third (35%, N=51 of 145). This trend of selecting the incorrect shape is slightly more pronounced in student populations beyond CU who have taken the multiple-response CUE (49%, N=79 of 161, 7 institutions). The most common error for this question was misidentifying the given point mass density (see Fig. 8) as a solid sphere or spherical shell with radius \vec{r}_1 (76%, N=99 of 130 incorrect responses, all institutions). This finding is interesting because, given students' tendency to misidentify charge densities as point charges on exams, we might expect that they would be more successful at identifying point mass densities. However, we hypothesize that the prevalence

of this particular error may have been exacerbated by the apparent similarity between the expressions for this point mass density and that of a spherical shell ($\rho(\vec{r}) = m_1\delta(r - r_1)$), which are distinguished only by the presence of the cube on the δ -function and vectors in the argument.

In interviews, 9 of 11 participants correctly identified the shape of one or more charge distributions from the mathematical expressions for the charge density. The remaining 2 students both sketched the charge distribution on 3D Cartesian axes as a very narrow spike originating at a point consistent with the argument of the δ -function and extending upwards parallel to one of the axes (e.g., see Fig. 11). This ‘spike’ representation was also seen in a small number of the exam solutions (5%, N=17 of 311). The spike drawn by these students is highly reminiscent of the 1D graphical representation of $\delta(x)$ as an infinitely tall and thin Gaussian distribution at $x = 0$ that is commonly used when first defining the δ -function. Students who draw this ‘spike’ when sketching the charge distribution may be attempting to apply this 1D representation of the δ -function to a 3D sketch. The examples shown in Fig. 11 have a number of additional interesting features. For example, the student in Fig. 11(d) places the spike at $z = 2c$ rather than $z = 2$. Additionally, the spikes are not always parallel

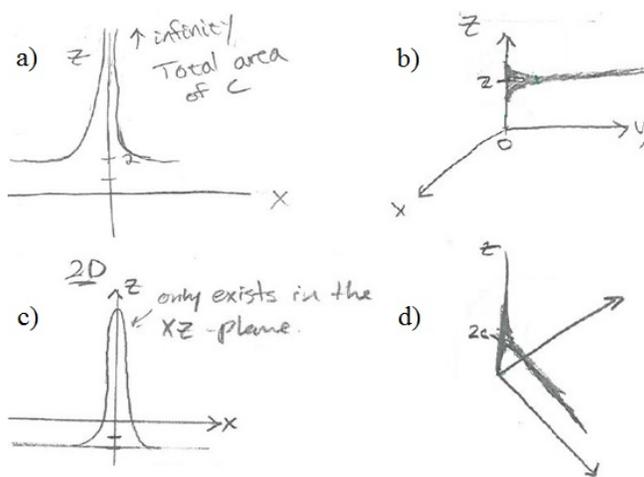


FIG. 11. Examples of the ‘spike’ representation used by 5% of exam students and 2 of 11 interview students. In this case, students were prompted to sketch the charge distribution described by $\rho(x, y, z) = c\delta(x)\delta(z - 2)$.

to the same axis despite all having the same prompt. These features may warrant further investigation in future studies; however, as they were observed in only a small number of exam solutions and in none of the interviews, we will not discuss them in further detail here.

The interviews also included an A2-type prompt (Fig. 9) to explore students' ability to translate a physical description of the charge distribution into a mathematical expression for the charge density. This process requires students to use the geometry of the charge distribution to select an appropriate coordinate system and number of δ -functions. Of the eight interview participants who attempted to use δ -functions in response to the A2-type prompt, three were able to correctly express the line charge density as the product of two 1D Cartesian δ -functions. Four of the remaining five students instead attempted to use a single δ -function whose argument was the difference between two vectors, for example $\rho \propto \delta(\vec{r} - \vec{r}_1)$ where $\vec{r}_1 = (1, 2, z)$. These students did not explicitly acknowledge this as a 3D δ -function either verbally or by cubing the δ -function. Three of these students also described the line charge as a continuous sum of point charges and integrated their δ -function expression over all z . This strategy reflects a fundamentally reasonable physical interpretation of the situation that can be used to construct a correct expression for the line charge density; however, none of these students were able to fully leverage this interpretation to correctly express the charge density.

Determining the location of the charge distribution (element DC2) was not a significant stumbling block for students in this study. None of the interview students and a tenth of the exam students (10%, N=39 of 372) drew an incorrect position for the distribution. The most common errors were switching the signs of the coordinates (36%, N=14 of 39, e.g., locating the plane described by $\rho(x, y, z) = c\delta(x - 1)$ at $x=-1$) or having the wrong orientation of line or plane distributions (38%, N=15 of 39). However, all relevant exam questions in this study dealt with locating charge distributions described by δ -functions in Cartesian coordinates, and it is possible that student difficulties with element DC2 would be more significant for other geometries or more abstract notation.

The third element in Construction relates to the need for a unitful constant in the expression for $\rho(\vec{r})$. For A1-type questions, this constant is provided as part of the given expression for the charge density, and fully interpreting this expression requires recognizing the physical significance of this constant. For example, in the expression $\rho(x, y, z) = c\delta(x - 1)$, the constant, c , represents the charge per unit area on the surface of the plane. Roughly a quarter of the exam students (27%, N=70 of 255) presented with an arbitrary constant spontaneously commented on its physical meaning and most of these (81%, N=57 of 70) had a correct interpretation. This fraction should be interpreted as a lower bound as additional students may have recognized the constant's physical significance but did not explicitly write it down on the exam. The interviews suggest that a students' interpretation of the constant can be facilitated or impeded by their identification of its units. This dynamic will be discussed in greater detail below in relation to the Reflection component.

The A2-type prompts used in interviews, on the other hand, do not include the unitful constant but instead require that a student recognize the need for this constant independently. Two of eight interviewees did not spontaneously include a multiplicative constant in their expression for the charge density. Alternatively, three of the remaining participants recognized the need for a multiplicative factor but expressed this factor as $\rho(\vec{r})$ (e.g., *volume charge density* = $\rho(\vec{r})\delta(x - 1)\delta(y - 2)$). These students appeared to be interpreting the $\rho(\vec{r})$ term as representing only the magnitude of the charge density rather than the magnitude at a specific point in space. This also suggests that these students are not treating ρ as the quantity defined by convention to be volume charge density and are instead treating it as the quantity conventionally defined as λ (i.e., line charge density). Ultimately, four of the eight participants included a multiplicative constant *and* articulated a correct physical interpretation of that constant before moving on from this question.

The final two elements in Construction (elements 4 & 5) are not specific to questions involving δ -functions, but rather apply to any physics problem involving multivariable integration. These elements deal with expressing the differential volume element and selecting

limit of integration. We have previously examined student work around these two elements in the context of Coulomb’s law integrals (see Sec. III B 4) and observed a number of difficulties. However, only a small number of students (21%, N=12 of 56) struggled with either the Cartesian differential volume element or limits of integration in the single semester where the exam question asked for total charge in a finite region of space. Furthermore, only one of eleven interviewees had clear difficulties expressing the differential charge element in spherical coordinates, and one other selected incorrect limits for their Cartesian integral. Thus, setting up integrals for the relatively simple Cartesian and spherical geometries used in this study was not a significant challenge for our upper-division students.

Execution of the mathematics: The Execution component of the framework deals with the procedural aspects of working with mathematical tools in physics. The exams provide limited insight into this component as students were asked to actually calculate an integral of a δ -function in only one semester. In this case, the students were given the expression for 3 point charges and asked to calculate $\int \rho(\vec{r})d\tau$. Roughly a quarter of the students (27%, N=15 of 56) made significant mathematical errors related to the δ -function while executing this integral (element DE1). The most common error (73%, N=11 of 15) amounted to a variation of equating the integral of the δ -function with the integral of the zero point of its vector argument (e.g., $\int \delta^3(\vec{r} - \vec{r}_1)d\tau = \int \vec{r}_1 d\tau$). Ultimately, only a small fraction of students (7%, N=4 of 56) had errors *only* with the elements in the Execution component (i.e., no mistakes in Activation or Construction).

More than a third of students (37%, N=53 of 145 at CU, 41%, N=66 of 161 at external institutions) selected an incorrect value for the integral of a point mass density in response to the CUE question. The most common error (81%, N=43 of 53 at CU, 68% N=45 of 66 at external institutions) was equating the integral of the δ -function with the value (vector or magnitude) of \vec{r} at which the argument was zero (e.g., $\int \delta^3(\vec{r} - \vec{r}_1)d\tau = \vec{r}_1$). This issue is different from, though potentially related to, the most common issue seen in the exam; however, none of the response options on the multiple-choice CUE question matched the

outcome of the incorrect setup seen commonly on exams. Ultimately, these types of errors indicate that students have correctly recognized that the δ -function somehow picks out the value $\vec{r} = \vec{r}_1$ but apply this reasoning in an *ad hoc* fashion in the integration process.

The first set of interviews investigated Execution in the context of calculating the total charge on a uniformly charged spherical shell. Two of the five participants were not able to complete this calculation and explicitly stated this was because they could not recall how to integrate the δ -function. When asked for the value of $\int_{-\infty}^{\infty} \delta(x)dx$ these participants guessed it would be x or $1/x$. Extending this logic into three dimensions, this response appears consistent with some of the common incorrect responses on the CUE and exam questions (e.g., $\int \delta^3(\vec{r} - \vec{r}_1)d\tau = \vec{r}_1$ or $1/|\vec{r}_1|$).

The second interview set (N=6) targeted the first element in Execution differently by asking students to perform the context-free integrations in Fig. 10. Two students stated that the integral in *b)* would be equal to x without evaluating this expression at $x = 0$, but none of the six participants had difficulty with the integrals in *a)* or *c)*. Moreover, the 3D δ -functions used on the exam and CUE seemed to evoke different, and potentially more fundamental, difficulties than the 1D δ -functions used in interviews. Three of six interviewees also evaluated the r integral in part *d)* in the following way, $\int \delta(r - r')r^2dr = \frac{1}{3}r'^3$, despite correctly executing parts *a)-c)*. Two of these students explicitly stated that the effect of the δ -function was only to pick out the value $r = r'$. This result is consistent with that from the exams and CUE, and suggests that a significant fraction of our upper-division students have internalized the idea that the δ -function picks out a particular value of the variable, but do not recognize the other impacts of the δ -functions on the result of the integral, particularly when dealing with 3D or non-Cartesian δ -functions.

Reflection on the result: The Reflection component deals with aspects of the problem-solving process related to interpreting and checking intermediate steps and the final solution. For the questions used in this study, one of the most powerful tools available for checking and interpreting the various δ -function expressions is looking at units (elements DR1 and DR2).

In particular, for A1-type problems, looking at the units of the given constant (e.g., the c in $\rho(x, y, z) = c\delta(x - 1)$) can facilitate interpretation of that constant's physical meaning. Five of the seven exam prompts explicitly asked students to comment on the units of the given constant and two-thirds of the students (70%, N=178 of 255) responded with the correct units (element DR1).

Beyond just commenting on the units of the given constant, we would like our students to recognize the physical interpretation of this constant (element DC3), and ensure that the units and physical interpretation are consistent (element DR2). However, it was often difficult to assess if students did this on the exams, in part because only a quarter of our students (N=70 of 255) explicitly commented on the physical interpretation of the unitful constant (element DC3). Most of these students (83%, N=58 of 70) also provided units for the constant that were consistent with their physical interpretation of its meaning. However, a third of the students (32%, N=82 of 255) gave units that were inconsistent with the shape they identified regardless of whether they commented on the constant's physical meaning. Some of these students had the correct shape but incorrect units (40%, N=33 of 82), and others had the incorrect shape but correct units (30%, N=25 of 82). For an expert, the units of the constant, its physical meaning, and the shape of the charge distribution are tightly linked; however, these results suggest that this relationship may still be developing for the students.

5. *Summary and Implications*

Our upper-division students encountered a number of issues when using/interpreting the δ -function. For example, our students often struggled to activate the δ -functions as the appropriate mathematical tool when not explicitly prompted. As our upper-division students progress forward through the undergraduate and graduate curriculum, it becomes increasingly important that they be able to recognize situations in which particular tools will be useful. Additionally, we found that students have a greater degree of difficulty translating

a verbal description of a charge distribution into a mathematical formula for volume charge density than the reverse process. While interpreting a mathematical formula for the charge density is a valuable skill for our physics majors, the ability to construct that same formula from scratch is potentially an even more valuable skill that our physics majors are likely to use in the future.

Our students also encountered difficulties with the procedural aspects of integrating 3D and/or non-Cartesian δ -functions despite often recognizing that the δ -function picks out the value of the integral at a single point. These difficulties manifest both in solving integrals embedded in a physics context and those presented in a purely mathematical context. Finally, our students demonstrated significant difficulty determining the units of the δ -function, thus limiting their ability to interpret or check their expressions for the charge density.

This study was not designed to investigate the impacts of different instructional strategies or curricular materials on the prevalence or persistence of students' difficulties with the Dirac δ -function; however, our findings do suggest several implications for teaching or using the δ -function at the upper-division level. For example, instructors should be aware that students are unlikely to encounter the δ -function in their required mathematics courses, and that it may or may not be covered in a math methods course run within a Physics department. Thus, the assumption that all students in a junior-level course will be aware of the δ -function and its properties may not be justified. Students' struggles with the procedural aspects of integrating more complex δ -function expressions may be one manifestation of this lack of sufficient prior experience with the δ -function. These students may benefit from additional opportunities to practice integrating both 1D and 3D delta functions and in multiple coordinate systems.

Additionally, canonical δ -functions questions rarely, if ever, require a student to consider *when* the δ -function is an appropriate tool. Questions that describe a charge distribution and ask for an expression for the charge density can provide a baseline assessment of student ability to activate the δ -function when not prompted explicitly. This type of question also addresses another finding: constructing a mathematical expression for the charge density is

a distinct and potentially more challenging task for our students than interpreting that same expression. As the former task is arguably the more authentic, students may benefit from additional opportunities to construct various δ -function expressions in multiple coordinate systems.

The belief that the δ -function is unitless was a surprisingly prevalent and persistent idea. This belief may be exacerbated by presenting the δ -function as a purely abstract mathematical construct. Moreover, the idea of a unitless δ -function can interfere with students' interpretation of the unitful constant. To facilitate student reflection on problems involving the δ -function, specific emphasis should be placed not only on the fact that the δ -function *can* have units, but also on *how* to determine them based on the argument of the δ -function.

D. Separation of Variables in the context of Laplace's Equation

1. *Previous Research on Student Difficulties with Separation of Variables*

Partial Differential Equations (PDEs) appear in multiple contexts in the upper-division, undergraduate physics curriculum (e.g., waves on a string, Maxwell's equations, the Schrödinger equation). One of the most common approaches to solving a PDE is to turn it into multiple Ordinary Differential Equations (ODEs) using a technique known as separation of variables (SoV). In this thesis, we use the term SoV to refer to the technique of guessing a general solution with a functional form that allows the PDE to be separated into several ODEs and then solving these ODEs individually with appropriate boundary conditions. This technique is not to be confused with the strategy (also conventionally referred to as separation of variables) used to solve separable ODEs by isolating terms with the function and the variable on either side of the equals sign and integrating. We are not aware of any existing research specifically targeting student difficulties with SoV, though some work has been done looking at solving ordinary differential equations generally. Much of this differential equations literature focuses on students' use of graphical techniques to solve linear ODEs

[68–70], or numerical solutions to more complex differential equations that cannot be solved analytically [71].

2. *Operationalizing ACER for Separation of Variables*

At CU, physics students encounter SoV several times in their undergraduate courses. The first exposure is often in sophomore classical mechanics as a technique to solve Laplace’s equation in Cartesian coordinates, typically in the context of solving for the temperature as a function of position in a mechanical system with given boundary conditions. Students may also encounter this technique in a differential equations course taken from the Math department. Junior electrostatics is the next place where CU’s physics majors see SoV in the context of solving Laplace’s equation for the electric potential, V , in 2D and 3D Cartesian coordinates and spherical coordinates with azimuthal symmetry. Students do not typically encounter spherical SoV without assuming azimuthal symmetry until quantum mechanics where it is used to solve the Schrödinger equation for the hydrogen atom. In discussions with the physics faculty at CU, some instructors have expressed the sense that students do not begin to demonstrate mastery of the SoV technique until they see it for a third time in quantum mechanics, and sometimes not even then.

Here, we focus on student use of SoV in junior-level electrostatics. To guide our data collection and analysis, we operationalized the ACER framework for problems involving solutions to Laplace’s equation in Cartesian coordinates (i.e., $\nabla^2 V(x, y, z) = 0$) and spherical coordinates with azimuthal symmetry (i.e., $\nabla^2 V(r, \theta) = 0$). These types of problems typically ask for an expression for the voltage in a charge-free region and provide an expression for the voltage along the boundary of that region. To operationalize the framework, we created a network diagram expressing the conceptual and procedural resources used by experts when solving these types of SoV questions (Fig. 12). Note that while this diagram includes both Cartesian and spherical geometries, there are several points where the framework highlights procedural differences between the solutions in these two different coordinate systems.

An interesting aspect of this network diagram, relative to those for integration (Fig. 3) and delta-functions (Fig. 6), is the degree of interplay between Construction and Execution. Particularly for Cartesian SoV, expert solution paths move back and forth between Construction and Execution several times before arriving at a complete solution. The implications of this dynamic will be discussed in greater detail in the following sections.

The elements of the flattened ACER framework extracted from the SoV network diagram are detailed below. Recall, element codes are for labeling purposes only and are not meant to suggest a particular order, nor are all elements always necessary for every problem. In particular, the elements of Construction and Execution are unlikely to occur in the specific order listed.

Flattened ACER framework

Activation of the tool: The first component of the framework involves identifying SoV as the appropriate mathematical technique to solve for the voltage. We identified three elements in the form of cues present in a prompt that are likely to activate resources associated with SoV.

SA1: The question provides boundary conditions and asks for separation of variables directly or provides the expression for the general solution

SA2: The question provides boundary conditions and uses language associated with separation of variables (e.g., infinite sum, Legendre polynomials, Fourier series, general solution)

SA3: The question provides boundary conditions and asks for the electric potential or voltage in a charge free region

For element SA1, it is more common for a question to provide the general solution for the voltage in spherical geometries than in Cartesian, in part because the functional form of the general solution in Cartesian depends on the boundary conditions (element SC2).

Construction of the model: Elements in this component deal with modifying the general expression for the solution to Laplace's equation so that it matches the boundary

conditions. As there is no ambiguity in the signs of the separation constants for SoV in spherical coordinates due to the nature of the θ coordinate, the second element of Construction is specific to Cartesian problems.

SC1: Determine any relevant boundary conditions, both those explicitly given in the prompt/figure and those implicit from the physical situation (e.g., $V(r \rightarrow \infty) \rightarrow 0$)

SC2: (Cartesian only) Choose the signs of the separation constants or select the functional forms for the solution that are appropriate for the boundary conditions

SC3: Apply each boundary condition to the general solutions in order to solve for all unknown constants (set up only)

Note that these elements do not necessarily occur sequentially, either with respect to one another or with respect to elements of the Execution component.

Execution of the mathematics: This component of the framework deals with elements involved in executing the mathematical operations related to SoV. As students are rarely (if ever) asked to actually produce and solve the ODEs resulting from SoV in spherical coordinates, the first two elements of Execution are specific to Cartesian (and potentially Cylindrical) problems.

SE1: (Cartesian only) Manipulate the PDE into ODEs using the separated form of the potential (e.g., $V(x, y, z) = X(x)Y(y)Z(z)$)

SE2: (Cartesian only) Know or look up the solution to these ODEs given the signs of the separation constants

SE3: Calculate values for all unknown constants based on applying the boundary conditions through zero matching, term matching, or Fourier's trick integrals.

SE4: Manipulate algebraic expressions into forms that can be readily input into $V(\vec{r})$ and interpreted

Element SE3 can be accomplished in a variety of ways sometimes involving several smaller steps depending on the particular boundary conditions. The strategies we refer to as zero matching, term matching, and Fourier's trick are explicitly described in Sec. III D 4.

Reflection on the result: The final component includes elements related to checking and interpreting aspects of the solution, including intermediate steps and the final result. While many different techniques can be used to reflect on a physics problem, the following four are particularly common when dealing with SoV.

SR1: Check the units of the final expression

SR2: Check that the solution matches all boundary conditions

SR3: (Spherical only) Check that the solution behaves as expected in known limits (may include calculating lowest order moments)

SR4: Confirm that the solution satisfies Laplace's Equation

Element SR3 refers specifically to checking the functional dependence, rather than the value, of the voltage in known limits. For example, checking that $V(r \rightarrow \infty) = 0$ would be considered SR2 while showing that V goes to zero as $1/r$ would be SR3. The final element of Reflection (SR4) was added to the framework after initial analysis of student work where we observed that mistakes in the Construction and Execution components often resulted in solutions that did not satisfy Laplace's equation.

In the next section, we will apply this operationalization of ACER to investigate student work on canonical SoV problems involving both Cartesian and spherical geometries.

3. Interview and Exam Prompts

Midterm exam data were collected from 10 semesters of the E&M 1 course (N=474) taught by 8 different instructors. Of these, six were traditional research faculty and two were physics education researchers. Two of these instructors, including one of the PER faculty members (SJP), taught the course twice during data collection. Questions on the exams were developed solely by the instructor for that semester. In four cases, the instructor asked one or more SoV questions on both the midterm and final exams, thus the following section reports the analysis of 15 distinct exam questions for a total of N=744 unique solutions.

A rectangular pipe, running parallel to the z -axis, extending from $-\infty$ to ∞ , has three grounded metal sides and a fourth side maintained at a constant potential, V_o , as in the figure. Find the potential $V(x, y)$ at all points inside the pipe.

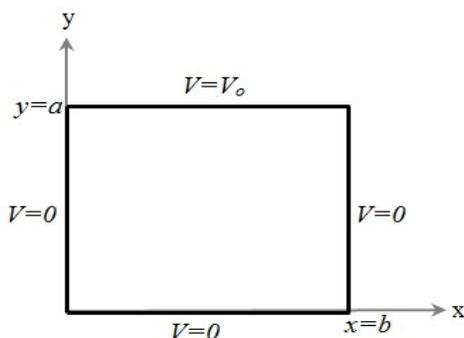


FIG. 13. An example of a canonical exam problem targeting Cartesian SoV. Variations on this question include providing a non-constant potential on the fourth side (e.g., $V_o(x, y = a) = V_o \sin \pi x/a$), placing the non-grounded side at $y = 0$ rather than $y = a$, or placing one of the grounded sides at infinity.

A thin spherical shell of radius R and centered on the origin has a voltage on its surface given by $V(R, \theta) = V_o \cos^2(\theta)$.

Find the potential $V(r, \theta)$ everywhere (inside *and* outside the sphere).

FIG. 14. An example of a canonical exam problem targeting spherical SoV. Variations on this question include providing a simpler or more complex expression for the boundary (e.g., $V(R, \theta) = V_o \cos \theta$), giving the boundary condition in terms of Legendre polynomials (e.g., $V(R, \theta) = P_9(\cos \theta)$), or only asking for the potential inside or outside the sphere.

As our goal is to identify the presence of common student difficulties, the remainder of this analysis will report N as the number of solutions rather than number of students.

Exam questions requiring the use of SoV in Cartesian and spherical coordinates are both common at CU. Four of the exam questions in our sample ($N=235$ solutions from 3 semesters) provided students with a rectangular pipe or gutter with given values for the voltage on each side and asked for an expression for the voltage valid everywhere inside (e.g., Fig. 13, see Appendix E for exact prompts). The remaining 11 exam questions ($N=509$ solutions from 9 semesters) provide students with an azimuthally symmetric expression for the voltage on the surface of a spherical shell and asked for an expression for the voltage valid inside and/or outside the shell (e.g., Fig. 14, see Appendix F for exact prompts).

Student responses to the multiple-response CUE (Colorado Upper-division Electrostatics diagnostic - a research-based, end-of-semester conceptual assessment described in Chapter IV) provide an additional data source. Two questions on the CUE deal with SoV: one in Cartesian (Fig. 15), and one in spherical (Fig. 16). CUE data were collected from four semesters (N=145) of the E&M 1 course at CU; three of these were courses for which we also have exam data. In addition to the CU data, we also collected multiple-response CUE data from 9 courses at 7 external institutions (N=161) ranging from small liberal arts colleges to large research institutions.

Think-aloud interviews (N=11) were conducted in two sets performed roughly two years apart (see Appendix G for full interview protocols) in order to further probe preliminary difficulties identified in student exams. The first interview set (N=4) was designed to target student difficulties with SoV in Cartesian coordinates and was conducted prior to the development of the ACER framework. The students were asked to determine the voltage inside a semi-infinite plate with one side held at a constant potential (similar to the question

Q13 - To solve for V inside the box by separation of variables, which form of the solution should you choose? Select only one.

A. $V(x,y) = (A e^{ky} + B e^{-ky})(C \sin kx + D \cos kx)$

B. $V(x,y) = (A \sin ky + B \cos ky)(C e^{kx} + D e^{-kx})$

C. $V(x,y) = (A \sin ky + B \cos ky)(C \sin kx + D \cos kx)$

D. $V(x,y) = (A e^{ky} + B e^{-ky})(C e^{kx} + D e^{-kx})$

E. More than one of these could be used

because ... (select ALL that support your choice of method)

a. it is harder (but not impossible) to fit the boundary conditions on x with exponentials

b. it is harder (but not impossible) to fit the boundary conditions on y with exponentials

c. oscillatory solutions *cannot* fit the boundary conditions on x

d. oscillatory solutions *cannot* fit the boundary conditions on y

e. exponential solutions *cannot* fit the boundary conditions on x

f. exponential solutions *cannot* fit the boundary conditions on y

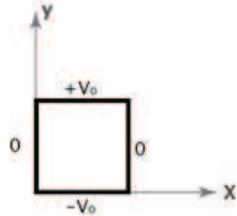
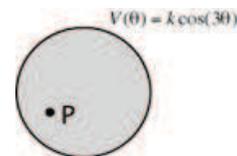


FIG. 15. The multiple-response CUE question related to the Cartesian SoV. The prompt has been paraphrased, see Appendix H for full prompt.

Q1 - A hollow, insulating spherical shell with radius R , with a voltage **on its surface**, $V(\theta) = k\cos(3\theta)$. Find \vec{E} or V inside the sphere at point P.

Select only one: **The easiest method would be ...**

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. symmetry allows you to calculate \vec{E} using a spherical Gaussian surface
- c. the boundary conditions is azimuthally symmetric (i.e., symmetric in ϕ)
- d. there is not appropriate symmetry to use other methods
- e. you can use $\vec{E}(\vec{r}) = -\vec{\nabla}(k\cos 3\theta)$ and evaluate this at point P
- f. $\nabla^2 V = 0$ inside the sphere and you can solve for V using Legendre Polynomials

FIG. 16. The multiple-response CUE question related to spherical SoV. This prompt is consistent with element SA3 of the framework.

in Fig. 13). The students were directly prompted to approach this problem by using the SoV technique to solve Laplace's equation. They were also provided the expression for ∇^2 in Cartesian coordinates along with the solutions to the relevant ODEs and the integral expression needed to determine the coefficients in a Fourier series. From the perspective of the ACER framework, this prompt clearly targeted the Construction and Execution components, but bypassed Activation.

The second interview set (N=6) began with a spherical SoV problem in which the students were given a spherical shell with a known voltage on the surface (see Fig. 14). To directly target Activation, the prompt did not specifically mention Laplace's equation or prompt the students to use the SoV technique. Students who were able to complete this problem in the allotted time were also explicitly prompted to come up with a way to check their solution in order to convince themselves their expression was correct (i.e., Reflection). The second interview set ended by asking students to begin working though the Cartesian question shown in Fig. 13, though no student had time to fully complete this question. While the

Cartesian prompt also did not directly prompt SoV, this question did not provide insight into spontaneous Activation because it came immediately after a spherical SoV question.

While we intentionally operationalized the framework for problems in both Cartesian and spherical geometries, our analysis found that the type and frequency of errors was often quantitatively and/or qualitatively different for the two geometries. For this reason, we report on student difficulties with these two geometries separately. Where appropriate, we also synthesize these findings in Sec. III D 6.

4. Student difficulties with Separation of Variables in Cartesian Coordinates

This section presents the identification and analysis of common student difficulties with the separation of variables technique in Cartesian coordinates organized by component and element of the operationalized ACER framework (Sec. III D 2).

Activation of the tool: Canonical Cartesian SoV questions (e.g., Fig. 13) are highly distinctive both because they provide boundary conditions and because they do not provide information needed to solve the problem using other methods (e.g., $\rho(\vec{r})$). Elements SA1-SA3 describe different types of prompts that can cue students to activate resources related to SoV, loosely organized according to the likelihood that they will do so. One of the four exam questions (N=69) explicitly prompted students to use SoV and presented them with the appropriate separated form for the solution (i.e., $V(x,y) = X(x)Y(y)$); thus, these solutions provide minimal insight into Activation. Alternatively, prompts consistent with SA2 or SA3 require students' to identify SoV as the appropriate technique, and only one of the 166 solutions to implicit prompts used a method other than SoV. Due to the distinctive nature of Cartesian SoV questions, we are hesitant to interpret this result as evidence that our students have a solid understanding of *when* SoV is the correct approach; however, it does suggest they are adept at recognizing these canonical questions.

Construction of the model: The Construction component deals with mapping between the physics and mathematics of a problem. For SoV, this process includes identifying all

necessary boundary conditions (element SC1). Cartesian SoV problems typically provide these boundary conditions explicitly in the prompt (see Sec. III D 5 for discussion of implicit boundary conditions). Consistent with this, only a small fraction of solutions (5%, N=12 of 234) used incorrect values for the boundary conditions. Common errors included putting the non-zero boundary condition on the wrong side (N=4), including inappropriate implicit boundary conditions (N=4, e.g., $V(x \rightarrow \infty) = 0$), or listing the value of the potential at a point (usually a corner) rather than along a side (N=2). One interview participant also listed the boundary conditions at each corner. This student recognized and corrected the error after attempting to apply a boundary condition and finding that this did not help him solve for any unknowns. Thus, extracting boundary conditions from the prompt or figure for Cartesian SoV questions was not a significant stumbling block for the majority of our students.

After identifying boundary conditions, the next step is to produce a general expression for the voltage that can satisfy these boundaries (element SC2). For Cartesian SoV questions, this amounts to deciding which direction (x or y) gets the exponential dependence. One exam prompt (N=69) asked students to select the appropriate general solution for $V(x, y)$ from the two possible Cartesian solutions to Laplace's equation (see Appendix E), and all students selected the correct expression. The multiple-response CUE asked a similar question but provided two additional response options which featured sinusoidal or exponential dependence in both directions (Fig. 15). In contrast to the results on the exam question, only two-thirds of CU students (66%, N=95 of 145) selected the correct expression, while the majority of the remaining students selected either the solution with flipped functional dependence (12%, N=17 of 145) or one of the two response options that did not satisfy Laplace's equation (14%, N=21 of 145). This trend is even more pronounced in student populations at other institutions, with almost a quarter of students (23%, N=37 of 161) selecting either purely sinusoidal or purely exponential dependence.

The remaining three exam prompts did not provide possible expressions for the voltage.

In practice, this meant that students could explicitly work through the process of separating Laplace's equation (elements SE1, SC2, and SE2) or jump straight to a general expression for the potential without deriving this expression. Using the former strategy, element SC2 requires deciding which separation constant gets the negative value (and thus which direction gets the sinusoidal dependence). Roughly two-thirds of solutions (71%, N=117 of 165) explicitly commented on the signs of the separation constants, and most (87%, N=102 of 117) assigned the negative constant such that it was consistent with the boundary conditions. Similarly, just over three-quarters of solutions that jumped straight to a general expression (79%, N=38 of 48) also gave a functional form that was consistent with the boundary conditions. As with the responses on the multiple-response CUE, the most common errors included either flipping the functional form or having sinusoidal solutions (or negative separation constants) in both directions.

Of the six interview students who progressed far enough in their solution to begin one of the Cartesian SoV questions (see Appendix G), four derived the general expression directly from Laplace's equation. All four correctly identified which coordinate (x or y) should be given the negative sign and justified this based on the boundary conditions. The two remaining students jumped straight to a general expression for the potential without derivation. One of these two argued for needing both exponential and sinusoidal dependence, but used complex rather than real exponentials. The second of these students argued for sinusoidal behavior in both x and y directions and justified this by stating that the boundary conditions in both directions could be matched by sines. Only when directly asked to show that this expression solved Laplace's equation did this student recognize that one of the two directions must have exponential dependence. This result suggests that students who argued for purely sinusoidal or purely exponential dependence on the exams and CUE may have been focusing on satisfying the boundary conditions without considering that, ultimately, the solution must also satisfy Laplace's equation.

The final element of Construction deals with setting up the equations that are used to solve

TABLE III. Difficulties setting up expressions match the boundary conditions in Cartesian. To account for the fact that certain difficulties were not applicable to all exam prompts, percentages are given with respect to the subset of incorrect solutions taken from the applicable semesters. Codes are not exhaustive or exclusive but represent the most common themes, thus the total N in the table need not sum to 71.

Difficulty	N	Percent
Incorrect application of non-zero boundary condition e.g., not plugging in a value for x or y , or using $Y(y = a) = V(x)$	32	20 (of N=71)
Missing + or - exponential term (relevant for box questions only excludes gutters)	16	24 (of N= 66)
Never applied non-zero boundary condition	14	20 (of N=71)
Incorrect setup of Fourier's trick integral e.g., missing or extra sum, not multiplying by $\sin(n'\pi x/a)$	8	12 (of N=66)

for the unknown constants in the general solution in order to match the boundary conditions (note, algebraic mistakes related to solving these equations will be discussed in relation to the Execution component). Almost a third of solutions (30%, N=71 of 234) included issues with setting up the equations to solve for one or more constants. The exact details of these errors were often strongly tied to the nature of the specific boundary conditions in question, but common issues included (see Table III): inappropriately arguing that one exponential term should be eliminated, incorrectly setting up or failing to utilize the non-zero boundary condition, or setting up an integral (i.e., Fourier's trick) incorrectly. As we might expect, the majority of these issues centered around setting up the expression to match the non-zero boundary condition.

In interviews, four students attempted to solve for unknowns (element SC3), and three had difficulty doing so. Consistent with student performance on exams, the most common issue related to using the non-zero boundary condition to solve for the final constant(s). None of these three students spontaneously included a sum in their solution even after applying the non-zero boundary condition and finding that the resulting equation could not be solved (e.g., $V(x, y = 0) = V_o = A\sin(\pi x/a)$, where A is a constant). After being prompted that they actually had an infinite number of solutions (rather than just one), all three students

recalled that they needed to introduce a sum of these solutions, but none could clearly justify how this step helped. One potential explanation for the increased frequency of this issue relative to the exams is that, on exams, students are recalling an algorithm that includes introducing a sum but have not internalized the motivation for that sum.

Execution of the mathematics: The Execution component deals with the procedural aspects of working through the mathematics of a physics problem. For Cartesian SoV, this can include the process of separating Laplace's equation into ODEs by assuming the separated form of the potential (i.e., $V(x, y) = X(x)Y(y)$, element SE1). Excluding solutions from semesters where the general solution for the potential was provided, more than half of the exam solutions (59%, N=98 of 165) explicitly included this process and only a small fraction (10%, N=10 of 98) had difficulties with it. The most common error (N=7 of 10) was using different constants in the x and y ODEs, resulting in a solution with the correct functional form that does not actually satisfy Laplace's equation.

In interviews, six of seven participants commented on or attempted to work through this process of separating Laplace's equation. Of these, four students spontaneously suggested assuming the separated form of the potential (element SE1), though one student noted that he did not understand the motivation for making this assumption. Of the remaining two participants, one clearly articulated that the goal was to separate Laplace's equation into ODEs, but could not recall how to do this on his own. The other student neither recalled the separated form nor recognized its purpose without being explicitly told. Additionally, four of the interviewees either did not recognize that the expression $f(y) + f(x) = 0$ implies $f(y) = -f(x) = c$ (where c is a constant), or attempted to apply this logic before having fully separated x and y dependent terms (e.g., arguing $X(x)''Y(y) = c$). Given that there is little (if any) physical motivation for assuming $V(x, y) = X(x)Y(y)$, it becomes particularly important that students understand the mathematical motivation for this move. However, interviews suggest that, even when students correctly use this assumption in their solution, they may not have a clear sense of the motivation or justification for this assumption.

The second element in Execution involves solving the ODEs that result from separating Laplace's equation (e.g., $X(x)'' = \pm k^2 X(x)$, where k^2 is the separation constant). More than two thirds of solutions (70%, N=116 of 165) included an expression for one or more ODEs either derived from Laplace's equation or stated without work. In practice, it is typical for students to simply write down the solutions to these ODEs either by memory or from an equation sheet, and just under a fifth of solutions (16%, N=19 of 116) provided a general solution that was inconsistent with the ODE they were solving. Common mistakes included providing a solutions whose functional form was inconsistent with the sign of the separation constant (N=7 of 19), or using the separation constant (rather than its square root) in the expression for the general solution (N=5 of 19). Thus solving the relatively simple ODEs required for Cartesian SoV questions was not a significant barrier to our students' success.

The Execution component of ACER also deals with the procedural mathematics of determining values for each of the unknown constants (element SE3) in order to match the boundary conditions (element SC3). Our initial analysis of both expert and student work suggested that there were three common strategies used to solve for these constants:

Zero matching: – Setting unknown constants to zero in order to enforce boundary conditions on which $V = 0$.

Fourier's trick: – The strategy used to solve for the coefficients in a Fourier series by exploiting the integral properties of orthogonal functions.

Term matching: – The strategy of exploiting the properties of orthogonal functions to directly match the coefficients of like terms.

Zero matching is nearly always necessary in Cartesian SoV questions, and nearly all of the exam solutions demonstrated some form of zero matching (94%, N=220 of 234). Alternatively, whether Fourier's trick or term matching is used to solve for the final unknown constant(s) often depends on the nature of the final boundary condition. For one of our three exam questions, the final boundary had a constant voltage (e.g., Fig. 13), making it necessary to solve using Fourier's trick. However, the two remaining exams provided a voltage of the

TABLE IV. Common difficulties when executing the procedural mathematics of solving for constants in the general solution. Percentages are of just the students who exhibited these difficulties (45%, N=105 of 234). Codes are not exhaustive or exclusive but represent the most common themes, thus the total N in the table need not sum to 105.

Difficulty	N	Percent
Off by a constant factor or sign e.g., factor of 2, V_o , or length a	33	31
Problems with a Fourier's trick integral e.g., pulling non-constant terms like $V(x, y = a)$ out of the integral, or not collapsing the sum	30	29
Dropping the factor of $Y(y = a)$ from the final solution	18	17
Not finishing the calculation	27	26

form $V(x, y = a) = V_o \sin(\pi x/a)$ (see Appendix E). In these cases, it is possible to use either Fourier's trick or term matching, though term matching is considerably simpler. Despite this, more of our students' solutions utilized Fourier's trick (44%, N=52 of 117) than term matching (36%, N=42 of 117). This result may indicate that our students have a strong association of Fourier's trick with SoV in Cartesian coordinates, and/or that they have not internalized the properties of orthogonal functions enough to see term matching as a viable strategy. Without interview data on student reasoning for this specific type of boundary condition, we are not able to distinguish between these two possible explanations.

When solving for the values of the unknown constants in their general solution (element SE3), roughly half of students' solutions (45%, N=105 of 234) contained various mathematical mistakes. Common errors included (see Table IV): losing or gaining a constant factor, incorrectly executing a Fourier's trick integral, (when applicable) not including the constant factor $Y(y = a)$, and not finishing the calculation. For the two exam questions that could be solved using either Fourier's trick or term matching, the fraction of solutions with mathematical errors was higher in solutions that utilized Fourier's trick (77%, N=40 of 52) than in those that utilized term matching (33%, N=14 of 42). This is likely due, at least in part, to the fact that Fourier's trick requires the set up and execution of an integral, and thus is a more mathematically demanding strategy.

For solutions in which the student finished solving for the final constant(s) (78%, N=130 of 166), it was then necessary to compile all aspects of the solution into a single expression for the voltage (element SE4). Just under a quarter of solutions (22%, N=29 of 130) either did not compile a final expression or made various mathematical mistakes not related to previous Execution or Construction errors (e.g., incorrectly simplifying exponentials to hyperbolic trig functions, or dropping or adding non-constant factors such as $Y(y)$). In practice, the interviews provided limited insight into the procedural aspects of solving for the unknown constants as students were, at most, asked to set up the expression for the final constant(s); however, none of our interview participants had difficulty with the simple manipulations required to match the $V = 0$ boundaries.

Ultimately, roughly a quarter of the solutions (26%, N=61 of 235) included *only* errors related to the elements in the Execution component (i.e., no previous mistakes in Activation or Construction). This number is high relative to the fraction of students who had difficulty only with Execution when utilizing multivariable integration (8%, Sec. III B 4) or the Dirac delta function (7%, Sec. III C 4). This result suggests that, particularly for problems involving Fourier's trick, the procedural mathematics involved in problems requiring Cartesian SoV can be a significant barrier for our junior-level electrostatics students.

Reflection on the Result: The Reflection component deals with the process of checking and/or interpreting the final expression. It is often the case in Cartesian SoV that mistakes in the Construction or Execution components resulted in an expression for the potential that had the wrong units, did not match the boundary conditions, or did not satisfy Laplace's equation (i.e., elements SR1, SR2, and SR4 respectively). Overall, we found that very few of our students (N=2 of 234) made explicit, spontaneous attempts reflect on their solution using any of these checks. This number should be interpreted as a lower bound on the frequency of spontaneous reflections as is possible that more of the exam students made one of these checks and simply did not write it down explicitly on their exam solution. However, only two of seven interview participants made spontaneous attempts to reflect on their solutions,

TABLE V. Number of exam students who explicitly utilized each of the three possible reflective checks (N explicit) along with the number of solutions that included an error that would have been detected by this check (N incorrect). Total N represents the total number of solutions that could have utilized that reflective check.

Reflective check	Total N	N incorrect	N explicit
Units (SR1)	125	19	0
Boundary conditions (SR2)	138	97	1
Laplace's equation (SR4)	154	30	1

and exclusively by checking that their general solution satisfied Laplace's equation (element SR4). One other student executed this check only after being directly prompted.

Another strategy for understanding Reflection involves looking at the number of solutions where the final expression included an error that would have been detected by one or more of these checks. Table V lists this along with the number of solutions that explicitly included each reflective check. Overall, these results suggest that an explicit check of boundary conditions would likely be the most effective reflective practice for students in terms of detecting errors, but that our students are rarely executing this (or other) checks spontaneously.

5. *Student difficulties with Separation of Variables in Spherical Coordinates*

This section presents the identification and analysis of common student difficulties with the separation of variables technique in spherical coordinates organized by component and element of the operationalized ACER framework (Sec. III D 2).

Activation of the tool: One instructor exclusively used SA1-type prompts on both the midterm and final exams in his course (N=138); thus, these solutions provide minimal insight into Activation. However, of the solutions to exam questions with implicit prompts (i.e., consistent with elements SA2 or SA3), very few (4%, N=16 of 371) utilized a method other than SoV (e.g., Coulomb's law, $\vec{E} = -\vec{\nabla}V(R, \theta)$, etc.). Alternatively, on the multiple-response CUE question (Fig. 16) asked at the end of the semester (but before the final) just under half of our students (41%, N=59 of 145) did not select SoV as the appropriate solution

method. This trend is slightly increased for other institutions with just over half the students selecting other methods (60%, $N=96$ of 161). The most common alternatives were Direct Integration via Coulomb's law (26%, $N=40$ of 155 incorrect responses, all institutions) and Gauss' law (45%, $N=69$ of 155 incorrect responses, all institutions). This may be a reflection of the fact that spherical SoV questions, while still distinctive from an expert point of view, are potentially less recognizable to students than their Cartesian counterparts due to their superficial similarity to problems that might be solved by Coulomb's law or Gauss' law.

The second set of interviews provided additional insight into Activation of spherical SoV through a question like the one shown in Fig. 14 (see Appendix G for the exact prompt). Of the six interview participants, three spontaneously brought up Laplace's equation and suggested SoV as the correct solution method. Of the remaining students, one mentioned Laplace's equation only after being prompted to consider the fact that all the charges would be confined to the surface of the shell, while the other two needed to be explicitly told to consider Laplace's equation. Moreover, these three students only suggested using SoV after being reminded that Laplace's equation is a complex PDE and asked how we generally deal with PDEs in physics. This result suggests that, as we might expect, the activation of the SoV technique for these students was more closely linked to the formal mathematics of the problem, rather than the physical context.

Students' overall success at Activation on the midterm and final exam questions seems to contradict the significantly lower success rate seen on the CUE and in interviews. One potential explanation for this is that students are simply pattern matching on the exams rather than internalizing a clear motivation for when and why SoV is appropriate. This interpretation is supported by the following comment made by one of the interview participants: "I remember these questions; I used to love these questions, and I don't remember how to do them anymore ... I guess I didn't understand this problem as well as I should have; I just remember going through a mathematical, like, process to get it, and I knew that one really well."

Construction of the model: The Construction component deals with mapping between the physics and mathematics of a problem. For spherical SoV, this process includes identifying all necessary boundary conditions (element SC1), both those provided explicitly in the prompt and those that are implicit in the underlying physics of localized charge distributions (i.e., $V(r \rightarrow \infty) \rightarrow 0$ and $V(r \rightarrow 0) \neq \infty$). Of the solutions that utilized SoV on the exams (N=488), almost two-thirds (61%, N=298 of 488) included correct expressions for all explicit and implicit boundary conditions. Of the remaining solutions, more than half (62%, N=117 of 190) never expressed the relevant implicit boundary conditions at $r = 0$ and/or $r = \infty$. Despite this, the majority of these solutions (89%, N=104 of 117) correctly eliminated either the A_l (outside) or B_l (inside) terms. This move was often accompanied by seemingly axiomatic statements like “ A_l ’s go to zero outside.” This finding is also consistent with the idea that some students are using pattern matching to guide their solution rather than clearly justifying their steps from the underlying physics. Other issues with expressing the boundary conditions (element SC1) included using incorrect or inappropriate implicit boundary conditions (12%, N=22 of 190, e.g., enforcing $V(r \rightarrow \infty) \rightarrow 0$ when solving for V inside a sphere), or incorrectly expressing the surface boundary condition (22%, N=42 of 190, e.g., arguing $V(R) = V_o \cos^2 \theta \rightarrow V_o P_2(\cos \theta)$).

As upper-division students are rarely (if ever) expected to derive the general expression for the voltage from Laplace’s equation in spherical coordinates, the second element of constructions does not typically apply to spherical SoV questions. Alternatively, the final element of Construction deals with setting up the equations to solve for the unknown constants in the general solution in order to match the boundary conditions (note, algebraic mistakes related to solving these equations will be discussed in relation to the Execution component). Ultimately, just under a fifth of solutions (19%, N=90 of 485) included issues with setting up the equations to solve for one or more constants. The most common issues included (see Table VI): not plugging in $r = R$ when matching the boundary condition at the surface, problems expressing or eliminating P_l terms, and including both A_l ’s and B_l ’s when matching the

TABLE VI. Difficulties setting up expressions match the boundary conditions in spherical. Percentages are of just the students who had difficulties setting up the boundary conditions (19%, N=90 of 485). Codes are not exhaustive or exclusive but represent the most common themes, thus the total N in the table need not sum to 90.

Difficulty	N	Percent
Not plugging in $r = R$ for the surface boundary	21	23
Problems with P_l terms e.g., expressing P_l terms incorrectly, dropping P_l terms inappropriately	16	18
Including both A_l 's and B_l 's in a single expression e.g., $(A_l R^l + \frac{B_l}{R^{l+1}})P_l = V_o P_l$	10	11
Never applied the surface boundary condition	18	20

boundary condition at the surface despite previously setting A_l or B_l to zero.

In interviews, students tended to move quickly back and forth between identifying boundary conditions and setting up equations to match them. For example, all five participants who solved the spherical SoV question (see Appendix G) began by identifying one of the two implicit boundary conditions (element SC1) and using it to correctly eliminate either the A_l or B_l terms (element SC3). All of these students then moved on to matching the boundary condition at the surface without commenting on either the second implicit boundary condition or what region their expression would be valid for. Three interviewees correctly set up an expression to match the surface boundary condition (element SC3). One of the remaining students did not plug in $r = R$ into his expression until prompted, while the other student did not initially isolate like terms when solving for constants. When asked where their final expression was valid, all five interviewees initially argued it would be valid everywhere. Once they were specifically directed to consider limiting values of r , all interviewees recognized their solution was inconsistent with the remaining implicit boundary conditions, but only one student spontaneously considered the possibility of having separate expressions for $V(r)$ inside and outside the sphere. Thus, the interviews suggest the tendency of both exam and interview students to not spontaneously acknowledge some or all of the implicit boundary conditions may discourage them from recognizing that their solution is valid only for certain regions of space or *vice versa*.

Execution of the mathematics: The Execution component deals with the procedural aspects of working through the mathematics of a physics problem. The first two elements of Execution address the process of separating Laplace’s equation into ODEs by assuming the separated form of the potential (i.e., $V(r, \theta) = R(r)\Theta(\theta)$). In spherical coordinates, this process yields a single, general solution for the potential. Students in junior electrostatics are typically shown this derivation once and are rarely (if ever) expected to replicate it. Thus elements SE1 and SE2 are not typically necessary for problems involving spherical SoV.

Once a student has used the boundary conditions to set up expressions for the unknown constants (element SC3), there are any number of mathematical manipulations that may be necessary to solve for these constants (element SE3). As described previously, we have noted three common strategies that can be used in this process (see Sec. III D 4): zero matching, Fourier’s trick, and term matching. Of the exam solutions that showed explicit evidence of Execution (92%, N=469 of 509), nearly all (97%, N=455 of 469) used some form of zero matching to eliminate one set of constants (A_l ’s or B_l ’s). The majority of solutions also used term matching (89%, N=405 of 455) to solve for the non-zero constants, while only a small fraction (12%, N=56 of 455) used Fourier’s trick. This strong preference for term matching is appropriate and is likely a reflection of the fact that nearly all surface boundary conditions given on exams at CU can be expressed as a sum of 1-3 Legendre polynomials. Moreover, several of the exam prompts provide or explicitly ask students to express $V(R)$ in terms of Legendre polynomials.

When solving for the values of the unknown constants (element SE3), roughly a quarter of students’ solutions (27%, N=125 of 469) contained various mathematical mistakes. Common issues included (see Table VII): losing or gaining a constant factor, keeping or losing P_l terms inconsistent with the boundary condition, and not finishing the calculation. The fraction of solutions with mathematical errors was higher in solutions that utilized Fourier’s trick to determine non-zero constants (60%, N=34 of 56) than in solutions that utilized term matching (21%, N=84 of 405). This is likely due, at least in part, to the fact that Fourier’s

TABLE VII. Common difficulties when executing the procedural mathematics of solving for constants in the general solution. Percentages are of just the students who exhibited these difficulties (27%, N=125 of 469). Codes are not exhaustive or exclusive but represent the most common themes, thus the total N in the table need not sum to 125.

Difficulty	N	Percent
Incorrect term matching e.g., keeping too many or not enough P_i 's in the expansion	32	26
Off by a unitless constant factor or sign e.g., factor of 2	27	22
Off by a unitfull constant factor or sign e.g., V_o , or radius R	22	18
Not finishing the calculation	24	19

trick represents an inherently more mathematically demanding strategy.

The final element in Execution (element SE4) deals with compiling all aspects of the solution into a single expression for the voltage. Roughly three-quarters of the solutions (73%, N=374 of 509) were completed enough to potentially include a final expression for the voltage, and most (83%, N=313 of 374) did so correctly. Common mistakes included not compiling a final expression (23%, N=14 of 61), dropping or adding terms (25%, N=15 of 61), and not including the r -dependence from the general solution (21%, N=13 of 61). Ultimately, only a small fraction of students (8%, N=43 of 509) had difficulties *only* with elements of the Execution component (i.e., no mistakes in Activation or Construction).

The interviews provided relatively minimal insight into student difficulties in the Execution component, in part because only two of the five students made mathematical errors of any kind while solving the spherical SoV problem. Both of these students initially failed to include the r -dependence from the general solution when compiling their expression for the voltage. Comments made by these two students suggested that they were focusing on how their final expression matched the boundary condition at $r = R$. As the boundary condition does not have r -dependence, this may account for these students leaving the r -dependence out of their final expression. The overall success of the interviewees with respect to Execution may be due in part to both the simplicity of the given boundary condition

($V(R) = V_o(1 + \cos \theta)$) and the fact that all of the interviewees used term matching rather than Fourier's trick to solve for the non-zero constants. Thus the mathematical manipulations required for this problem were minimal and purely algebraic. Overall, analysis of both the interviews and exam solutions suggest that Execution rarely represents the primary barrier to student success on spherical SoV problems.

Reflection on the Result: We identified four reflective checks that a student could use to gain confidence in (or detect problems with) their solution to problems involving spherical SoV (elements SR1-4). Only a small fraction of our students made explicit, spontaneous attempts to check their final expressions (8%, N=27 of 360) and the majority of these did so only by checking boundary conditions (70%, N=19 of 27). In interviews, two of five students made spontaneous attempts to check their solution, one by looking at units and the other at boundary conditions. One additional student suggested checking units after being asked how he might convince himself his solution was correct.

Two of the exam prompts directly targeted element SR3 by asking students to comment on why they might expect the first term in the potential outside the sphere to behave as $1/r$ (the given surface voltage was everywhere positive, see Appendix E). A completely correct response requires that the student recognize that if the voltage is everywhere positive, then the sphere must have net positive charge on its surface, and thus would look like a point charge in the limit of large r . Of the solutions to these two exams, only a small fraction (8%, N=6 of 72) articulated this argument fully. Common alternative justifications included that $1/r$ was the dependence for a point charge but made no comment about the charge on the sphere, or that $1/r$ goes to zero at infinity which matches the boundary condition. Similarly, all three of the interview students that were asked about limiting behavior of the potential needed explicit guidance before recognizing that the overall sign of the potential can be used to infer the sign of the total charge. If a significant fraction of our students have difficulty producing an expectation for the behavior of the potential at large r , this may have contributed to why spontaneous checks of limiting behavior were so rare.

TABLE VIII. Number of exam students who explicitly utilized each of the four possible reflective checks (N explicit) along with the number of solutions that included an error that would have been detected by this check (N incorrect). Total N represents the total number of solutions that reached a point where they could have utilized that reflective check. The limiting behavior check applies only to exams that asked for $V_{outside}$, and the two semesters in which students were directly prompted to consider limiting behavior are excluded; this accounts for the lower Total N.

Reflective check	Total N	N incorrect	N explicit
Units (SR1)	315	39	2
Boundary conditions (SR2)	382	119	22
Limiting behavior (SR3)	157	27	5
Laplace's equation (SR4)	365	51	1

It is also possible that more of the exam students performed one of these reflective checks spontaneously, but did not explicitly write it down on their solution. To address this, we can also examine the fraction of solutions that included errors in their final expressions that would have been detected by one or more of these checks. Table VIII lists this along with the number of students who explicitly utilized each reflective check. These results suggest that explicit checks of boundary conditions are both the most common (though still rare) and potentially the most effective in terms of catching errors.

6. *Summary and Implications*

We investigated upper-division student difficulties when using the separation of variables technique to solve Laplace's equation in the context of junior electrostatics by examining students' solutions to exam questions and think-aloud student interviews. We found that our students encountered a number of issues when solving SoV problems, and that these difficulties differed for problems involving spherical and Cartesian geometries. The ACER framework helped us to organize and categorize these difficulties within the problem-solving process.

For Cartesian SoV, we found that our students were highly successful in terms of recognizing SoV as the appropriate mathematical technique when presented with canonical Cartesian

SoV questions. Alternatively, a subset of our students used a general expression for the potential that did not satisfy Laplace's equation, possibly because they were focusing primarily on satisfying the boundary conditions. Moreover, despite a relatively high rate of success on exams, we observed a number of issues in interviews relating to recalling/justifying the separated form of the potential (i.e, $V(x, y) = X(x)Y(y)$), applying the correct logic to separate Laplace's equation into several ODEs, and recalling/justifying the need for the infinite sum. We suspect that this apparent disconnect between student performance on exams and interviews may be a reflection of students pattern matching on exams without being able to fully justify their steps. Pattern matching, as a less robust strategy, is unlikely to be as effective in the interviews which take place after the course was completed.

We also found that when solving for non-zero constants in Cartesian SoV, our students had a strong preference for Fourier's trick over term matching when both strategies were possible. This preference, while understandable given the types of boundary conditions students are accustomed to, may have exacerbated student difficulties with the procedural mathematics as Fourier's trick represents a more mathematically demanding strategy. Finally, very few of our students made spontaneous attempts to reflect on their solutions despite the fact that these strategies, particularly checking boundary conditions, can be highly effective at detecting errors made earlier in the solution.

For spherical SoV, we found that our students were sometimes less successful in terms of recognizing SoV as the appropriate mathematical technique than for Cartesian SoV. Interviews suggest that this may be due, in part, to a failure to activate Laplace's equation as the underlying equation that needs to be solved. Many of the issues that arose in Cartesian SoV related to working through the process of separating Laplace's equation and solving the resulting ODEs; however, as students are typically not required to work through this process in spherical, these difficulties were not observed for spherical SoV.

We also found that many students did not identify and/or spontaneously identify all implicit boundary conditions on the potential in spherical coordinates. As boundary conditions

are typically explicitly given in Cartesian SoV, this difficulty was not observed in Cartesian problems. Moreover, we found that in contrast to the tendency to use Fourier's trick for Cartesian SoV, our students had an appropriate preference for term matching when solving for non-zero constants in spherical SoV problems. Students' tendency to prefer one strategy over the other is likely a reflection of the canonical kinds of boundary conditions that are used for these two different geometries. Consistent with this and the idea that Fourier's trick represents a more mathematically demanding strategy, fewer students had difficulty with the procedural mathematics in spherical problems.

The idea that students were solving SoV problems algorithmically was also supported by student work around spherical SoV. For example, we consistently found that students make unjustified simplifications on spherical exam problems (e.g., setting A_l 's to zero without explanation), and one interview student made explicit comments about recalling that there was an explicit procedure for solving these problems but was not able to reproduce it. Consistent with the results for Cartesian, we again found that students rarely made spontaneous attempts to reflect on their final solutions when solving spherical SoV problems despite making a number of errors that would have been detected through one or more of these checks.

While it was not the goal of this study to investigate the impacts of different instructional strategies or curricular materials on the prevalence or persistence of students' difficulties with separation of variables, our findings do suggest several implications for teaching SoV in electrostatics. First, for problems in Cartesian coordinates, both the introduction of the separated form of the potential and the infinite sum are critical pieces of the solution that students have difficulty clearly justifying and/or coming up with spontaneously. It may be particularly important to directly target these two issues in order for students to form a more robust conception of the SoV technique. For example, getting students to come up with the need for the infinite sum on their own (i.e., asking them to try solving for the final non-zero constant(s) without it), rather than simply telling them it was necessary, seemed

to be a particularly productive exercise for our interview students.

It is also worth acknowledging that solving SoV problems algorithmically, while undesirable, is often a highly effective strategy, in part because there are a finite number of solvable SoV questions and they are all fairly similar. However, we have identified several variations on these canonical questions that may help to discourage students from purely pattern matching, particularly for Cartesian questions. For example, placing the non-zero boundary condition on either the $y = 0$ or $x = 0$ sides of a rectangular box can complicate the simplification of the exponential term. Alternatively, providing a function rather than a constant for the non-zero boundary condition (i.e., $V(x, y = 0) = V_o \sin(\pi x/a)$) can also force students to adapt their normal Cartesian SoV procedure. This latter strategy would also provide an avenue for an explicit discussion of when Fourier's trick vs. term matching represent the most efficient strategy for solving for the unknown constants. For spherical SoV, asking for the potential between two nested spherical shells may also discourage pattern matching as neither the A_l 's or B_l 's go to zero in this case.

E. Synthesis and Discussion

Having applied the ACER framework to three distinct mathematical tools and techniques, we can begin to identify common themes around students' use of mathematics in upper-division electrostatics. For example, there is a general sentiment expressed by some physics faculty that "students just don't know the math; they really just need to learn the math better and they will be successful at solving physics problems." However, we find while up to half of our students made mistakes related to the procedural mathematics (i.e., Execution) of exam questions, this was the sole issue for less than a tenth of students. All of the remaining students who did not have perfect solutions had difficulty with elements of Activation and Construction in addition to any issues they may have encountered with the procedural mathematics. The only exception to this trend for our students was for separation of variables in Cartesian coordinates where roughly a quarter of the students' solutions

included *only* Execution errors.

We also found that as many as a quarter to half our students had difficulty identifying the appropriate mathematical tool for a given problem (i.e., Activation) when not explicitly told what tool to use. Once again, Cartesian separation of variables was the exception to this trend; however, this may well be due to the highly distinctive nature of canonical Cartesian separation of variables problems. We also found that standard exam questions at CU often explicitly prompt students to use a particular tool or technique to solve, and thus provide little to no insight into whether students understand *when* these tools are appropriate. Alternatively, canonical exam questions rarely directly prompt students to reflect on their solutions (i.e., Reflection), and for all three mathematical tools discussed here, less than a tenth of CU students made explicit, spontaneous attempts to check or interpret their final expressions.

Often, the most significant and persistent difficulties encountered by our students when utilizing these mathematical tools in electrostatics related to customizing each of these tools to the specific physical situation in each problem (i.e., Construction). The issues that arose in this component were highly dependent on both the tool in question and the nature of the specific prompt. Elements of the Construction component are often those that are unique to physics problems and depend on both a student's conceptual and mathematical resources. Students do not typically encounter this aspect of mathematical problem solving in the context of solving purely mathematical problems in their math courses.

The applications of the ACER framework to the specific mathematical tools and techniques presented in the last three sections not only offers insight into common student difficulties, but also helps us to discuss the framework itself. In all cases, the ACER framework provided a consistent scaffold from which we grounded our analysis of what students actually did when solving mathematically demanding physics problems. Moreover, the process of operationalizing the framework forced us to fully and clearly explicate the resources and strategies (both tacit and explicit) that are required to solve specific types of physics prob-

lems. This helped to more clearly identify the sources of students' difficulties in terms of what is actually involved in mathematical problem solving. It also helped to provide actionable implications for instructors trying to overcome these difficulties in their classroom. Additionally, both the framework and the findings that resulted from applying it to different mathematical tools were designed to be accessible to physics instructors who are not familiar with qualitative methods or theoretical constructs related to the nature of knowledge and learning.

There are several important limitations to the ACER framework. As with any expert-guided description, it should not be assumed *a priori* that the operationalized ACER framework will span the space of all relevant aspects of actual student problem solving; however, additional research comparing the operationalized framework, as produced by the expert task analysis, to interviews and group problem-solving sessions will be necessary to explore the limitations of ACER in terms of capturing emergent aspects of students' work. The framework was designed to target the intersection between mathematics and physics in upper-division physics courses, and it is not well suited to describing student reasoning around purely conceptual or open-ended problems. Additionally, the framework inherently incorporates some aspects of representation because the translation between verbal, mathematical, graphical, and/or pictorial representations is almost always required to solve physics problems; however, the exact placement of multiple representations within the framework is likely to be highly content dependent. Furthermore, we have not commented on the integration of prediction and metacognition into the framework, in part because we rarely observe our students showing explicit signs of either without prompting.

IV. CHAPTER IV: DEVELOPING THE MULTIPLE-RESPONSE CUE DIAGNOSTIC

A. Introduction

1. *Background and Motivation*

Research-based conceptual assessments represent one of the most commonly adopted tools to come out of the PER community in the last several decades. At the introductory level, the FCI [25] is arguably the most well-known of these assessments; however, many other instruments, spanning multiple topical areas, have also been developed (see Refs. [25, 26, 72] for examples and Ref. [73] for a more comprehensive list).

Fewer conceptual assessments have been developed to target upper-division physics content, in part because conceptual assessment at the advanced undergraduate level presents some unique challenges. For example (as discussed in Chapter II), advanced physics content requires students to employ sophisticated mathematical tools and techniques. This increased emphasis on mathematics makes it more difficult, and perhaps less desirable, to create assessments that focus only on students' conceptual understanding. Additionally, the increased complexity of the physics content makes it challenging to construct clear, level-appropriate questions that can be answered within a reasonably short time frame. The relatively small body of existing research on students' difficulties also makes it more difficult to create questions that specifically target areas where students struggle. Various logistical constraints on the development of standardized assessments also represent a more significant barrier at the upper-division level than the introductory level. For example, less consistency in content coverage between different instructors and institutions makes it more difficult to create a single instrument that matches the learning goals of a majority of courses/instructors. Additionally, small class sizes at the upper-division level hinder efforts to collect enough

Significant fractions of Chapter IV are taken from Refs. [51, 74] on which BRW is first author. All text taken verbatim from these publications was written by BRW.

early-implementation data to achieve sufficient statistical power to ensure the validity and reliability of a new instrument.

Despite these challenges, several conceptual assessments have been developed for the upper-division level, targeting a range of content areas that include (but are not limited to): sophomore classical mechanics [28], junior electricity and magnetism [11, 27, 75], quantum mechanics [12, 24, 76], and several engineering assessments targeting thermodynamics [77–79] and waves [80]. Additionally, several assessments developed for the introductory and sophomore level have been used productively as pre/post-tests at the upper-division level [81, 82]. Note that, here, we are using the term ‘conceptual assessment’ broadly and include in this category assessments that target aspects of mathematical thinking (rather than procedural mathematics) and strategic processes and practices (e.g., identifying the correct solution method).

Here, we will focus on one of these upper-division conceptual assessments which targets content from junior-level electrostatics: the Colorado Upper-division Electrostatics (CUE) Diagnostic [27]. The CUE was developed at CU in order to facilitate assessment of the effectiveness of the newly developed transformed course materials for E&M 1 [10]. Topics on the CUE align with the first 5 chapters of Griffiths’ text [30] including some minimal magnetostatics. The scope and format of the CUE was specifically designed to address explicitly articulated learning goals developed through collaboration between physics faculty and PER post-docs at CU [35]. These goals represent a consensus of what these faculty want students to be able to do after completing our upper-division courses. Several examples of these consensus learning goals are given below (see Ref. [83] for the full list).

Math/physics connection: Students should be able to translate a physical description of an electrostatics problem to a mathematical equation necessary to solve it.

Communication: Students should be able to justify and explain their thinking and/or approach to a problem or physical situation.

Problem-solving techniques: Students should be able to choose and apply the problem-

solving technique that is appropriate to a particular problem.

These course-scale learning goals are tightly linked to the physics content; however, they also emphasize more meta-level outcomes related to the problem-solving strategies and habits of mind characteristic of professional physicists and highlight the importance of students' ability to synthesize and generate responses. For this reason, the developers of the CUE decided that a free-response (FR) format would more adequately test these consensus learning goals, and thus would be more valuable to faculty than a multiple-choice instrument. Additionally, relatively little literature on student difficulties around problems related to upper-division electrostatics was available to inform the development of tempting multiple-choice distractors. The developers anticipated that once established, the FR format might provide the insight into students' reasoning necessary to craft a multiple-choice version of the assessment at a later date.

While the open-ended nature of the questions on the CUE has the potential to elicit a large variety of student responses, creating clear and reliable rubrics for such questions requires a complex and nuanced grading scheme that details correct scoring for common student answers (e.g., Fig. 17). Tests of inter-rater reliability for the CUE grading rubric showed that some training was necessary for new graders to produce consistent scores using this rubric [27]. Since its development, the CUE has been given in multiple courses and institutions. CUE scores correlate strongly with other measures of student learning, such as overall course and BEMA score, and are sensitive to different types of instruction (e.g., interactive vs. traditional lecture) [27]. Data from student responses on the CUE have also been used by researchers to investigate and identify student difficulties (e.g., [18, 19, 60, 84]).

While the CUE has proved relatively successful as a standard measure of student learning across a small number of institutions and courses, this process has been heavily facilitated by researchers at CU. A significant fraction of the external instructors have sent their students' CUE's to the PER group at CU for grading and analysis. This strategy is highly impractical if the CUE is to be used as a large-scale assessment tool like the multiple-choice instruments used at the introductory level. However, the time-consuming nature of both the grader

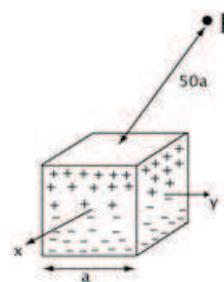
Give a brief outline of the EASIEST method you would use to solve the problem.

DO NOT SOLVE the problem, we just want to know:

- (1) **The general strategy (half credit) and**
- (2) **Why you chose that method (half credit)**

Q3. A neutral, non-conducting cube with side length a and charge density, $\rho(r) = kz$.

Find \vec{E} (or V) at point P, where P is **off-axis**, at a distance $50a$ from the cube.



Q3 Rubric

Correct Answer (3 pts)	<p>Correct answer is multipole expansion using the dipole component</p> <p>+1 point for saying direct integration</p> <p>+2.5 for dipole only</p> <p>+0.5 for approximation or multipole</p> <p>+2 for multipole</p> <p>+1 for dipole or approximation</p>
Explanation (2 pts)	<p>Full credit for saying dipole dominates because the observation point is far away</p> <p>+1.5 points for “multipole because $r \gg a$”</p> <p>+1 point if said that it’s a dipole but give no further explanation</p> <p>+1 point if mention higher order multipoles (but not a dipole)</p> <p>+0.5 for saying the integration is hard because P is off-axis.</p> <p>If answer “direct integration” must give explanation of how they would solve this integral. 0.5 for a poor explanation of how they would go about it (e.g., writing down Coulomb’s Law)</p>

FIG. 17. An example question from the CUE along with the associated scoring rubric. Graders undergo roughly 5 hours of training in which they independently score 10-15 example exams and compare their scores to accepted scores in order to use this rubric to produce consistent scores that can be compared to published results. Question prompt has been paraphrased; see Ref. [31] for full question prompt, scoring rubric, and scoring materials.

training and grading process itself represent a significant barrier to independent faculty adoption of the CUE. If the CUE is to be used as an assessment tool by a wide range of faculty, it must be adapted to a more easily graded format without sacrificing its ability to provide a meaningful measure of students’ conceptual understanding in upper-division electrostatics.

To address these scalability and usability issues, we set out to craft a more easily graded version of the CUE using student solutions from previous semesters to construct distractors. Early attempts quickly showed that a standard multiple-choice assessment, with a single unambiguously correct answer and 3-5 clearly incorrect (though tempting) distractors, was insufficient to capture the variation in responses found on the majority of the FR questions. Instead, we developed a version where students select multiple responses and receive partial credit depending on the accuracy and consistency of their selections in order to match the more nuanced grading of a FR assessment [85]. For the remainder of this section we will refer to this new version as the coupled, multiple-response (CMR) CUE.

2. Previous Research on Multiple-choice testing

Instructors and researchers have been considering the relative advantages of multiple-choice (MC) and free-response (FR) testing formats for many years [86–89]. While the difficulties inherent in writing and interpreting MC questions have been well documented [90–92], the ease and objectivity of grading afforded by MC formats makes them ideal for large-scale and cross-institutional testing. These logistical advantages have motivated considerable work to develop strategies for constructing valid and objective multiple-choice items [93–95]. This includes discussion of how many MC distractors should be used [96] and methods for accounting for guessing [97].

The increase in the number and usage of MC tests has not, however, marked the end of the tension between MC and FR testing formats, and direct comparisons of the two have yielded mixed results. Kruglak [98] looked at tests of physics terminology with both MC and FR formats. He found that while the MC questions showed a slightly higher correlation with overall achievement in the course than the FR questions, scores on the two different formats showed only low to moderate correlation with one another. This finding disagrees with results from Hudson and Hudson [94], who found a high degree of correlation between students' scores on MC tests and a set of FR pop-tests particularly when using an aggregate

score from all the MC tests given in the semester. A related study by Scott et. al. [99] compared students' scores on a MC final exam to their performance on the same problems in a think-aloud interview setting. Ranking the students based on each of these scores, they found a high correlation between the two ranks.

Additional work comparing MC and FR formats has focused specifically on the FCI. For example, Lin and Singh [85] looked at two FCI questions in MC and FR formats. As with all FCI items, the distractors for these questions were based on common student difficulties. However, rather than being scored dichotomously, scores were weighted to reflect different levels of understanding. Comparing scores on MC and FR versions of these questions, they found that average scores on the two formats did not differ significantly, and that both formats had equivalent discrimination. A similar study reported a moderate correlation between student performance on FCI questions and FR exam questions targeting the same physics content [100]. However, this study also found that the nature of individual student's responses varied significantly between the FCI and exam questions.

There have also been attempts to narrow the gap between MC and FR formats by modifying the 'standard' MC format characterized by a single unambiguously correct answer and 3-5 distractors. For example, a conceptual inventory developed by Rhoads and Roedel [101] contained a number of items with a 'multiple-correct' answer format. These items offered four response options including 2-3 correct options targeting different levels of Bloom's Taxonomy [102] in order to differentiate between higher and lower level learning. Kubinger and Gotschall [103] also described a multiple-response format where each item has five response options, of which any number may be correct. Students were scored as having mastered the item only if they selected all correct responses and none of the incorrect ones. This study found that measures of item difficulty did not differ significantly between the multiple-response format and FR formats.

Overall, previous research on MC testing has highlighted a number of similarities and differences between MC and FR formats. Agreement between the two formats appears

augmented when: the MC distractors are based on known student difficulties, the scoring of each MC distractor is adjusted to reflect varying levels of student understanding, the format of the MC items is adjusted to reduce the impact of guessing, and/or a large number of MC items are used in the comparison. We specifically leveraged the first three of these heuristics in our development of the new CMR version of the CUE.

3. *Testing Theory*

A key step in the development of a standardized conceptual assessments in either a MC or FR format is establishing the validity and reliability of the instrument. Two common perspectives on test development are Classical Test Theory (CTT) [104] and Item Response Theory (IRT) [105]. The majority of conceptual assessments in physics, both at the introductory and upper-division level have been validated using CTT, while only a small number have been developed or analyzed using IRT [106–109]. CTT posits that a high quality assessment must have 5 characteristics [104, 110]:

1. Validity - Does the test measure what it claims to measure?
2. Reliability - Does the test measure what it measures consistently?
3. Discrimination - Does the test differentiate between high and low performers?
4. Good comparative data
5. Tailored to the population

Additionally, CTT provides a number of test statistics that quantify how well an instrument matches these characteristics. For polytomously scored assessments, these statistics include [104]: item difficulty as measured by the average score on each individual item, item discrimination as measured by Pearson Correlation Coefficients [111] of item scores with the rest of the test, internal consistency as measured by Cronbach’s Alpha [112], and whole test discrimination as measured by Ferguson’s Delta [105]. Each of these test statistics will be discussed in greater detail in Chapter V. For dichotomously scored assessments, several of the test statistics used are slightly different (see Ref. [104] for an overview); however, as

both versions of the CUE are polytomously score, we will not provide further elaboration on statistical tests for dichotomously scored instruments.

One significant drawback of CTT is that all test statistics are population dependent. As a consequence, there is no guarantee that test statistics calculated for one student population (e.g., physics students at a community college) will hold for another population (e.g., physics students at a university). For this reason, scores on assessments validated through the use of CTT can only be clearly interpreted inasmuch as the student population matches the population with which the assessment was validated. For additional discussion of the limitations of CTT, see Ref. [113]. To address the shortcomings of CTT, psychometricians later developed Item Response Theory. In the simplest IRT model (i.e., the Rasch model [114]), a students' performance on individual items is assumed to depend only on their latent ability and the item difficulty. More complex IRT models also include parameters to account for item discrimination and student guessing. For test items that fit this model, all item and student parameters can be determined in such a way as they are independent of both population and test form [105, 115].

Despite the appeal of generating population-independent parameters, there are several significant drawbacks to IRT as a potential tool to develop upper-division physics assessments. Even the simplest dichotomous IRT models require large N (>100) to produce estimates of item and student parameters that are reliable enough for low-stakes testing [105, 116]. This number increases for more complex models that, for example, include item discrimination parameters, or for instruments with polytomous scoring [116]. The small class sizes typical of upper-division physics would necessitate classroom testing at multiple institutions, possibly over multiple semesters, to collect this volume of data. Additionally, in order for the parameters generated by IRT to be truly population independent, they must fit the appropriate IRT model. Crafting a large number of items that fit these models often requires multiple iterations of preliminary testing, further increasing the number of students necessary to develop and validate an assessment. Due in large part to the logistical barriers to IRT, the

analysis in this thesis will exclusively utilize CTT.

B. Creating a Multiple-response Version

1. *Adapting the Questions*

As our goal was to create a new version of an already established and validated conceptual assessment, we explicitly avoided making substantive changes to the content coverage or questions on the CUE. There are several distinct question styles on the original FR CUE, but roughly two-thirds of the items present students with a physics problem and ask them to state the appropriate solution method or select an answer and then to justify that answer/method (Fig. 17 gives an example of this type of item). These items are scored using a validated grading rubric that includes established point values for common responses [27].

Preliminary distractors for the CMR CUE were developed using both the FR grading rubric and student responses to the FR items. We began by compiling a list of *a priori* codes based on common responses identified on the FR rubric (see Fig. 17). These codes were then used to characterize multiple semesters of student responses to the FR version. The initial list of codes was expanded during this process to account for emergent response patterns not already encompassed by the *a priori* codes. By looking at the frequency of each code, we identified the most common justifications provided on each question. For many of the CUE questions, a completely correct justification requires the student to connect several distinct ideas. For example, a complete response to the item shown in Fig. 17 must include several pieces: (1) that the potential can be expressed using the Multipole expansion, (2) that the cube looks like a dipole, and (3) at large distances the first non-zero term (the dipole term) will dominate. However, both the *a priori* and emergent coding showed that many students gave partially correct justifications that were missing one or more key elements. This was especially true for the Method/Reasoning style items like the one in Fig. 17. A standard MC format requires identifying and articulating a single, unambiguously correct

answer along with three to five tempting but unambiguously incorrect distractors. Our early attempts to construct distractors satisfying these requirements failed to capture either the variation in justifications or the partially correct ideas that were coded from the FR items.

To accommodate the wide range of correct, partially correct, and incorrect reasoning that students gave on the FR version, we switched to a less restrictive multiple-response format. After asking the students to select the correct answer or easiest method, we provide a set of what we are calling reasoning elements, of which students can select all that support their choice of method (see Fig. 18). Reasoning elements were taken from common codes identified in the student responses to the FR version, and each one may be correct, incorrect, or irrelevant in the context of justifying a particular answer or choice of method. Full credit requires that a student select all (and only) the reasoning elements that together form a complete justification; however, they can receive partial credit for selecting some, but not all, of the necessary elements (see Sec. IV B 2). This format allows for a wider range of justifications than a standard multiple-choice test as students can combine reasoning elements in many different ways. It also requires students to connect multiple ideas to construct a complete justification, thus providing an avenue for students to demonstrate partially correct or incomplete ideas.

This multiple-response format is similar to the one discussed by Kubinger and Gotschall [103]. However, the response options they used were individually marked as true or false, making it unnecessary for the student to make any kind of connections between the statements. In contrast, our items often include reasoning elements that are true but irrelevant, true in some physical situations, or true but incomplete. A fully correct response with this format requires that students be consistent both between reasoning elements and method/answer selection, and between reasoning elements.

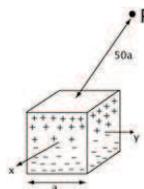
The CMR version of the question shown in Fig. 17 is given in Fig. 18. The boxes next to each reasoning element are intended to facilitate students' interaction with the new format by resembling 'check all' boxes that the students are familiar with from online surveys. The

Q3 - A solid, neutral, non-conducting cube as below, with side length 'a' and $\rho(z) = kz$.

Find \vec{E} or V at point P, where P is **off-axis**, at a distance $50a$ from the cube.

Select only one: **The easiest method would be ...**

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. the cube will look like a dipole; approximate with \vec{E} or V for an ideal dipole
- c. symmetry allows you to calculate \vec{E} using a cubical Gaussian surface
- d. symmetry allows you to calculate \vec{E} using a spherical Gaussian surface
- e. the observation point is **far** from the cube
- f. there is not appropriate symmetry to use other methods
- g. $\nabla^2 V = 0$ outside the cube and you can solve for V using Fourier Series

FIG. 18. A sample item from the CMR CUE. Question prompt has been paraphrased; see Appendix H for full prompt and Appendix I for the rubric and scoring materials.

wording of the question prompts was adjusted only when necessary to accommodate the new format. Two questions on the FR version (Q11 & Q15, see Ref. [31]) dealt with mathematical expressions for the boundary conditions on E and V . One of these questions (Q15) was not included on the new CMR version as it already had a multiple-response format that became redundant once Q11 had been redesigned to be multiple-response. Thus the FR version has 17 numbered items, while the CMR version has only 16.

Roughly two thirds of the questions on the FR CUE explicitly ask for students to express their reasoning, and their CMR counterparts have the tiered format shown in Fig. 18. The remaining items have various formats that include interpreting or generating specific formulas (e.g., boundary conditions) as well as sketching graphs, vector fields, and/or charge distributions. These items were translated into more standard multiple-choice and multiple-response formats. We selected the simplest format for each item that allowed us

to encompass the majority of the coded student solutions to the FR version. It is not possible to provide examples of all question formats on the CMR CUE here; however, the full instrument is included in Appendix H and can also be accessed at Ref. [31].

2. *Scoring*

While allowing for a more nuanced measure of student understanding, the ‘select all’ format of the CMR CUE also sacrifices one of the logistical advantages of standard MC questions because it is not as straightforward to score this format using automated machine grading (e.g., Scantron). Student responses to the CMR CUE must instead be entered into an electronic grading spreadsheet. The grading spreadsheet is an Excel workbook that has been pre-programmed to translate students’ letter responses to each question (e.g., ‘a’, ‘B’, etc.) into numerical scores for each student using an appropriate grading scheme. Thus, once students’ responses have been entered, the electronic gradesheet instantly scores each student, preserving the fast and objective grading of a MC assessment.

Additionally, the new format also allows for considerable flexibility in terms of scoring. The CMR CUE can easily be scored using multiple grading schemes simply by modifying the grading spreadsheet. FR tests, on the other hand, require significant time and resources to regrade with a new grading scheme. There are several different potential grading schemes for a question like the one in Fig. 18 ranging from very simple to very complex. However, Lin and Singh’s previous work [85] suggests that a more complex rubric designed to reflect different levels of student understanding may be more effective at achieving consistency between the FR and CMR versions. Here we will investigate two different grading schemes: (1) a simple grading scheme designed to preserve the straightforward scoring of a standard MC rubric, and (2) a more complex rubric designed to closely replicate the nuanced grading used to score the FR CUE [27]. These two rubrics (described in greater detail below) are not the only potential rubrics that might be developed, but represent two extremes that will allow us to explore the impact of different grading schemes.

In the grading scheme used on the simpler grading rubric (R1), students are awarded full points for making correct selection and zero points for any response not including a correct selection. For the example item shown in Fig. 18, students are awarded three points for selecting the Multipole Expansion ('D') as the easiest method and up to two points for also selecting any combination of the relevant reasoning elements 'b' (0.75 pt), 'e' (0.75 pts), and 'f' (0.5 pts), for a total of five points on the question. In this scheme, reasoning points can only be received when the correct method is selected, and selection of incorrect or irrelevant reasoning elements does not impact the student's score.

Alternatively, in the grading scheme used on the more complex grading rubric (R2), students are awarded full points for making correct selections but can also receive some credit for partially correct responses. For example, the CMR CUE scoring rubric for the Method/ Reasoning questions awards full points for selecting the easiest method, and also awards partial credit for selecting methods that are possible, even if they are not easy. Additionally, the rubric awards points for reasoning elements that are consistent with the choice of method. For example, on the item in Fig. 18, it is also possible, though difficult, to use Direct Integration to solve for \vec{E} or V . The rubric awards students who select method 'A' one point for the Method and an additional half point for selecting the consistent reasoning element, 'a'. On items without the tiered format, students can still receive some credit for selecting distractors that reflect partially correct ideas or demonstrate an appropriate degree of internal consistency. The point distribution and weighting of each answer or method/reasoning combination was designed to closely match the established scoring on the FR version (see Appendix I for the full R2 grading scheme).

In addition to offering additional credit for consistency, the R2 rubric also subtracts points from students with reasoning elements that are inconsistent with their choice of method. Typically, selecting an inconsistent or incorrect reasoning element will prevent a student from getting more than three out of five points on questions that ask for explicit justifications. These consistency checks in the R2 grading scheme help to reduce the credit a student can

get by chance. On standard MC tests, a student can expect to get a score of roughly 20-25% just by guessing. Using the R2 scoring rubric on 100 computer generated responses simulating random guessing, we found an average score of 13%. Alternatively, these same responses resulted in an average score of 21% when using the R1 grading scheme which does not subtract points for inconsistent selections. Differences in students' scores using these two rubrics will be presented in Chapter V.

C. Establishing Content Validity

There are several distinct aspects to establishing the validity of an assessment instrument. Recall that in CTT (see Sec. IV A 3), validity deals with whether a test measures what it claims to measure. This includes examining how well the test items cover the targeted content domain, which is referred to as content validity [104].

1. *Expert Validation*

The content validity of an assessment is established in part by ensuring that the content is seen as accurate, clear, and valuable to content experts in the discipline. The FR CUE was designed to align with learning goals for upper-division electrostatics that were developed in collaboration with physics faculty at CU. The original instrument was also reviewed by physics experts and revised based on their feedback [27]. Since the CMR CUE has the same prompts, the validity of its physics content is, to a large extent, already established. However, the operationalization of this content has changed significantly in the new format. We solicited and received feedback from nine experts in physics content or assessment spanning six institutions, all with experience teaching upper-division physics. Small modifications were made to several items as a result of this feedback.

The majority of the changes suggested by these experts were minor adjustments to the wording of several prompts and reasoning elements to clarify either the language or the

physics content (e.g., the addition of the word ‘isotropic’ to describe the linear dielectric in Q14, see Appendix H). In one case (Q12, part iii), several experts expressed a subtly incorrect idea relating to the functional dependence of the electric field along the z-axis above a uniform disk of charge when $z \ll R$. This idea led them to the (incorrect) conclusion that the fully correct graph of $E(z)$ was not provided in the given list of response options. If a subset of physics experts believe that the correct response to this question is not included, it may impact their perception of the content validity of the assessment. To address this, we added an additional response option in which the behavior of $E(z \ll R)$ corresponded to this incorrect idea. As the difference between this graph and the correct one is subtle, the R2 grading scheme offers close to full credit (3.5 out of 4 pts) for selecting this distractor.

Overall, the expert reviewers expressed enthusiasm for the CMR CUE and offered no critiques that questioned the overall validity of the new format. However, several of the reviewers did point out that, as with all multiple-choice formats, this new format only requires students to recognize a correct answer, which is a potentially easier task than requiring them to generate it. Ideally, the research-generated distractors combined with the multiple-response format reduce the potential for students to simply recognize correct answers, particularly on the Method/Reasoning type questions where the individual reasoning elements rarely represent complete justifications. In Chapter V, we present empirical evidence that, on average, our students do not score higher when asked to select the correct answers/justifications on the CMR version than when asked to generate answers/justifications on the FR version.

With respect to the expert-guided ACER framework described in Chapters II & III, the CUE offers a much clearer measure of student performance in the Activation component than is typical of standard exam and homework questions. Alternatively, the Construction and Execution components are largely missing from the CUE. This was, in part, an intentional feature of the original CUE as it was specifically designed to target aspects of problem solving not captured by traditional exams which target primarily Construction and Execution.

Reflection is also not explicitly targeted by the CUE; however, the use of reflective checks would likely facilitate student performance on a number of questions.

2. *Student Validation*

Content validity not only requires that the content, format, and language of an assessment is clear and understandable to experts, but also to the targeted student population [104]. Think-aloud validation interviews in which individual students work through the diagnostic while explaining their reasoning helps to ensure that students are interpreting the questions, formatting, instructions, and distractors as intended. For the CMR CUE student interviews were also crucial because we were concerned that the ‘select ALL that apply’ format might be unfamiliar or confusing. We performed thirteen interviews with the full 16 question CMR CUE and three interviews with a 7 question subset of the full instrument to demonstrate the validity of the new format. All interview participants had completed an upper-division electrostatics course one to four weeks prior to the interview with final course grades ranging from A to C. During these interviews, participants were first allowed to complete the assessment without input from the interviewer. After completing the assessment, the interviewer probed the students in more detail where it was unclear why they selected or rejected certain distractors.

The student validation interviews were analyzed with a particular focus on identifying instances where the phrasing of an item caused students to select responses that were inconsistent with the reasoning they articulated verbally, or where students interacted with the new format in a way that caused an artificial inflation (e.g., selecting answers based on superficial similarities in wording) or deflation of their score (e.g., not following directions or not reading all the reasoning elements). Minor wording changes were made to several of the prompts and reasoning elements as a result of these interviews. For example, the prompt to one question (Q14) originally asked for students to comment on the value of \vec{E}_{net} inside a capacitor filled with linear dielectric; however, we found that a number of students in inter-

views responded to this question as if it was asking for the electric field from the dielectric only despite clearly recognizing that the superposition of the fields from the dielectric and capacitor would result in zero electric field inside. To counter this, we modified the prompt to explicitly express the net electric field as the superposition of the field from the capacitor and dielectric (i.e., $\vec{E}_{net} = \vec{E}_{plates} + \vec{E}_{dielectric}$).

The interviews also informed several changes to the grading scheme. For example, some items contain reasoning elements that are true but irrelevant statements. These were typically included because they appeared as a common justification for an incorrect method selection on the FR version. We found in interviews that students who knew the correct method often selected these reasoning elements simply because they were true statements. To account for this, we modified the R2 grading rubric so students who did this would not be penalized or rewarded for selecting a true reasoning element that did not directly support their choice of method.

A concern raised by one faculty reviewer was that students who did not know how to start a problem might figure out the correct approach by examining the given reasoning elements. This possibility also relates to the validity of the CUE in terms of achieving an unambiguous measure of the Activation component of the ACER framework. We did observe instances in the interviews where students would explicitly refer to the reasoning elements in order to inform their choice of method. However, this technique seemed most useful to students with higher overall CUE and course scores, and, in all such cases, the student provided additional reasoning that clearly demonstrated their understanding of the correct method. Alternatively, some students in the interviews were led down the wrong path by focusing on an inappropriate reasoning element. This suggests that using the reasoning elements to figure out the correct method does not result in a significant inflation of scores.

V. CHAPTER V: STATISTICAL VALIDATION OF THE CMR CUE

A. Comparing the Multiple-response and Free-response Versions

Following the initial development and validation of the CMR (coupled multiple-response) version of the CUE, we began classroom testing in order to produce measures of the statistical validity and reliability of the instrument across a range of student populations. As part of this process, we also performed a side-by-side comparison of the new version with the original FR (free-response) version.

1. *Methods*

Comparison data were collected during two semesters of the E&M 1 course at CU, the first of which was taught by a PER faculty member (SJP) who incorporated a number of materials designed to promote interactive engagement, such as in-class tutorials and clicker questions [10]. The second semester was taught by a visiting professor who utilized primarily traditional lecture with some minimal interactive engagement interspersed.

To make a direct comparison of the two versions of the CUE, each semester of the E&M 1 course was split and half the students were given the CMR version and half the FR version. The two groups were preselected to be matched based on average midterm exam score but were otherwise randomly assigned. Attendance on the day of the diagnostic was typical in both semesters and ultimately a total of 45 students took the CMR version and 49 students took the FR version of the CUE (75% response rate overall). The analysis presented in the remainder of this section will focus exclusively on the comparisons of the FR and CMR versions of the CUE; Sec. V B will report on the larger scale validation of the CMR CUE for independent implementation using data from different instructors and additional institutions.

The majority of Section V A is taken from Ref. [51] on which BRW is first author. All text taken verbatim from this publication was written by BRW.

2. Results: Statistical Comparison

This section presents the quantitative comparison of test statistics from both versions of the CUE. These test statistics are pulled exclusively from Classical Test Theory (see Sec. IV A 3). Using the nuanced R2 grading rubric described in Sec. IV B 2, the average score on the CMR version, $54.3 \pm 2.8\%$ ($\sigma = 19.1\%$), was not significantly different (Student's t-test [117], $p = 0.9$) from the average on the FR version, $54.6 \pm 2.8\%$ ($\sigma = 19.6\%$). Score distributions for both versions (Fig. 19) were nearly normal (Anderson-Darling test [118], CMR: $p = 0.9$, FR: $p = 0.6$) and had similar variances (Brown-Forsythe test [119], $p = 0.9$).

The average score on the CMR version from the simpler grading scheme (R1, see Sec. IV B 2) was $50.1 \pm 2.9\%$ ($\sigma = 19.7$). This score, while lower, does not differ statistically from the average scores on either the FR version or the R2 grading scheme ($p = 0.3$). We found this result to be somewhat surprising given the significant differences between the two grading rubrics. R1 gives no partial credit but also does not penalize inconsistent reasoning/method

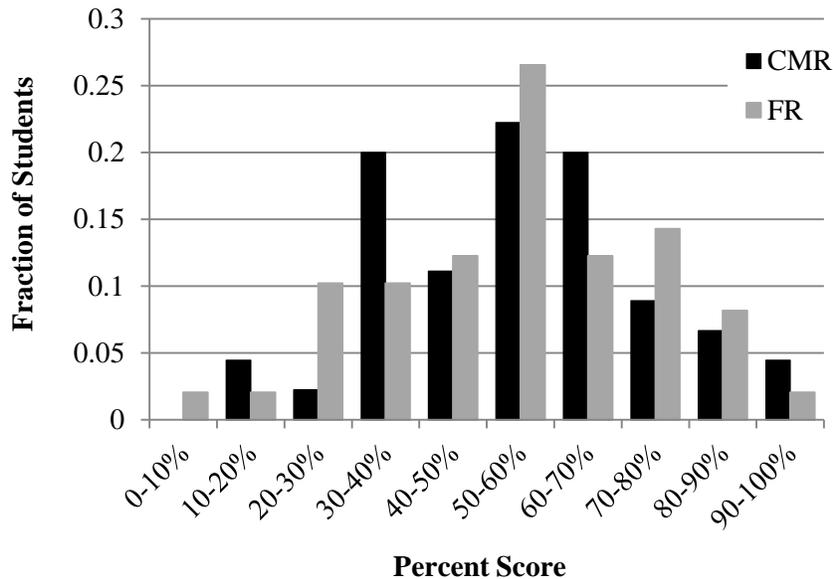


FIG. 19. Distributions of scores on the CMR ($N=45$) and FR ($N=49$) CUE for the E&M 1 course at CU. There is no statistically significant difference between averages for the two the distributions (Student's t-test, $p = 0.9$).

selections. A concern with R1 is that students who simply selected many or all reasoning elements could still get full credit. However, we find that, both in the test semesters and in interviews, our students rarely select more than 1-3 reasoning elements. On the other hand, R2 offers significant partial credit and penalizes inconsistent selections. While individual student's scores sometimes shift significantly from one rubric to the other, on average the opportunity for partial credit and the penalty for inconsistency balance out to give similar average scores on R2 as R1. The more nuanced grading of R2 was also designed to better reflect subtle differences in levels of understanding between students, and it may be that the size of this initial sample was not large enough to highlight these differences.

To further investigate the impact of different grading rubrics on the agreement between these two versions, we also scored the CMR CUE using a strict right/wrong rubric. In this grading scheme, students receive full credit for selecting the correct method and all the necessary reasoning elements. Not selecting one or more of the important reasoning elements or selecting inconsistent reasoning elements both result in zero points on the reasoning portion of that item. Using this 'perfect' scoring rubric, the score on the CMR version falls to $42.8 \pm 2.8\%$ ($\sigma = 18.5\%$). The difference between this score and the average on the FR version is statistically significant (Student's t-test, $p < 0.01$). This finding suggests that a nuanced grading rubric designed to reflect different levels of student understanding does improve agreement between multiple-choice and free-response formats. Because it more closely matches the FR rubric and our higher level learning goals for the course, the remainder of this section exclusively utilizes scores from the R2 grading scheme.

The start and stop time of each student was also recorded. On average, students spent a comparable amount of time on the CMR version, 35.0 ± 1 min ($\sigma = 7.5$ min), as on the FR version, 34.8 ± 1 min ($\sigma = 7.9$ min). The average time spent on both versions was also the same for the two semesters individually. We were initially concerned that the multiple-response format might encourage students to go through the CMR version quickly and carelessly. Given the amount of writing required on the FR version, the fact that students

took the same amount of time to complete the CMR CUE suggests that they were reading the distractors carefully and putting thought into their responses.

Criterion Validity

Another property of the CMR CUE is how well its scores correlate with other, related measures of student understanding. The most straightforward comparison is with the more traditional, long-answer course exam scores. Students in both semesters of E&M 1 took two midterm exams and one final exam. The CMR CUE scores correlate strongly with aggregate exam scores (Pearson Correlation Coefficient [111] $r = 0.79$, $p < 0.05$). For comparison, the correlation for the 49 students who took the FR version was also high ($r = 0.79$, $p < 0.05$). Similarly the scores for both versions are strongly correlated with final course score which includes roughly 30% homework and participation (CMR: $r = 0.76$, FR: $r = 0.73$). To account for differences between the average exam, course, and CUE scores between the two semesters, the correlations above are based on standardized scores (z-scores) calculated separately for each class using the class mean and standard deviation. These correlations are not statistically different from the correlation, $r = 0.6$ ($p = 0.8$, determined using the procedure described in [120]), reported previously for the FR CUE [27].

Item Difficulty

In addition to looking at the overall performance of students on the CMR and FR versions of the CUE, we examined their performance on individual items. Fig. 20 shows the average scores on both versions for each question (item numbers are consistent with the CMR version). The score given for Q11 on the FR version is an average of the scores from two items originally numbered 11 and 15 on the FR version (see Sec. IV B 1). Differences between the scores are significant for 3 of 16 items (Mann-Whitney U-test [121], $p < 0.05$; see Fig. 20). In all three cases, the difficulty went down for the new version. The decrease in difficulty in these three questions is balanced by a marginal (but not statistically significant) increase in difficulty on several of the remaining question; thus the whole-test average is the same for both versions.

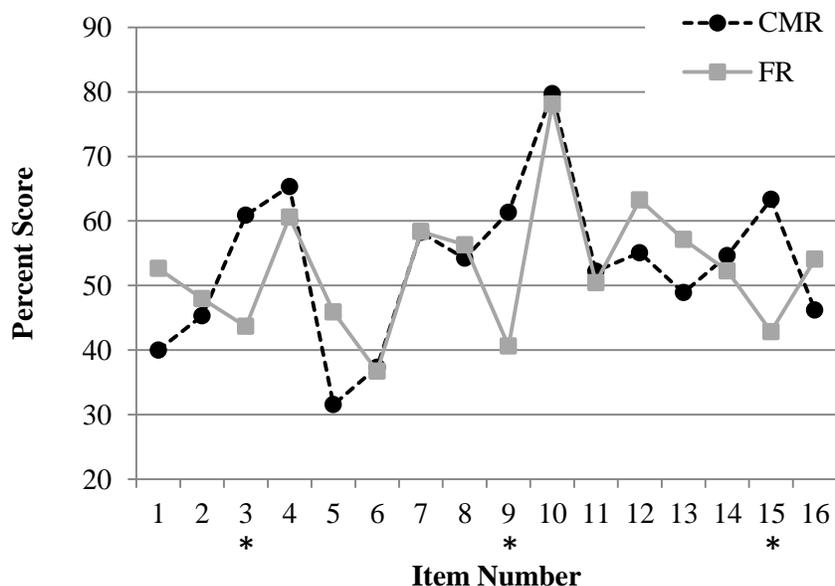


FIG. 20. Average scores on each item on the CUE. Statistically significant differences between the CMR and FR versions are indicated by an asterisk (Mann-Whitney, $p < 0.05$). See Appendix H for the questions on the CMR version and Ref. [31] for the FR version. Item numbering is consistent with the CMR version.

The FR versions of two of the items with statistically significant differences (Q9 and Q15, see Appendix H) were particularly challenging to adapt to the new format and, ultimately, underwent the most significant modification of all questions. Student interviews suggest that for Q15 the decrease in difficulty arises because the FR version contains an explanation component that was eliminated on the CMR version. Alternatively, on Q9 interviewees often recognized the appropriate justifications among the given reasoning elements even when they did not generate them which may account for the decreased difficulty. For the remaining item (Q3, Fig. 18), we had no *a priori* reason to expect that the CMR version would be significantly different than the FR version. However, student interviews suggest that, for this item, one particularly tempting reasoning element ('b') can help students to determine the correct method ('D'). See Sec. V A 3 for more discussion of Q3.

Discrimination

We also examined how well performance on each item compares to performance on the rest of the test (i.e., how well each item discriminates between high and low performing

students). Item-test correlations were between 0.24 and 0.71 for all items on the CMR and FR CUE with the exception of the CMR version of Q15 ($r = 0.17$, see Fig. 21). Q15 was also the only item that had a statistically significant difference between the item-test correlations for the CMR and FR versions. As stated previously, Q15 was modified significantly for the CMR version. A common criterion for acceptable item-test correlations is $r \geq 0.2$ [26]; however, for $N = 45$, correlation coefficients less than 0.24 are not statistically significant. Q15 on the CMR CUE is the only item on either version that falls below the cutoff for acceptability or statistical significance.

To look more closely at the source of this anomalously low correlation, we can break this out by course. The correlation is stronger and statistically significant for one of the two semesters ($r = 0.44$, $p = 0.02$, $N=25$); whereas the correlation is negative but not statistically significant in the other semester ($r = -0.11$, $p = 0.6$, $N=20$). A scatterplot of item score vs. test score for Q15 shows that the low correlation is due to a large amount of scatter in the scores rather than one or two outliers. This may indicate that this topic (dipole moments)

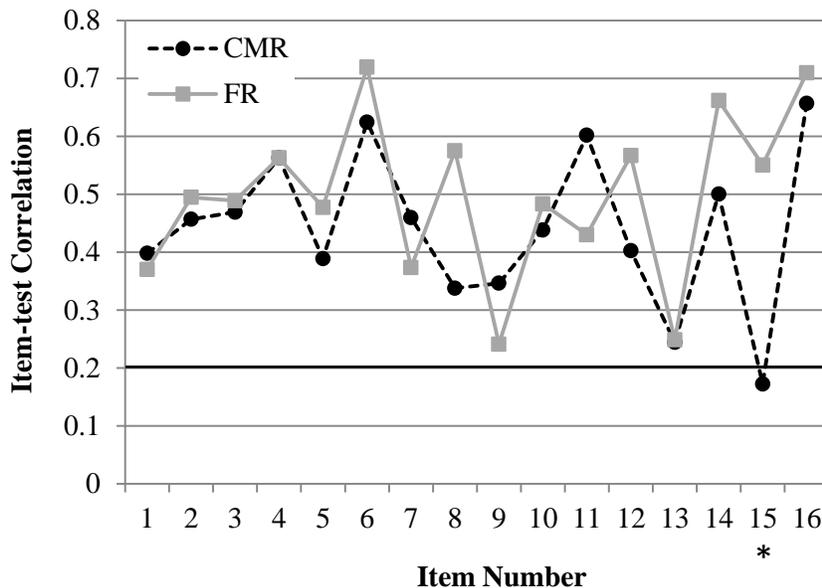


FIG. 21. Comparison of item-test correlations as a measure of item discrimination on the CMR and FR versions of the CUE. The conventional cutoff for an acceptable correlation (0.2) is marked with a bold line. Statistically significant differences between the two versions are indicated by an asterisk ($p < 0.05$, determined using the procedure described in [120]).

was not covered as effectively in this semester as compared to the other semester. Sec. VB2 will show that the low correlation demonstrated by this one course is not representative of the broader student population.

As a whole-test measure of the discriminatory power of the CMR CUE, we calculate Ferguson's Delta [105]. Ferguson's Delta is a measure of how well scores are distributed over the full range of possible point values (total points: CMR - 93, FR - 118). It can take on values between [0,1] and any value greater than 0.9 indicates good discriminatory power [26]. For this student population, Ferguson's Delta for both the CMR and FR versions of the CUE is 0.98. This is similar to the previously reported FR value of 0.99 [27].

Internal Consistency

The consistency of scores on individual items is another important property of an assessment. To examine this, we calculate Cronbach's Alpha for both versions of the test as a whole. Cronbach's Alpha can be interpreted as the average correlation of all possible split-half exams [112]. Using the point value of each item to calculate alpha, we find $\alpha = 0.82$ for the CMR version and $\alpha = 0.85$ for the FR version. Again, this is consistent with the value of 0.82 reported historically for the FR CUE [27]. For a unidimensional test the commonly accepted criteria for an acceptable value is $\alpha \geq 0.8$ [122]. While we have no *a priori* reason to assume that the CUE measures a single construct, multidimensionality will tend to drive alpha downward [112]; thus we argue Cronbach's Alpha provides a conservative measure of the internal consistency of the instrument.

3. Results: Reasoning Comparison

The previous section demonstrated a high degree of consistency between the CMR and FR versions on the CUE in terms of scores and a variety of test statistics. However, one of the primary goals of the CUE's original creators was to gain insight into student thinking and the nature of common student difficulties with electrostatics [27]. Gauging how much of this insight is preserved in the new CMR version requires a comparison of what students

wrote/selected on each version. To do this we performed a qualitative analysis of student responses to a subset of the CUE questions, Q1-Q7. We focused on these seven items because they represent all the Method/Reasoning type questions (see Figs. 17 & 18) and typically elicit the richest and most detailed explanations on the FR version.

Method Selection

We started by comparing just the students' method selections on both versions of the CUE. This approach required coding student responses to the FR version into one of the method options offered on the CMR version. The method coding process was relatively straightforward because the FR version directly prompts the students to select a solution method and provides them a list of methods at the beginning of the exam that matches the list provided on the CMR version. In a few cases, some interpretation was necessary to assign a method selection to students who did not use the precise name of the method in their response (e.g., 'use the multipole expansion' vs. 'use a dipole approximation'). Inter-rater reliability was established by two people independently coding 20% of the FR tests. Agreement on the coded method selection was 96% before discussion and 100% after discussion.

A comparison of the method selections of students taking the CMR and FR versions of one question (Q2) is given in Fig. 22 (See Appendix J for comparisons for all 7 questions). Visually, the two distributions are strikingly similar, and this trend is representative of five of the seven questions. The remaining two questions showed greater variation (see Fig. 23). To quantitatively compare the two versions, we constructed 2x9 contingency tables detailing the number of students in each group who selected each method for each of the seven questions. While χ^2 is a common statistic for determining statistical significance from contingency tables, it loses considerable statistical power when cells have $N < 5$ [123]. As many of the cells in our tables fell below this cutoff, statistical significance was determined using Fisher's Exact Test [124, 125]. Fisher's Exact Test determines the probability of obtaining the observed contingency table given that the two variables in the table have no association.

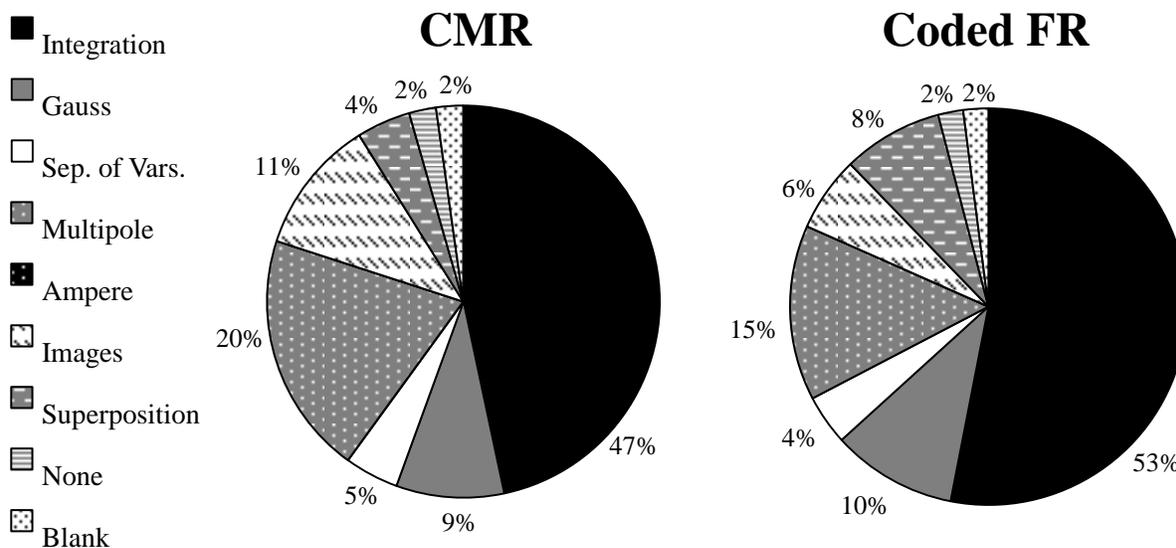


FIG. 22. Percent of students who selected each method on Q2 for each version of the CUE. The left chart represents student selections from the CMR version ($N=45$), while the right chart represents coded method selections from the FR version ($N=49$). The difference between the two distributions is not statistically significant (Fisher’s exact test, $p = 0.9$).

It then sums the probability of the observed table along with all more extreme tables to return a p-value for having observed that particular table given the null hypothesis.

Ultimately, only the two questions with visually different distributions had statistically significant differences ($p < 0.05$) between the method selections of students taking the CMR and FR CUE. For one of these two questions (Q3, see Fig. 23), students taking the CMR version were more likely to select the correct method (Multipole Expansion) and less likely to select the possible but harder method (Direct Integration). This trend is consistent with the decrease in difficulty observed for this item (see Sec. V A 2). As stated earlier, this shift may be attributable to the presence of a particularly tempting correct reasoning element. For the second of the two questions identified by Fisher’s Exact Test (Q5), students were less likely to select the correct method (Superposition) and more likely to select a common incorrect method (Gauss’s Law) on the CMR version. In this case, student interviews suggest that this trend may be due to the presence of a particularly tempting, but this time incorrect, reasoning element justifying the use of Gauss’s Law.

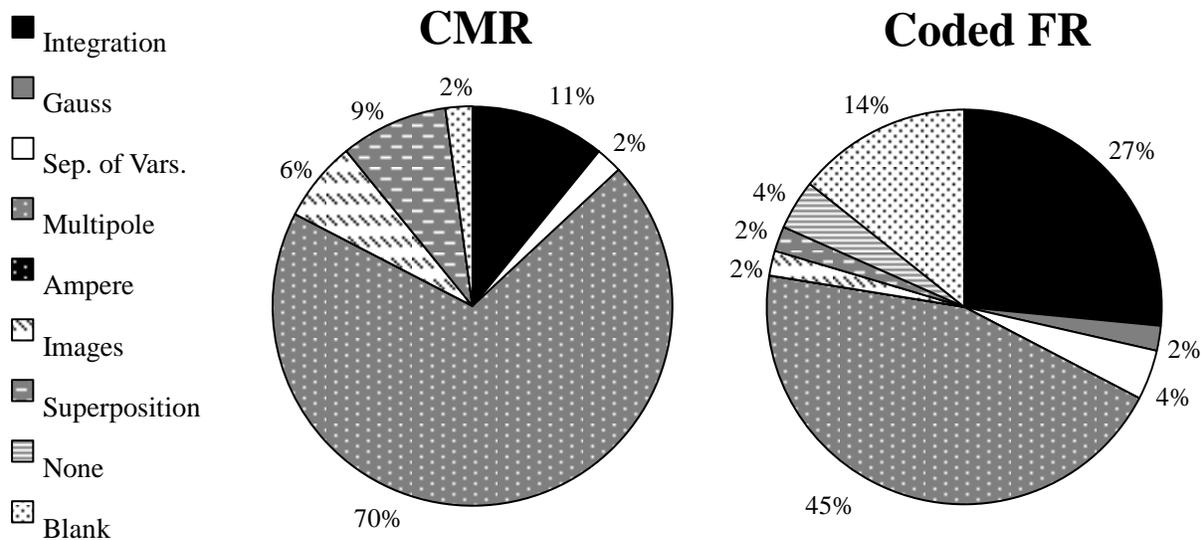


FIG. 23. Percent of students who selected each method on Q3 (see Fig. 18) for each version of the CUE. The left chart represents student selections from the CMR version ($N=45$), while the right chart represents coded method selections from the FR version ($N=49$). The difference between the two distributions is statistically significant (Fisher's exact test, $p = 0.02$).

Reasoning Selection

The reasoning portion of the FR questions was more challenging to code than the method portion because students are no longer constrained to a finite list of methods and are free to justify their answer in any way they choose. We started by coding students' free responses using the reasoning elements provided on the CMR version. We also identified aspects of student responses that did not match one of the available reasoning elements. These aspects were coded into two 'other' categories: satisfactory and unsatisfactory. Satisfactory codes were given to elements of students' justifications that represented correct physical statements that supported the choice of method but that did not get coded into one of the CMR categories. Unsatisfactory codes were given to elements that represented incorrect or irrelevant statements. Students could receive multiple codes, meaning that a student could be awarded an 'other' code even if some elements of their response fit into one of the CMR categories.

Due to the higher degree of difficulty inherent in coding the reasoning portion, inter-rater

reliability was established in two stages. Additionally, because the coding on the reasoning portion allows for multiple codes for each student, we determined inter-rater reliability statistics for both complete agreement (i.e., no missing or additional codes) and partial agreement (i.e., at least 1 overlapping code). First stage reliability statistics were generated from independent coding of 20% of the FR exams. For this initial set, complete agreement was 65% before discussion and 94% after discussion, and partial agreement was 79% before discussion and 96% after discussion. In the second stage, an additional 10% of the FR exams were independently coded, and both complete and partial agreement before discussion rose to 89%. There is not a well-accepted threshold for an acceptable percent agreement because this cutoff must account for the possibility of chance agreement and thus depends on the number of coding categories [126]. However, given our large number (7-10) of non-exclusive coding categories, the potential for complete agreement by chance is low. Thus, we consider 89% agreement to be acceptable for the general comparisons made here.

Ultimately, an average of 74% of FR students who did not leave the reasoning portion blank were coded as having one or more of the CMR reasoning elements per question. The remaining 26% received only ‘other’ codes (11% satisfactory and 15% unsatisfactory) meaning that no aspect of their justification matched one of the CMR reasoning elements. Overall, 33% of students who took the FR version received ‘other’ codes, including those who also received one or more CMR codes. In other words, one third of the information on students’ reasoning that is accessed by the FR version is forfeit on the CMR version. This result is not surprising as it is not possible to capture the full range of student reasoning with a finite number of predetermined response options.

This section presented an analysis of student responses to the Method/Reasoning type questions on both versions of the CUE. We found that the two versions elicited matching Method selections for five of seven questions. On the remaining two questions, there was a statistically significant shift in the fraction of students selecting each of the two most common method choices. In both cases, this shift may be attributable to the presence of a particularly

attractive reasoning element. Additionally, we find that roughly three-quarters of responses to the FR version contained elements that matched one or more of the reasoning options provided on the CMR CUE. However, roughly a third of these responses also contained elements that did not match the CMR reasoning options; thus, the logistical advantages of our CMR assessment come at the cost of reduced insight into student reasoning.

B. Validating the CMR CUE

1. *Methods*

Following the initial comparison of the CMR and FR versions of the CUE, we set out to more robustly establish the validity and reliability of the new version as an independent instrument. To do this, we expanded our data collection with an emphasis on including additional students and instructors at multiple institutions. We recruited instructors to pilot the CMR CUE in several ways including soliciting participants during talks and posters presenting the results of the initial comparison study at professional meetings and workshops (e.g., the American Association of Physics Teachers summer meetings). The new version was also uploaded to the online materials repository (see [31]) where it can be accessed by any physics instructor interested in using our transformed course materials. We also contacted a number of colleagues working in PER who facilitated putting us in contact with the instructor in their department who was teaching electrostatics.

Ultimately, we collected post-test CUE data from 14 courses spanning 9 institutions and 12 instructors. Institution and course characteristics are shown in Table IX. We also have pretest data from 12 of these courses. Pretests were administered in the first week of class as either 20 min in-class activities (N=6) or as an out-of-class online survey (N=6). For all courses but one, post-tests were administered in the last week of class as a 50 min in-class activity. In one case, the post-test was given as an out-of-class online survey.

To what extent in-class implementations of the pre and post-tests can be compared to out-

TABLE IX. General characteristics of each institution where we collected post-test CMR CUE data. N indicates the number of responses rather than the total number of students enrolled.

* Highest degree offered directly by the Physics Department.

**These courses include the 2 semesters in which we conducted the comparison study described in Sec. V A. Only data from students who took the CMR version is included in the total N.

† The post-test was taken online at this institution.

Institution Code	Institution Type	Highest Degree*	Size (undergraduates)	Number of courses	N
R1-A	Public	Ph.D.	25,000	4**	145
R1-C	Public	Ph.D.	37,000	1	40
R1-D	Public	Ph.D.	40,000	1	30
R1-E	Public	Ph.D.	29,000	1†	67
R2-A	Public	Ph.D.	19,000	2	33
BG-B	Private	B.S.	4,000	2	23
BG-C	Private	B.A.	3,000	1	8
BG-E	Private	B.S.	2,000	1	8
BG-F	Private	B.S.	3,000	1	19

of-class, online implementations is still an open question that we will not attempt to robustly answer here. However, pretest data from CU show an average score of $31.0 \pm 1.7\%$ when the pretest was taken online (N=78) compared to $30.5 \pm 2.2\%$ when taken in-class (N=56). This indicates that, for the pretest, in-class and online implementations are likely comparable. The post-test, however, is a considerably longer and harder instrument, and it may be that scores on a 50 min assessment administered online and in-class are not directly comparable. However, for the single course where the post-test was given online, the average score and standard deviation were consistent with that from in-class implementations. Moreover, the inclusion of this course does not significantly change any of the statistics or conclusions reported in the rest of this section. As such, we have opted to include these data in the following analysis in order to realize greater statistical power.

2. Results: Statistical Validation

This section presents the statistical validation of the CMR CUE independent of the FR version. As before, our analysis will be guided by Classical Test Theory (see Sec. IV A 3).

Using the nuanced R2 grading rubric (Sec. IV B 2), the overall average on the CMR CUE is $51.4 \pm 1\%$ when treating students as data points. The distribution of $N=373$ scores are shown in Fig. 24 and can be treated as normal (Anderson-Darling test [118], $p = 0.2$). Averaging by students differentially weights impact of large courses, which, in these data, come exclusively from large research institutions (Table IX). This effect can be offset by considering performance by course, rather than students. Taking the mean of the average scores for each course, the overall performance on the CMR CUE is $49.5 \pm 2.5\%$. With only $N=14$ courses, the difference between the by course and by student averages is not statistically significant; however, we argue that a difference of 2% is also not of practical significance.

The overall average (by students) falls to $48.0 \pm 1\%$ when using the simpler R1 scoring rubric (Sec. IV B 2). Using the ‘perfect’ scoring rubric, this average drops again to $41.0 \pm 0.9\%$. This is consistent with the findings from the comparison study (Sec. V A 2). Additionally, the internal consistency students’ scores as measured by Cronbach’s Alpha [112] fall to 0.78 and

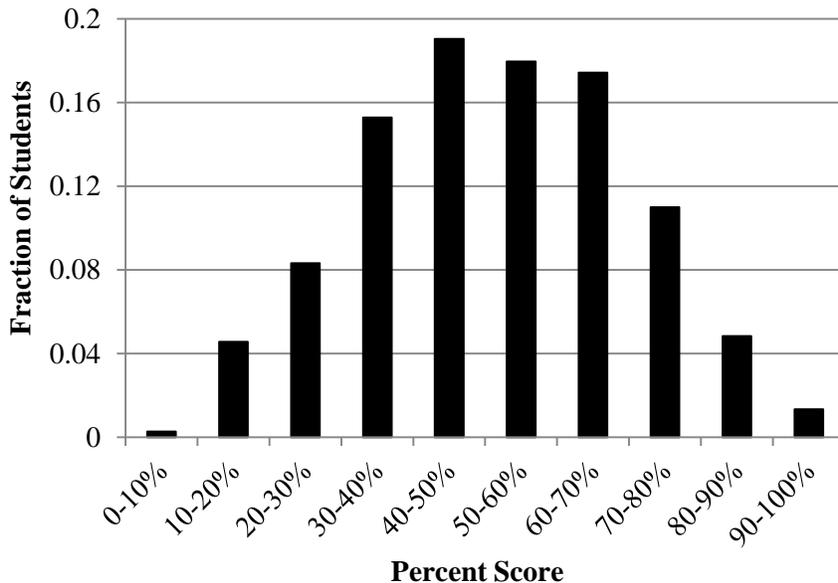


FIG. 24. Distributions of scores ($N=373$) on the CMR CUE from 14 courses at the institutions described in Table IX. These data passed a statistical test for normality (Anderson-Darling test, $p = 0.2$).

0.77 using the R1 and perfect rubric respectively. Given the increased internal consistency and overall consistency with the FR version offered by the more nuanced R2 grading scheme (Sec. V A 2), we will exclusively utilize the R2 rubric for the remainder of this section.

Criterion Validity

To establish the extent to which scores on the CMR CUE are consistent with other, related learning outcomes, we would ideally correlate these scores with final course grades and/or aggregate exam scores for all students in our sample. Unfortunately, we only have access to final course and exam scores for a subset of the students at CU and for none of the external institutions. Sec. V A 2 reported a high correlation of CMR CUE scores with aggregate exam score (Pearson's Correlation Coefficient [111] $r = 0.79$) as well as final course score ($r = 0.76$). This finding establishes the criterion validity of the CUE for the student population at CU; however, we are not able to extend this conclusion to external institutions with the available data.

Item Difficulty

To characterize the difficulty of each item, we look at the average score by question (Fig. 25). Item difficulties for all questions fall between 30-75%. We are not aware of a well-established range of acceptable values for item difficulty on polytomously scored items. However, for dichotomously scored items where item difficulty is measured as the percent of students who answer each item correctly [26], it is typically argued that ideal values should fall half-way between 100% and the percent expected by random guessing [127]. This maximizes the potential discriminatory power of each item. Since not all items will achieve this ideal, one standard range for acceptable values is 30-90% [26]. Extending this same logic of maximizing the potential discriminatory power of each item as well as the test as a whole, we argue item difficulties for our polytomously scored items fall within an acceptable range, with no single item being too easy or too hard to contribute to the overall discrimination of the test.

Scores on each individual item are rarely normally distributed. This is in part an artifact

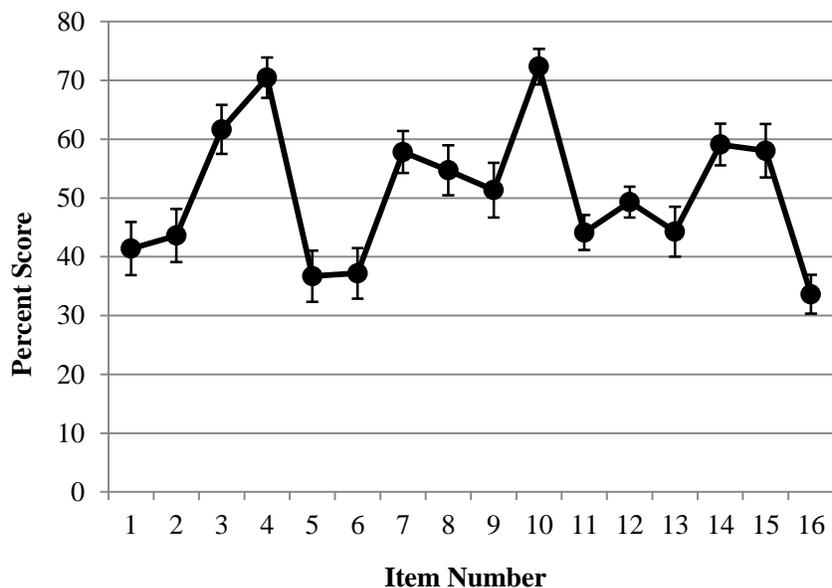


FIG. 25. Average scores for each item on the CMR CUE (N=373). Error bars represent a 95% confidence interval (double the standard error on the mean). Score distributions for each individual item are not necessarily normally distributed.

of the grading scheme in which there are a finite number of potential point combinations (typically between 0-5 pts in 0.5-1 pt intervals). For this reason, the median score on each item is often different from the average score. In two cases (Q1 and Q5, see Appendix H), the median score for this population of students is 0pts. For Q1, this is because there is only one possible solution path and any student who does not select the correct method receives zero points on the question. For Q5, the low median score is instead explained by the presence of a tempting incorrect method that receives no credit but is selected by nearly half the students.

Discrimination

One preliminary indication of the whole-test discrimination of the CMR CUE comes from the overall spread in the distribution of students' scores (Fig. 24). These scores span nearly the full range of possible scores (from 0-100%) with a minimum score of 4.3% and a maximum score of 91.9%. Thus the students are well distributed across the range of possible scores. As another measure of the whole-test discrimination of the CMR CUE, we again use Ferguson's

Delta [105]. For the full population of students, Ferguson’s Delta is 0.99. Recall, Delta can take on values between [0,1] and anything above 0.9 indicates good discriminator power [26].

We also examine the discrimination of each individual item by comparing a student’s score on that item to their performance on the rest of the test. Item-test correlations for all 16 items are shown in Fig. 26, and all correlation coefficients fall between 0.25-0.55 and are statistically significant given $N=373$ [128]. As has been done before [27], we adopt the standard cutoff of $r = 0.2$ used for dichotomously scored items [26] to argue that all items on the CMR CUE demonstrate acceptable discriminatory power. Note that this includes Q15, which previously demonstrated a lower item-test correlation ($r = 0.17$, see Sec. V A 2). Ultimately, only three courses (including the one in the initial comparison study) demonstrated lower item-test correlation on this item; however all other courses showed significant positive correlations, resulting in overall positive discrimination for this item.

Consistency

Once again, we utilize Cronbach’s Alpha as a conservative measure of the internal consis-

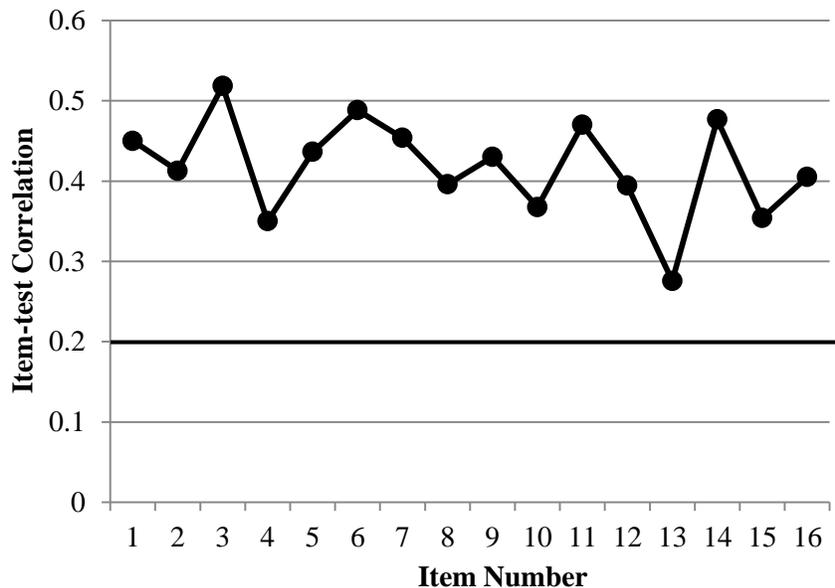


FIG. 26. Item-test correlations (as measured by Pearson’s r) for each of the items on the CMR CUE. For $N=373$ any correlation greater than 0.1 is significant at the $p < 0.05$ level [128], thus item-test correlations are statistically significant for all items. The conventional cutoff for an acceptable correlation (0.2) is marked as a bold line.

tency of the CMR CUE as a whole. For our population of students, we calculate a value of $\alpha = 0.81$. Recall that, for a test used to assess individuals rather than just groups, the commonly accepted threshold is $\alpha > 0.8$ [104]. Thus the CMR CUE demonstrates an acceptable level of internal consistency.

In terms of the new CMR format, there is another aspect of consistency that is important to consider. As the name implies, the majority of the questions on the *coupled* multiple-response CUE have several subparts whose scores and/or content are coupled, either explicit (as with the Method/Reasoning type items, Fig. 18) or implicitly (i.e., there is an opportunity for a student to be consistent or inconsistent in their responses to consecutive subparts). For example, the distribution of method selections for the item shown in Fig. 18 are given in Fig. 27(a). The two most common methods are Direct Integration and Multipole Expansion, both of which are possible solution paths in this case, though Multipole expansion is considerably easier. Fig. 27(b) breaks down the reasoning choices for students who selected each of these methods. There is a clear qualitative difference between the reasoning elements selected

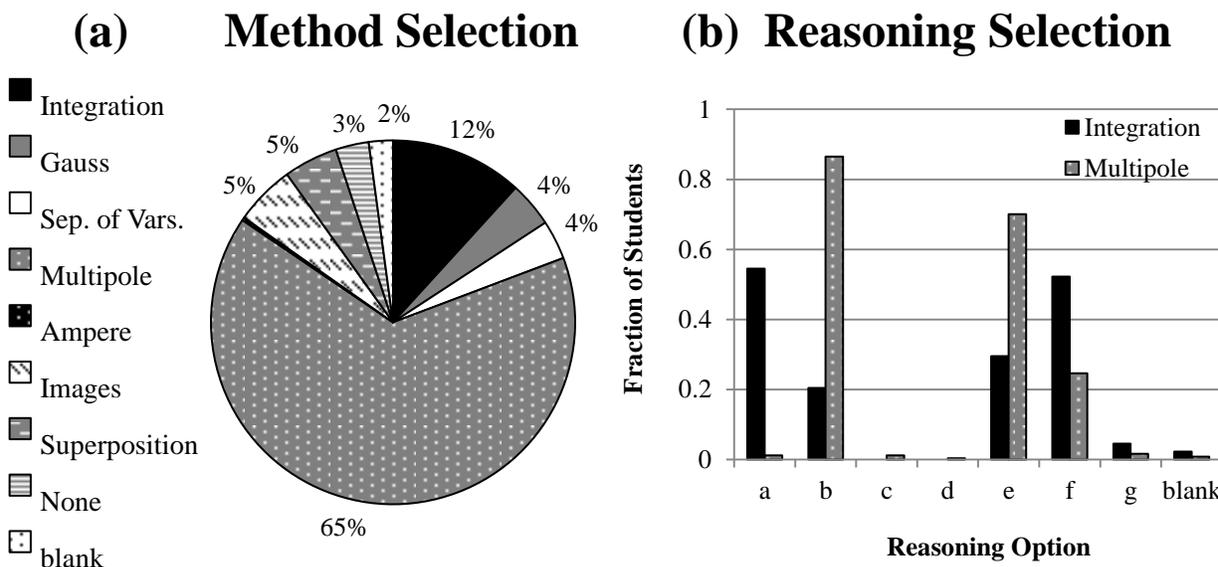


FIG. 27. (a) Method selections for N=373 students on Q3 (Fig. 18). (b) Reasoning selections for the subset of students who selected each of the two most common Methods: Direct Integration and Multipole Expansion (N=288). Each distribution is normalized with respect to the number of students who selected that method; however, students can select multiple reasoning elements so each distribution need not integrate to 1.

by these two sets of students. Students who chose Multipole were more likely to select reasoning elements ‘b’ and ‘e’, which represent the two elements required to fully justify Multipole expansion as the easiest method. Alternatively, students who selected Integration were more likely to select reasoning elements ‘a’ and ‘f’. Both of these elements are the commonly expressed justifications for using Direct Integration to solve this problem.

While Fig. 27 qualitatively suggests a certain degree of consistency between students’ method and reasoning selections, we also wanted to get a more quantitative sense of students’ consistency. To do this, we assigned a consistency code to students’ response to each question (excluding Q8, Q11, Q14, & Q15 which have no consistency check). Students were coded as ‘consistent’ if they selected at least one of the reasoning elements that supported their specific choice of method/answer and no inconsistent elements. Alternatively, if they selected any reasoning elements that were directly inconsistent with their choice of method, they were coded as ‘inconsistent’ regardless of whether they also selected some consistent reasoning elements. The remaining subset of students were coded as ‘neither’, meaning they left one of the two parts blank, chose the ‘None of These’ method option, or selected only reasoning elements that were neither directly consistent nor inconsistent with their choice of method. For example, on Q3 (Fig. 18), the combinations $(Method, Reasoning)=(B,d)$ or (A,af) would both be coded as ‘consistent’, whereas the combinations (B,bd) or (A,bf) , would be coded as ‘inconsistent, and the combinations (B,e) or (A,f) would be coded as ‘neither.’

The breakdown of the fraction of students receiving each consistency code is given in Fig. 28. On all questions but one, the fraction of consistent students is ≥ 0.5 , and the fraction of inconsistent student is ≤ 0.3 . Consistency between Q12 subparts iii and iv is noticeably lower than on other questions. These two subparts ask for qualitative graphs of E_z and V from a finite disk of charge and, for any given response to subpart iii, there is, at most, one consistent response to subpart iv. The relatively small number of potential consistent response patterns and the fact that consistency between these subparts is not explicit in the problem statement both contribute to the greater degree of inconsistency on this question. For all questions,

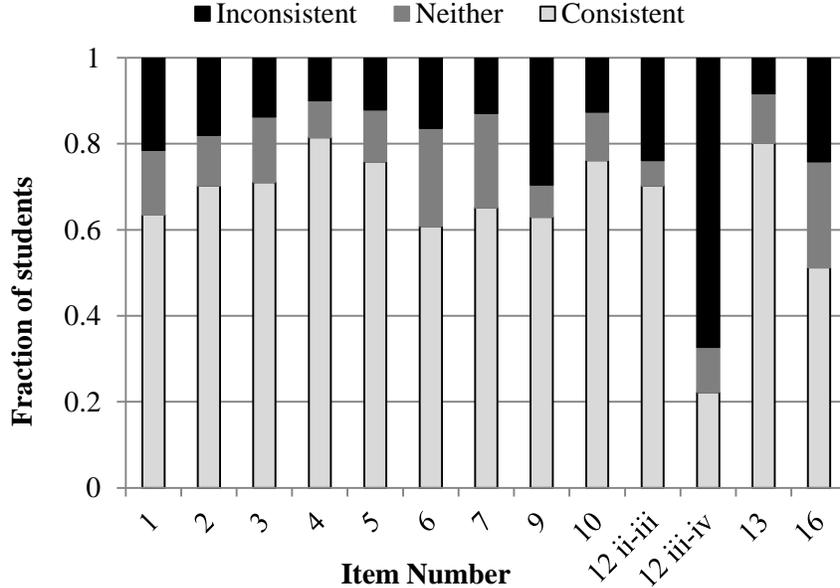


FIG. 28. Fraction of students coded as ‘consistent’ (at least one consistent reasoning element and no inconsistent ones), ‘inconsistent’ (any inconsistent reasoning elements), or ‘neither’ (neither consistent nor inconsistent reasoning elements) on each of the questions that have consistency checks. Note, there are no consistency checks on Q’s 8, 11, 14 & 15, but there are two possible consistency checks in Q12: one between subparts ii and iii, and one between subparts iii and iv.

consistent responses do not come exclusively from correct responses. In other words, many students are consistent even when they are incorrect. We take this as an indication that the majority of students are connecting their answers and reasoning selections in reasonable and meaningful ways rather than randomly guessing. This finding further supports the overall validity of the multiple-response format.

C. Accessing Student Difficulties with the CMR CUE

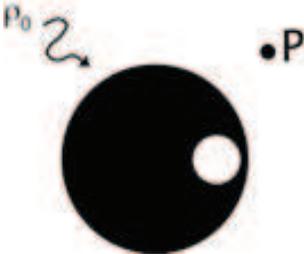
The previous sections have established the validity and reliability of the CMR CUE according to Classical Test Theory. However, in addition to providing a quantitative measure of student outcomes, the CUE also presents an opportunity to gain insight into student difficulties. For example, we utilized student responses to several CMR CUE questions in our investigations of student difficulties with the Dirac delta-function (Sec. III C) and separation of variables (Sec. III D). In this section, we focus on one additional question (Q5)

as an example of using the CMR CUE to think about student difficulties. The distributions of student responses to the remaining questions are given in Appendix K but will not be discussed in further detail here.

Q5 (Fig. 29) presents students with a solid sphere with an off-center, spherical cavity carved out of it and asks for the easiest method to find \vec{E} or V outside the sphere. The correct response is Superposition ('G') because you can treat this situation as two oppositely charged spheres ('e') and superpose the electric fields ('g') from each uniform sphere ('c') individually to determine the total electric field at point P. It is also possible, though *much* more difficult, to solve this problem through Direct Integration ('A') via Coulomb's Law ('a'). The distribution of method selections from this population of students is shown in Fig. 30(a). Almost half of the students (41%, N=156 of 373) correctly selected Superposition as

Q5 - A charged, insulating solid sphere of radius R with a uniform volume charge density ρ_o , with an off-center spherical cavity of radius r carved out of it (see Figure). Find \vec{E} or V at point P, a distance $4R$ from the sphere. Select only one: **The easiest method would be ...**

A. Direct Integration
 B. Gauss's Law
 C. Separation of Variables
 D. Multipole Expansion
 E. Ampere's Law
 F. Method of Images
 G. Superposition
 H. None of these



because ... (select **ALL** that support your choice of method)

a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
 b. the cube will look like a dipole; approximate with \vec{E} or V for an ideal dipole
 c. \vec{E} or V outside a uniform sphere is the same as from a point charge at the center
 d. the location of the cavity doesn't matter, you just need $Q_{enclosed}$ to calculate \vec{E}
 e. you can treat this as two uniform spheres, one with charge density ρ_o and one with charge density $-\rho_o$
 f. this will be the same as a uniform sphere with total charge $\frac{4}{3}\pi(R^3 - r^3)\rho_o$
 g. electric fields from multiple sources can be combined through a vector sum
 h. $\nabla^2 V = 0$ outside the cube and you can solve for V using Fourier Series

FIG. 29. Q5 from the CMR CUE. See Appendix I for the rubric and scoring materials.

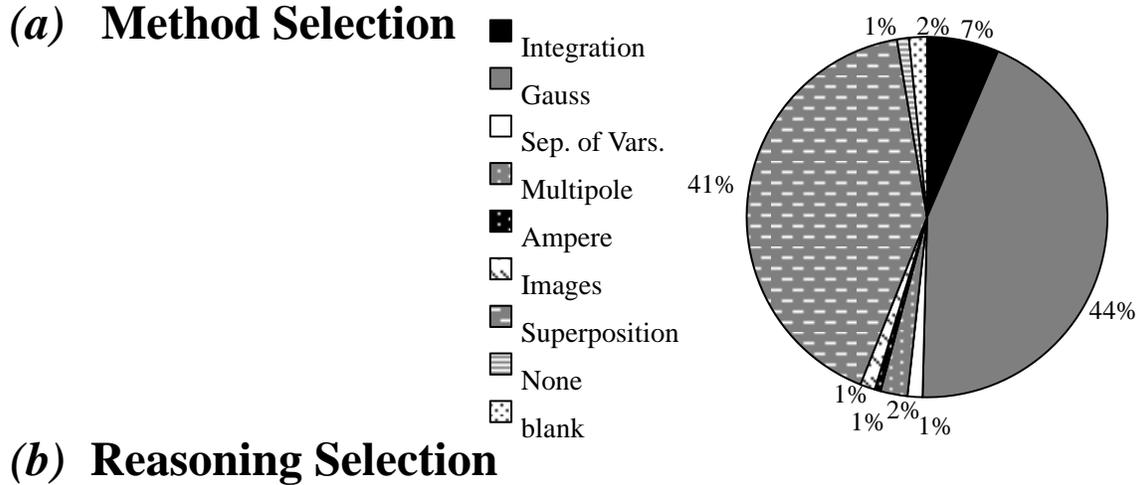


FIG. 30. (a) Method selections for $N=373$ students on Q5 (Fig. 29). (b) Reasoning selections for the subset of students who selected each of the three most common Methods: Direct Integration ($N=24$), Gauss’s Law ($N=164$), and Superposition ($N=154$). The relative number of student selecting each reasoning element is represented by the area of the circle.

the easiest method, and only a small number (7%, $N=24$ of 373) selected the more difficult method, Integration. Of the remaining students, the overwhelming majority (44%, $N=164$ of 373) selected Gauss’s Law.

There are at least two possible reasoning paths that could lead a student to select Gauss’s law as the method for this question. First, they are imagining using a single large Gaussian sphere centered on the origin of the solid sphere (not the cavity) to calculate \vec{E} from $Q_{enclosed}$ (consistent with reasoning elements ‘d’ and ‘f’). Alternatively, they could be imagining using two Gaussian spheres, one centered on a solid, uniform sphere and one on a solid, negatively

charged sphere in place of the cavity (consistent with reasoning elements ‘e’ and ‘g’). The latter strategy is correct while the former is fundamentally incorrect.

To distinguish between these two lines of reasoning, we must examine the reasoning selections of those students who selected Gauss’s law (Fig. 30(b)). The two most common reasoning elements selected by these students are ‘d’ and ‘f’, which supports the conclusion that the majority of these students were following the first (incorrect) line of reasoning. Indeed, of the students who selected Gauss’s law, only a tenth (10%, N=16 of 164) did not select one or both of reasoning elements ‘d’ or ‘f’. Only 6 of the remaining students selected both reasoning elements ‘e’ and ‘g,’ suggesting that they were using the second (correct) line of reasoning. This finding is consistent with previous research [18] and our own findings (Sec. III B 4) that suggest students often misapply Gauss’s law. In this case, the majority of students have argued that the location of the cavity does not matter, suggesting that either they have not recognized that the asymmetrical location of the cavity breaks the symmetry of the electric field, or that they have not recognized the asymmetry of the electric field eliminates Gauss’s law as a potential solution method. However, it is not possible to decide which of these two issues is at play for a particular student given only their response to this question.

As Superposition is the correct response to this question, it is tempting to assume that any student selecting method ‘G’ understands the correct solution method. This conclusion is generally supported by the observation that the most common reasoning elements selected by these students are ‘e’ and ‘g.’ However, more than a tenth of students who selected Superposition (16%, N=25 of 154) also selected one or both of reasoning elements ‘d’ and ‘f’, suggesting that these students were thinking only about superposition of charges (i.e., $Q_{large} - Q_{small}$) rather than fields (i.e., $\vec{E}_{large} - \vec{E}_{small}$). This distinction between superposition of charges rather than fields was also observed in previous research examining student responses to the free-response version of the CUE [129]. Both this result and the finding that a small number of students (N=6) selected Gauss’s Law along with reasoning

elements that suggest a correct strategy, underscore the importance of asking students to express their reasoning to avoid misinterpreting student responses.

D. Summary and Discussion

We have created a multiple-response version of an existing upper-division conceptual assessment, the CUE. Using student responses to the original free-response version of the instrument along with existing research on student difficulties, we crafted multiple-response options that reflected elements of common student ideas. This new version utilizes a novel approach to multiple-choice questions that allows students to select multiple reasoning elements in order to construct a complete justification for their answers. By awarding points based on the accuracy and consistency of students' selections, this assessment has the potential to produce scores that represent a more fine-grained measure of students' understanding of electrostatics than a standard multiple-choice test.

Direct comparison of the multiple-response and free-response versions of the CUE from two large, upper-division electrostatics courses yielded the same average score when using a nuanced grading scheme on both. The two versions also showed a high degree of consistency on multiple measures of test validity and reliability. Student interviews and expert feedback were used to establish the content validity of the CMR CUE for this comparison. Given the agreement between scores on the two versions and the ease of grading afforded by this new format, the CMR CUE is a considerably more viable tool for large-scale implementation. Additionally, while the FR version elicits greater variation in student reasoning, nearly three-quarters of students' responses to the FR version contain one or more elements that match the reasoning elements provided on the CMR version.

Moreover, we collected scores on the CMR CUE from multiple courses at multiple institutions. These data supported the validity and reliability of the instrument as measured by various test statistics including item difficulty, item discrimination, and internal consistency. We also examined the consistency of students' responses on consecutive subparts of individ-

ual questions. These data showed that the majority of students selected responses that were internally consistent even when the overall response was incorrect. Additionally, as an example of using the CMR CUE to gain insight into student difficulties, we demonstrated that student responses to one question support the findings from previous research that students tend to misapply Gauss's law in non-symmetric situations.

Our findings suggest that the CMR format can provide valid and easily-graded questions that produce scores that are consistent with scores from a FR format. We found this outcome surprising. When we began developing the CMR CUE we were skeptical that it would be possible to maintain more than a superficial level of consistency between the two versions. However, construction of the reasoning elements for the CMR CUE items relied heavily on the existence of data from student responses to the FR version. It is our opinion that, without this resource, our CMR CUE would not have been as successful at matching the FR CUE. An important limitation of the CMR format may be its reliance on pre-existing data from a FR version of the item or test. Additionally, as with the majority of conceptual assessments, the CMR CUE took several years to develop and validate even when building off the already established FR CUE. This time requirement places significant constraint on the development of similar assessments by instructors.

Another potential limitation of the CMR format comes from the relative complexity of the prompts. It is important that students read the directions fully for each question in order for an instructor to meaningfully interpret their response patterns. Interviews and overall scores from the two electrostatics courses discussed here suggest that our students followed directions and engaged with the question format as expected. However, these students were all junior and senior-level physics students taking a course required for their major. More experimentation is necessary to determine if the CMR format is viable for use with less experienced students who are not as invested in the content (e.g., introductory students or non-majors).

Additionally, not all questions easily lend themselves to the CMR format. For example,

one question on the CUE (Q9) was particularly challenging to translate into a CMR format. This item deals with determining the sign of the electric potential from a localized charge distribution given an arbitrary zero point. Students can leverage multiple, equally valid ways of determining the correct answer (e.g., by thinking about work, electric field, or shifting the potential). Capturing all of the correct, incorrect, and partially correct ideas expressed by students on the FR version of this question would have required a prohibitively large number of reasoning elements. To avoid this, we crafted a smaller number of reasoning elements with the goal of including each of the different types of justification (i.e., work, electric field, potential); however, we recognized that these elements did not encompass the variety of partially correct or incomplete ideas present in the FR version.

VI. CHAPTER VI: DISCUSSION & FUTURE WORK

In the preceding chapters, we described the development and implementation of two methodological tools designed to address our research questions and thus support efforts to increase student learning in upper-division electrostatics. Here, we briefly summarize each of these two tools and discuss future research directions associated with each.

In Chapter II, we presented an analytic framework – ACER – that is specifically targeted towards characterizing student difficulties with mathematics in upper-division physics. The ACER framework provides an organizing structure that focuses on important nodes in students’ solutions to complex problems by providing a researcher-guided outline that lays out the key elements of a well-articulated, complete solution. To account for the complex and highly context-dependent nature of problem solving in advanced undergraduate physics, ACER is designed to be operationalized for specific mathematical tools in different physics contexts rather than as a general description. In Chapter III, we utilized the operationalized ACER framework to inform and structure investigations of student difficulties with multivariable integration, the Dirac delta function, and separation of variables in upper-division electrostatics. This allowed us to more clearly identify prevalent difficulties our students demonstrated with each of these topics and to paint a more coherent picture of how these difficulties are interrelated.

The difficulties identified in this thesis represent only a subset of the difficulties that students may encounter when manipulating these specific mathematical tools. Future work could involve leveraging ACER to investigate students’ difficulties with specific mathematical tools in multiple contexts. For example, our investigation of multivariable integration focused exclusively on calculating the electric potential; however, the Coulomb’s law integral for the electric field uses many of the same ideas. Extending our investigation to target the electric field could potentially give us additional insight into student difficulties with multivariable integration and allow us to capture difficulties that arise when integrating vector quantities.

Similarly, multivariable integration appears again in the calculation of both the magnetic field and vector potential via the Biot-Savart law providing additional opportunities to explore the context-dependent aspects of student difficulties.

ACER could also be used to explore the evolution of students' difficulties over time and across topical areas. For example, Newton's law of gravity for extended bodies is mathematically very similar to the use of Coulomb's Law for continuous charge distributions but is typically encountered in sophomore physics. Similarly, both the Dirac delta function and separation of variables are important mathematical tools used in quantum mechanics (typically taken after completing electrostatics). Identifying students' difficulties with these mathematical tools in courses taken earlier or later in the undergraduate sequence and comparing them to the difficulties described here in junior electrostatics would allow us to investigate how these difficulties change (or not) as students advance through the curriculum.

Future work with the ACER framework might also include investigation of new mathematical tools and techniques such as vector calculus (div, grad, curl), line integrals, complex exponentials, etc. Investigation of new tools in new topical areas also provides opportunities to clarify the integration of representation in the framework. For example, students are often expected to work with multiple representations when dealing with complex exponentials in the context of electromagnetic waves. These representations include symbolic expressions, phasor diagrams, and vector diagrams in which the vectors represent polarizations of different fields. ACER could provide a useful framework to help determine which of these representations is most useful or difficult for students at different points in the problem solving process (i.e., Activation, Construction, Execution, Reflection).

The ACER framework was designed to be a tool not only for researchers but instructors as well. We have discussed a number of suggestions for instructors that may help students avoid or overcome the difficulties we identified. However, ACER can also be used to critique and design problems. Examining the prompt of a question can identify which components of the framework the problem targets and which ones it might short circuit (e.g., bypassing

Activation by instructing the student to use a Taylor series to approximate a function). This can help instructors to produce homework sets and exams that offer a balanced and complete assessment of all aspects of students' problem solving.

In Chapter IV, we presented the development and validation of another methodological tool: a novel, multiple response version of an existing free-response conceptual assessment targeting upper-division electrostatics (the CUE). This new CMR version of the CUE presents students with multiple response options of which they are allowed to select all that are relevant and awards points based on both the accuracy and consistency of their selections. In Chapter V, we demonstrated the validity and reliability of the CMR CUE using data from multiple institutions, and showed that student performance on the new version is statistically and, to a large degree, qualitatively comparable to student performance on the original free-response version.

Ongoing work with the CMR CUE includes continued data collection both at CU and other institutions in order to increase statistical power and broaden the student population for which the assessment has been validated. As we continue to aggregate additional data from multiple instructors and courses, it may also become possible to examine the impact of different curricular materials and/or pedagogical techniques on student performance on the CMR CUE. This information would help to guide instructors and researchers interested in improving student learning in electrostatics. The CMR CUE could also be used as a baseline assessment of graduate students both in terms of establishing their incoming preparation and measuring the impact of graduate E&M courses on student performance.

While it would be inappropriate to use the CMR CUE itself in any other context than upper-division electrostatics, the coupled multiple-response format could likely be adapted for use in other courses. Future work could include developing CMR versions of CU's other upper-division assessments which target, for example, electrodynamics or classical mechanics. Additionally, cross-disciplinary collaborations with education researchers from other scientific disciplines (e.g., biology or chemistry) could help to determine if the CMR for-

mat is appropriate for assessment within other fields. It would also be valuable to explore the viability of this format for the broader student populations including introductory and non-major students in physics courses. However, implementing the multiple-response format in larger introductory courses would require fully automating the scoring of the assessment using machine grading. Additional research would be necessary to ensure that the multiple-response format remained valid and reliable when student responses were collected on a multiple-choice bubble-sheet.

A significant part of the motivation for developing both the ACER framework and the CMR CUE was so that these tools might eventually support efforts to improve student learning through course transformations and materials development. Future work of this type might include developing and testing tutorial activities designed to specifically target student difficulties identified through ACER and/or the CUE. For example, a tutorial designed to guide students through the process of constructing the integral form of Coulomb's law by modeling a continuous charge distribution as a sum of many point charges could help overcome student difficulties with customizing both z and dq for specific charge distributions. Similarly, a separation of variables tutorial that experiments with different types of boundary conditions could help students to recognize both need for the infinite sum and which strategy for determining unknown constants is most efficient (i.e., Fourier's trick vs. term matching) depending on the boundary conditions. Moreover, both the CUE and ACER framework could also help to identify questions and content areas in which it could be particularly valuable to specifically target Activation and/or promote Reflection through modified homework problems, in-class activities, or concept tests.

Despite being influenced by CU's consensus and meta-level learning goals, both the ACER framework and the CUE diagnostic are still heavily content focused. Yet, there are many skills and characteristics related to a student's development as a physicist that extend beyond content knowledge and are rarely, if ever, assessed directly. For example, the capacity for independent learning, the ability to read and write scientific publications, and the ability

to work collaboratively are just a few characteristics of successful physicists that we ultimately want our physics majors to internalize. We argue that operationalizing and assessing these implicit goals represents an important outstanding issue for us and the broader PER community to consider. An open question is, can we begin to craft methodological tools that more accurately reflect and assess the full range of learning outcomes we value for our physics majors?

Both of the methodological tools described in this thesis were intentionally designed to be used either by education researchers to support research efforts investigating student learning or developing new curricular materials, or by traditional physics faculty and instructors to support less formal assessment and development within their classroom. In this way, we believe that these tools have the potential to support efforts to improve student learning over a wide range of scales, from individual classrooms all the way to broader transformation efforts that impact multiple courses and institutions.

-
- [1] R. J. Beichner. An introduction to physics education research. *Getting Started in PER*, 2009.
- [2] E. Mazur. *Peer Instruction: A User's Manual*. Series in Educational Innovation. Prentice Hall, Upper Saddle River, 1997.
- [3] L. C. McDermott and P. S. Shaffer. *Tutorials in introductory physics and homework package*. Prentice Hall, 2001.
- [4] J. Saul and E. F. Redish. Evaluation of the workshop physics dissemination project. Technical report, August 1998.
- [5] R. R. Hake. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.*, 66(1):64–74, 1998.
- [6] N. D. Finkelstein and S. J. Pollock. Replicating and understanding successful innovations: Implementing tutorials in introductory physics. *Phys. Rev. ST Phys. Educ. Res.*, 1:010101, Sep 2005.
- [7] S. Chasteen, K. Perkins, P. Beale, S. Pollock, and C. Wieman. A thoughtful approach to teaching: Course transformation for the rest of us. *Journal of College Science Teaching*, 40(4):70–76, March 2011.
- [8] S. J. Pollock. Longitudinal study of student conceptual understanding in electricity and magnetism. *Phys. Rev. ST Phys. Educ. Res.*, 5(2):020110, 2009.
- [9] S. V. Chasteen, S. J. Pollock, R. E. Pepper, and K. K. Perkins. Thinking like a physicist: A multi-semester case study of junior-level electricity and magnetism. *Am. J. Phys.*, 80:923, 2012.
- [10] S. V. Chasteen, S. J. Pollock, R. E. Pepper, and K. K. Perkins. Transforming the junior level: Outcomes from instruction and research in E&M. *Phys. Rev. ST Phys. Educ. Res.*, 8(2):020107, 2012.
- [11] C. Baily, M. Dubson, and S. Pollock. Research-based course materials and assessments for upper-division electrodynamics (E&M II). In *Physics Education Research Conference 2012*, volume 1513 of *PER Conference*, pages 54–57, Philadelphia, PA, August 1-2 2012.
- [12] S. Goldhaber, S. Pollock, M. Dubson, P. Beale, and K. Perkins. Transforming upper-division quantum mechanics: Learning goals and assessment. In *Physics Education Research Conference 2009*, volume 1179 of *PER Conference*, pages 145–148, Ann Arbor, Michigan, July 29-30 2009.
- [13] S. Pollock, R. Pepper, and A. Marino. Issues and progress in transforming a middle-division classical mechanics/math methods course. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference*, pages 303–306, Omaha, Nebraska, August 3-4 2011.
- [14] L. C. McDermott and E. F. Redish. Resource letter: PER-1: Physics education research. *American Journal of Physics*, 67(9), 1999.
- [15] J. L. Docktor and J. P. Mestre. Synthesis of discipline-based education research in physics. *Phys. Rev. ST Phys. Educ. Res.*, 10:020119, Sep 2014.
- [16] L. Hsu, E. Brewster, T. Foster, and K. Harper. Resource letter RPS-1: Research in problem solving. *Am. J. Phys.*, 72(9):1147–1156, September 2004.
- [17] D. E. Meltzer and R. K. Thornton. Resource letter ALIP-1: Active-learning instruction in physics. *Am. J. Phys.*, 80:478, 2012.

- [18] R. Pepper, S. Chasteen, S. Pollock, and K. Perkins. Our best juniors still struggle with gauss's law: Characterizing their difficulties. In *Physics Education Research Conference 2010*, volume 1289 of *PER Conference*, pages 245–248, Portland, Oregon, July 21–22 2010.
- [19] C. S. Wallace and S. V. Chasteen. Upper-division students' difficulties with ampère's law. *Phys. Rev. ST Phys. Educ. Res.*, 6:020115, Sep 2010.
- [20] C. Singh, M. Belloni, and W. Christian. Improving students' understanding of quantum mechanics. *Phys. Today*, 59(8):43–49, August 2006.
- [21] T. I. Smith, J. Thompson, and D. Mountcastle. Addressing student difficulties with statistical mechanics: The boltzmann factor. In *Physics Education Research Conference 2010*, volume 1289 of *PER Conference*, pages 305–308, Portland, Oregon, July 21–22 2010.
- [22] L. Deslauriers and C. Wieman. Learning and retention of quantum concepts with different teaching methods. *Phys. Rev. ST Phys. Educ. Res.*, 7(1):010101, January 2011.
- [23] M. Ayene, J. Kriek, and B. Dantie. Wave-particle duality and uncertainty principle: Phenomenographic categories of description of tertiary physics students' depictions. *Phys. Rev. ST Phys. Educ. Res.*, 7:020113, Nov 2011.
- [24] G. Zhu and C. Singh. Surveying students' understanding of quantum mechanics in one spatial dimension. *Am. J. Phys.*, 80(3):252–259, March 2012.
- [25] D. Hestenes, M. Wells, and G. Swackhamer. Force concept inventory. *The Physics Teacher*, 30(3):141–158, 1992.
- [26] L. Ding, R. Chabay, B. Sherwood, and R. Beichner. Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Phys. Rev. ST Phys. Educ. Res.*, 2:010105, Mar 2006.
- [27] S. V. Chasteen, R. E. Pepper, M. D. Caballero, S. J. Pollock, and K. K. Perkins. Colorado upper-division electrostatics diagnostic: A conceptual assessment for the junior level. *Phys. Rev. ST Phys. Educ. Res.*, 8(2):020108, 2012.
- [28] M. Caballero and S. Pollock. Assessing student learning in middle-division classical mechanics/math methods. In *Physics Education Research Conference 2013*, PER Conference, pages 81–84, Portland, OR, July 17–18 2013.
- [29] C. Henderson and M. H. Dancy. Impact of physics education research on the teaching of introductory quantitative physics in the united states. *Phys. Rev. ST Phys. Educ. Res.*, 5:020107, Dec 2009.
- [30] D. J. Griffiths. *Introduction to electrodynamics*. Prentice Hall, 1999.
- [31] <http://per.colorado.edu/Electrostatics>, 2015.
- [32] B. R. Wilcox, M. D. Caballero, D. A. Rehn, and S. J. Pollock. Analytic framework for students use of mathematics in upper-division physics. *Phys. Rev. ST Phys. Educ. Res.*, 9:020119, Nov 2013.
- [33] E. F. Redish. Problem solving and the use of math in physics courses. In *World View on Physics Education in 2005*, Focusing on Change, pages 1–10, Delhi, August 21–22 2005.
- [34] J. Wagner, C. Manogue, and J. Thompson. Representation issues: Using mathematics in upper-division physics. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference Invited Paper*, pages 89–92, Omaha, Nebraska, August 3–4 2011.
- [35] R. Pepper, S. Chasteen, S. Pollock, and K. Perkins. Facilitating faculty conversations: Development of consensus learning goals. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference*, pages 291–294, Omaha, Nebraska, August 3–4 2011.
- [36] M. D. Caballero, B. R. Wilcox, R. Pepper, and S. Pollock. ACER: A framework on the use

- of mathematics in upper-division physics. In *Physics Education Research Conference 2012*, volume 1513 of *PER Conference*, pages 90–93, Philadelphia, Pennsylvania, August 1-2 2012.
- [37] B. Wilcox, M. Caballero, R. Pepper, and S. Pollock. Upper-division student understanding of coulomb’s law: Difficulties with continuous charge distributions. In *Physics Education Research Conference 2012*, volume 1513 of *PER Conference*, pages 418–421, Philadelphia, Pennsylvania, August 1-2 2012.
- [38] F. Reif, J. H. Larkin, and G. C. Brackett. Teaching general learning and problem-solving skills. *Am. J. Phys.*, 44(3):212–217, March 1976.
- [39] D. S. Wright and C. D. Williams. A wise strategy for introductory physics. *The Physics Teacher*, 24:020105, April 1986.
- [40] P. Heller, R. Keith, and S. Anderson. Teaching problem solving through cooperative grouping. Part 1: Group vs. individual problem solving. *Am. J. Phys.*, 60(7):627–636, July 1992.
- [41] J. Docktor. *Development and Validation of a Physics Problem-Solving Assessment Rubric*. Dissertation, University of Minnesota, September 2009.
- [42] J. Tuminaro. *A Cognitive Framework for Analyzing and Describing Introductory Students’ Use and Understanding of Mathematics in Physics*. Dissertation, University of Maryland, January 2004.
- [43] D. Hammer. Student resources for learning introductory physics. *Am. J. Phys.*, 68(S1):S52–S59, July 2000.
- [44] A. Collins and W. Ferguson. Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist*, 28(1):25–42, 1993.
- [45] D. Tannen. *Framing in Discourse*. Oxford University Press, 1993.
- [46] T. J. Bing. *An Epistemic Framing Analysis of Upper-Level Physics Students’ Use of Mathematics*. Dissertation, University of Maryland, August 2008.
- [47] R. Catrambone. Task analysis by problem solving (TAPS): Uncovering expert knowledge to develop high-quality instructional materials and training. In *2011 Learning and Technology Symposium*, Columbus, Georgia, June 2011.
- [48] R. Catrambone. The subgoal learning model: Creating better examples so that student can solve novel problems. *J. Exp. Psychol. General*, 127(4):355–376, December 1998.
- [49] B. L. Sherin. How students understand physics equations. *Cognition and instruction*, 19(4):479–541, 2001.
- [50] R. Catrambone. Generalizing solution procedures learned from examples. *J. Exp. Psychol. Learn*, 22(4):1020–1031, July 1996.
- [51] B. R. Wilcox and S. J. Pollock. Coupled multiple-response versus free-response conceptual assessment: An example from upper-division physics. *Phys. Rev. ST Phys. Educ. Res*, 10:020124, Oct 2014.
- [52] A. Orton. Students’ understanding of integration. *Educational Studies in Mathematics*, 14(1):1–18, 1983.
- [53] P. W. Thompson and J. Silverman. The concept of accumulation in calculus. *Making the connection: Research and teaching in undergraduate mathematics*, 73:43–52, 2008.
- [54] R. R. Bajracharya, T. M. Wemyss, and J. R. Thompson. Student interpretation of the signs of definite integrals using graphical representations. In *2011 Physics Education Research Conference*, volume 1413, pages 111–114. AIP Publishing, 2012.
- [55] D.-H. Nguyen and N. S. Rebello. Students’ understanding and application of the area under the curve concept in physics problems. *Phys. Rev. ST Phys. Educ. Res.*, 7:010112, Jun 2011.

- [56] M. Artigue, J. Menigaux, and L. Viennot. Some aspects of students' conceptions and difficulties about differentials. *European Journal of Physics*, 11(5):262, 1990.
- [57] D. Hu and N. S. Rebello. Understanding student use of differentials in physics integration problems. *Phys. Rev. ST Phys. Educ. Res.*, 9:020108, Jul 2013.
- [58] D. Hu and N. S. Rebello. Using conceptual blending to describe how students use mathematical integrals in physics. *Phys. Rev. ST Phys. Educ. Res.*, 9:020118, Nov 2013.
- [59] N. Khan, D. Hu, D.-H. Nguyen, and N. S. Rebello. Assessing students' ability to solve introductory physics problems using integrals in symbolic and graphical representations. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference*, pages 227–230, Omaha, Nebraska, August 3-4 2011.
- [60] R. E. Pepper, S. V. Chasteen, S. J. Pollock, and K. K. Perkins. Observations on student difficulties with mathematics in upper-division electricity and magnetism. *Phys. Rev. ST Phys. Educ. Res.*, 8:010111, Mar 2012.
- [61] E. C. Sayre and M. C. Wittmann. Plasticity of intermediate mechanics students' coordinate system choice. *Phys. Rev. ST Phys. Educ. Res.*, 4:020105, Nov 2008.
- [62] B. R. Wilcox and S. J. Pollock. Upper-division student difficulties with the dirac delta function. *Phys. Rev. ST Phys. Educ. Res.*, 11:010108, Mar 2015.
- [63] H. Sadaghiani. *Conceptual and Mathematical Barriers to Students Learning Quantum Mechanics*. Dissertation, The Ohio State University, August 2005.
- [64] A. Mason. *Reflection on Problem Solving in Introductory and Advanced Physics*. Ph.d. dissertation, University of Pittsburgh, July 2009.
- [65] G. Zhu. *Improving Students' Understanding of Quantum Mechanics*. Ph.d., University of Pittsburgh, May 2011.
- [66] C. Singh. Helping students learn quantum mechanics for quantum computing. In *Physics Education Research Conference 2006*, volume 883 of *PER Conference Invited Paper*, pages 42–45, Syracuse, New York, July 26-27 2006.
- [67] M. J. Lighthill. *An introduction to Fourier analysis and generalized functions*. Cambridge University Press, 1958.
- [68] C. L. Rasmussen. New directions in differential equations: A framework for interpreting students' understandings and difficulties. *The Journal of Mathematical Behavior*, 20(1):55–87, 2001.
- [69] C. L. Rasmussen and K. D. King. Locating starting points in differential equations: A realistic mathematics education approach. *International Journal of Mathematical Education in Science and Technology*, 31(2):161–172, 2000.
- [70] S. Habre. Exploring students' strategies to solve ordinary differential equations in a reformed setting. *The Journal of Mathematical Behavior*, 18(4):455–472, 2000.
- [71] G. Duda and J. Ross. Problem-based learning in upper division courses: Student successes, perceptions, and reactions. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference*, pages 183–186, Omaha, Nebraska, August 3-4 2011.
- [72] R. K. Thornton and D. R. Sokoloff. Assessing student learning of newtons laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *Am. J. Phys.*, 66(4), 1998.
- [73] <http://www.ncsu.edu/per/TestInfo.html>, 2014.
- [74] B. R. Wilcox, M. D. Caballero, C. Baily, H. Sadaghiani, S. V. Chasteen, Q. X. Ryan, and S. J. Pollock. Development and uses of upper-division conceptual assessment. *Phys. Rev. ST*

- Phys. Educ. Res.*, Focused Collection on PER in Upper-division, In press.
- [75] B. M. Notaros. Concept inventory assessment instruments for electromagnetics education. In *Antennas and Propagation Society International Symposium, 2002. IEEE*, volume 1, pages 684–687. IEEE, 2002.
- [76] E. Cataloglu and R. W. Robinett. Testing the development of student conceptual and visualization understanding in quantum mechanics through the undergraduate career. *American Journal of Physics*, 70(3):238–251, 2002.
- [77] K. C. Midkiff, T. A. Litzinger, and D. L. Evans. Development of engineering thermodynamics concept inventory instruments. In *Frontiers in Education Conference, 2001. 31st Annual*, volume 2, pages F2A–F23. IEEE, 2001.
- [78] R. A Streveler, R. L. Miller, A. I. Santiago-Román, M. A. Nelson, M. R. Geist, and B. M. Olds. Rigorous methodology for concept inventory development: Using the ‘assessment triangle’ to develop and test the thermal and transport science concept inventory (ttci). *International Journal of Engineering Education*, 27(5):968, 2011.
- [79] D. L. Evans, G. L. Gray, S. Krause, J. Martin, C. Midkiff, B. M. Notaros, M. Pavelich, D. Rancour, T. Reed-Rhoads, P. Steif, et al. Progress on concept inventory assessment tools. In *Frontiers in Education, 2003. FIE 2003 33rd Annual*, volume 1, pages T4G–1. IEEE, 2003.
- [80] T. R. Thoads and R. J. Roedel. The wave concept inventory—a cognitive instrument based on bloom’s taxonomy. In *Frontiers in Education Conference, 1999. FIE’99. 29th Annual*, volume 3, pages 13C1–14. IEEE, 1999.
- [81] C. Singh. Student understanding of symmetry and gauss’s law of electricity. *Am. J. Phys.*, 74(10), 2006.
- [82] S. B. McKagan, K. K. Perkins, and C. E. Wieman. Design and validation of the quantum mechanics conceptual survey. *Phys. Rev. ST Phys. Educ. Res.*, 6:020121, Nov 2010.
- [83] <http://per.colorado.edu/sei>, 2014.
- [84] S. Chasteen and S. Pollock. A research-based approach to assessing student learning issues in upper-division electricity & magnetism. In *Physics Education Research Conference 2009*, volume 1179 of *PER Conference Invited Paper*, pages 7–10, Ann Arbor, Michigan, July 29–30 2009.
- [85] S.-Y. Lin and C. Singh. Can free-response questions be approximated by multiple-choice equivalents? *Am. J. Phys.*, 81(8), 2013.
- [86] A. Bork. Letter to the editor. *Am. J. Phys.*, 52(10), 1984.
- [87] T. R. Sandin. On not choosing multiple choice. *Am. J. Phys.*, 53(4), 1985.
- [88] B. L. Scott. A defense of multiple choice tests. *Am. J. Phys.*, 53(11), 1985.
- [89] M. Iona. In self-defense and defense of multiple choice tests. *Am. J. Phys.*, 54(4), 1986.
- [90] M. Iona. Multiple choice questions. *Am. J. Phys.*, 52(5), 1984.
- [91] R. N. Varney. More remarks on multiple choice questions. *Am. J. Phys.*, 52(12), 1984.
- [92] R. J. Dufresne, W. J. Leonard, and W. J. Gerace. Making sense of students answers to multiple-choice questions. *The Physics Teacher*, 40(3), 2002.
- [93] G. J. Aubrecht and J. D. Aubrecht. Constructing objective tests. *Am. J. Phys.*, 51(7), 1983.
- [94] H. T. Hudson and C. K. Hudson. Suggestions on the construction of multiple-choice tests. *Am. J. Phys.*, 49(9), 1981.
- [95] S. J. Sobolewski. Development of multiple-choice test items. *The Physics Teacher*, 34(2), 1996.
- [96] M. C. Rodriguez. Three options are optimal for multiple-choice items: A meta-analysis of 80

- years of research. *Educational Measurement: Issues and Practice*, 24(2):3–13, 2005.
- [97] J. Higbie. Multiple-choice exams: How well do we know the result? *The Physics Teacher*, 31(7), 1993.
- [98] H. Kruglak. Experimental study of multiple-choice and essay tests. I. *American Journal of Physics*, 33(12), 1965.
- [99] M. Scott, T. Stelzer, and G. Gladding. Evaluating multiple-choice exams in large introductory physics courses. *Phys. Rev. ST Phys. Educ. Res.*, 2:020102, Jul 2006.
- [100] R. N. Steinberg and M. S. Sabella. Performance on multiple-choice diagnostics and complementary exam problems. *The Physics Teacher*, 35(3), 1997.
- [101] T. R. Rhoads and R. J. Roedel. The wave concept inventory: A cognitive instrument based on bloom’s taxonomy. In *Frontiers in Education Conference, 1999. FIE’99. 29th Annual*, volume 3, pages 13C1–14. IEEE, 1999.
- [102] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. Taxonomy of educational objectives: Handbook I: Cognitive domain. *New York: D. McKay*, 19(56), 1956.
- [103] K. D. Kubinger and C. H. Gottschall. Item difficulty of multiple choice tests dependent on different item response formats—an experiment in fundamental research on psychological assessment. *Psychology science*, 49(4):361, 2007.
- [104] P. Engelhardt. An introduction to classical test theory as applied to conceptual multiple-choice tests. In *Getting Started in PER*, volume 2. April 2009.
- [105] L. Ding and R. Beichner. Approaches to data analysis of multiple-choice questions. *Phys. Rev. ST Phys. Educ. Res.*, 5(2):020103, September 2009.
- [106] L. Ding. Seeking missing pieces in science concept assessments: Reevaluating the brief electricity and magnetism assessment through rasch analysis. *Phys. Rev. ST Phys. Educ. Res.*, 10:010105, Feb 2014.
- [107] J. S. Aslanides and C. M. Savage. Relativity concept inventory: Development, analysis, and results. *Phys. Rev. ST Phys. Educ. Res.*, 9:010118, May 2013.
- [108] J. A. Marshall, E. A. Hagedorn, and J. O’Connor. Anatomy of a physics test: Validation of the physics items on the texas assessment of knowledge and skills. *Phys. Rev. ST Phys. Educ. Res.*, 5:010104, Mar 2009.
- [109] M. Planinic, L. Ivanjek, and A. Susac. Rasch model based analysis of the force concept inventory. *Phys. Rev. ST Phys. Educ. Res.*, 6:010103, Mar 2010.
- [110] P. Kline. *A handbook of test construction: Introduction to psychometric design*. Methuen, 1986.
- [111] J. Cohen. *Statistical power analysis for the behavioral sciences*. Routledge Academic, 2013.
- [112] J. M. Cortina. What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, 78(1):98, 1993.
- [113] C. S. Wallace and J. M. Bailey. Do concept inventories actually measure anything? *AstR.omy Education Review*, 9(1):010116, 2010.
- [114] L. Ding. Applying rasch theory to evaluate the construct validity of brief electricity and magnetism assessment. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference*, pages 175–178, Omaha, Nebraska, August 3-4 2011.
- [115] F. B. Baker. *The basics of item response theory*. ERIC, 2001.
- [116] R. J. De Ayala. *Theory and practice of item response theory*. Guilford Publications, 2009.
- [117] H. A. David and J. L. Gunnink. The paired t test under artificial pairing. *The American Statistician*, 51(1):9–12, 1997.

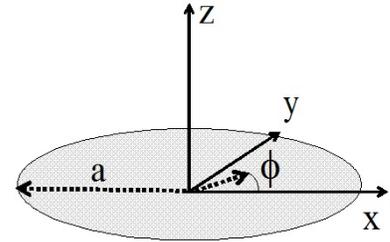
- [118] M. A Stephens. Edf statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737, 1974.
- [119] M. B. Brown and A. B. Forsythe. Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364–367, 1974.
- [120] R. A. Fisher et al. On the “probable error” of a coefficient of correlation deduced from a small sample. *Metron*, 1:3–32, 1921.
- [121] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.
- [122] J. M. Graham. Congeneric and (essentially) tau-equivalent estimates of score reliability: what they are and how to use them. *Ed. & Psych. Meas.*, 66:930, Dec 2006.
- [123] F. Yates. Contingency tables involving small numbers and the χ^2 test. *Supplement to the Journal of the Royal Statistical Society*, pages 217–235, 1934.
- [124] D. Howell. *Statistical methods for psychology*. Cengage Learning, 2012.
- [125] C. R. Mehta and N. R. Patel. A network algorithm for performing fisher’s exact test in $r \times c$ contingency tables. *Journal of the American Statistical Association*, 78(382):427–434, 1983.
- [126] M. Lombard, J. Snyder-Duch, and C. C. Bracken. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human communication research*, 28(4):587–604, 2002.
- [127] R. L. Doran. *Basic Measurement and Evaluation of Science Instruction*. ERIC, 1980.
- [128] B. L. Weathington, C. J. Cunningham, and D. J Pittenger. *Understanding business research*. J. Wiley & Sons, 2012.
- [129] J. Zwolak, M. B. Kustusch, and C. Manogue. Re-thinking the rubric for grading the cue: The superposition principle. In *Physics Education Research Conference 2013*, PER Conference, pages 385–388, Portland, OR, July 17-18 2013.

A. Exam Questions Targeting Multivariable Integration

Below are the three versions of the canonical Coulomb's Law question used on midterm exams in the E&M 1 course at CU during the four semesters from which data were collected.

Version 1

You have a flat insulating disk in the x - y plane: radius a , with a non-uniform charge density $\sigma = \sigma_o \cos(\phi/4)$ (where ϕ is the usual azimuthal angle in the x - y plane, as shown, and $\sigma_o > 0$.)



Please assume, as usual, that $V(\infty) = 0$

i) Describe briefly in words and/or pictures what this charge distribution “looks like”.

- What do you expect for the sign of V at the origin? (Physically, why?)

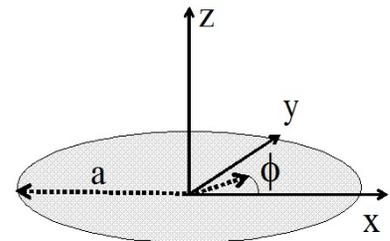
- Also, what do you expect for the limiting behavior of $V(z)$ along the z -axis, as z gets very large? (Why?) (Don't just say “it goes to zero” - HOW does it go to zero?)

(IMPORTANT: you will not need to explicitly check that your result, which you will obtain in the next part, behaves this way, to get full credit here!)

ii) Find the **voltage**, $V(0, 0, z)$, for points along the z -axis. (Express your answer in terms of given constants, namely z , a , and σ_o .)

Version 2 - Used twice

You have a flat insulating disk in the x - y plane: radius a , with a non-uniform charge density $\sigma = \sigma_o \cos(\phi/4)$ (where ϕ is the usual azimuthal angle in the x - y plane, as shown, and $\sigma_o > 0$.)



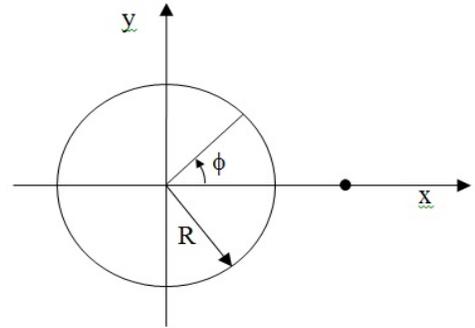
i) What is the total charge of this plate (in terms of σ_o and a ?)

ii) Find the voltage, $V(z)$, along the z -axis.

Version 3

A disk of radius R , in the xy -plane, centered on the origin, has surface charge density $\sigma = \sigma_o \cos^2(\phi)$.

A) Show with a sketch what this charge distribution looks like. Your sketch should show clearly where charges are located and whether the disk has a net charge.



B) Write down an explicit integral expression for the total charge Q of the disk. You do not have to perform the integral, but the integral must be in a form that is ready to be computed by Mathematica or a math student.

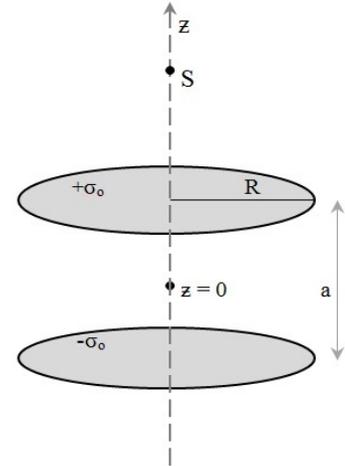
C) Write down an explicit integral expression for the voltage at a point on the x -axis. Assume $V = 0$ at infinity. Again, do not perform the integral, but the integral must be in a form that is ready to be computed by a Math student who knows no physics.

B. Interview Questions Targeting Multivariable Integration

Below are the two interview protocols targeting multivariable integration.

Interview Protocol 1 - Implemented prior to the development of the ACER framework.

Two thin circular disks of radius R are separated by a distance a . There is a uniform charge density $+\sigma_o$ on the top disk and $-\sigma_o$ on the bottom disk. The origin is between the two disks, as in the figure (so e.g. the top disk is at $z = +a/2$).



Equation Sheet - provided along with prompt

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_o} \iiint \frac{\rho(\vec{r}')}{z} d\tau'$$

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_o} \iiint \frac{\rho(\vec{r}')}{z^2} \hat{z} d\tau'$$

Interview Protocol 2 - Designed after the development of the ACER framework to target aspects of the framework not captured by the first set of interviews. Interviewees were asked this question after a set of 4 questions targeting the Dirac delta function (see Appendix D).

Q1. Calculate the electric potential on the z -axis from a charge distribution with volume charge density $\rho(\vec{r}) = \sigma \cos^2(\theta) \delta(r - R)$.

Q2. You have a friend taking E&M 1 who has asked for your help. Your friend tells you he is solving several separate problems asking for the potential from static, localized charge distributions with positive total charge Q . He shows you his answers (below) and several pages of work. You really don't want to look through all his work. Is there any way you can still help him? What questions might you want to ask?

1. $\frac{a}{2\epsilon_0} (\sqrt{r^2 + d^2} - r)$
2. $\frac{b}{4\pi\epsilon_0} \frac{d}{r^2}$
3. $\frac{c}{2\epsilon_0} \frac{d}{r}$

C. Exam Questions Targeting the Dirac Delta Function

Below are the four versions of delta-function questions used on midterm exams in the E&M 1 course at CU during the seven semesters of in which data were collected.

Version 1 - Used three times

Make a 3-D sketch the charge distribution: $\rho(x, y, z) = c\delta(x)\delta(z - 2)$

Describe it in words too. What are the units of c ?

Version 2 - Used twice

Make a 3-D sketch the charge distribution: $\rho(x, y, z) = a\delta(y - 2) + c\delta(x)\delta(y)\delta(z - 1)$

Describe it in words too. What are the units of a and c ?

Version 3

Consider the charge density $\rho(\vec{r}) = \sum_{i=1}^3 q_i \delta^3(\vec{r} - \vec{r}_i)$ where $q_1 = 1e$, $q_2 = 2e$, $q_3 = 3e$, and $\vec{r}_1 = (x = -1, y = 1, z = 0)$, $\vec{r}_2 = (1, 0, 0)$, and $\vec{r}_3 = (2, 2, 0)$.

A) Clearly describe this charge distribution with words and a sketch

B) What is the value of the volume integral $\int_{x>0} \rho d\tau$ (the integral is over the half-space $x > 0$)?

Version 4

A charge density is given as $\rho(x, y, z) = a\delta(x + 2)\delta(y - 3)$. What does this charge density physically represent?

- a) A point charge at $x = -2$, $y = 3$, and $z = 0$
- b) An infinitely long line or wire charge at $x = -2$ and $y = 3$
- c) An infinitely long line or wire charge at $x = 2$ and $y = -3$
- d) An infinitely large plane charge at $y = 3$
- e) An infinitely large plane charge at $x = 3$

D. Interview Questions Targeting the Dirac delta function

Below are the two interview protocols targeting the Dirac delta function.

Interview Protocol 1

Q1. What is the volume charge density for an infinitely long linear charge distribution running parallel to the z-axis and passing through the point (1,2,0). Be sure to define any new symbols you introduce.

Sketch this charge distribution.

Q2. Sketch the following charge distribution: $\rho(\vec{r}) = \alpha\delta(z - 2)$.

Describe the distribution in words.

What are the units of α ?

Q3. Sketch the following charge distribution: $\rho(\vec{r}) = \beta\delta(z)\delta(x + 1)$.

Describe the distribution in words.

What are the units of β ?

Q4. Make a 3-D sketch of the following charge distribution: $\rho(\vec{r}) = \gamma\delta(r - 2)$.

Describe the distribution in words.

What is the total charge?

Interview Protocol 2 -

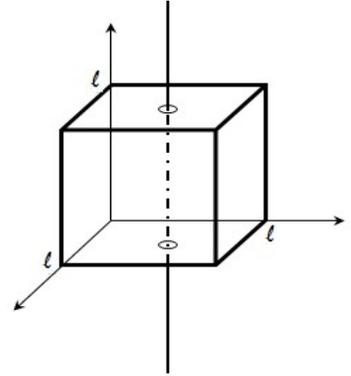
Q1. Using delta functions, provide a mathematical expression for the volume charge density, $\rho(\vec{r})$, of an infinite line of charge running parallel to the z-axis and passing through the point (1,2,0). Be sure to define any new symbols you introduce.

Q2. Sketch the following charge distribution: $\rho(\vec{r}) = \beta\delta(y - 2)\delta(x - 1)$.

Describe the distribution in words.

What are the units of β ?

Q3. Using your expression for the volume charge density, $\rho(\vec{r})$, calculate the total charge enclosed in the cube given below.



Q4. Calculate the following:

a) $\int_{-\infty}^{\infty} \delta(x) dx$

b) $\int_{-\infty}^{\infty} x \delta(x) dx$

c) $\int_0^{10} [a\delta(x-1) + b\delta(x+2) + c\delta(x-3)] dx$

d) $\iiint a\delta(r-r') r^2 \sin(\theta) dr d\phi d\theta$

E. Exam Questions Targeting Cartesian Separation of Variables

Below are the four versions of Cartesian SoV questions used on midterm and final exams in the E&M 1 course at CU during the 3 semesters in which data were collected.

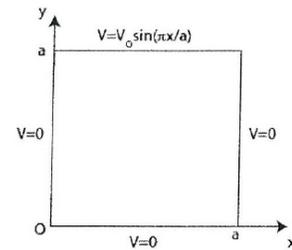
Version 1

A rectangular pipe, running parallel to the z -axis, extending from $-\infty$ to ∞ , has three grounded metal sides at $y = b$, $x = 0$, $y = a$, the fourth side at $y = 0$ is set to a constant potential V_0 .

Find the potential inside the pipe.

Version 2

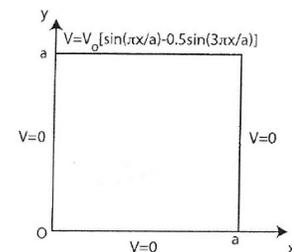
A square rectangular pipe (sides of length a) runs parallel to the z -axis (from $-\infty$ to ∞). The 4 sides are maintained with boundary conditions given in the figure. (Each of the 4 sides is insulated from the others at the corners).



Find the potential $V(x, y)$ at all points in this pipe. You need to write down the relevant equations, boundary conditions and solution procedures.

Version 3

A square rectangular pipe (sides of length a) runs parallel to the z -axis (from $-\infty$ to ∞). The 4 sides are maintained with boundary conditions given in the figure. (Each of the 4 sides is insulated from the others at the corners).

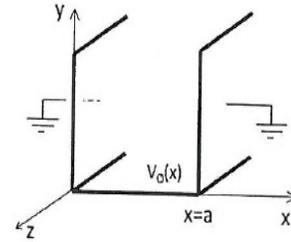


A) Write down the relevant differential equation and boundary conditions for $V(x, y)$.

B) Find the potential $V(x, y)$ at all points in this pipe. You need to write down the solution procedures.

Version 4

a) Consider a gutter which extends to infinity in the $\pm z$ -direction and to infinity in the $+y$ -direction, as shown in the figure. The two metal plates at $x = 0$ and $x = a$ are grounded. The base at $y = 0$ is maintained at a specific potential $V_o(x)$. Assume that the potential drops to zero at $y = \infty$.



Using separation of variables in Cartesian coordinates, the potential inside the gutter can be written as $V(x, y) = X(x)Y(y)$. Which of the following expression for $V(x, y)$ is correct for this set-up? Circle your answer, you do not need to explain your reasoning.

(i) $V(x, y) = [A_x \exp(kx) + B_x \exp(-kx)][A_y \sin(ky) + B_y \cos(ky)]$

(ii) $V(x, y) = [A_y \sin(kx) + B_y \cos(kx)][A_x \exp(ky) + B_x \exp(-ky)]$

where A_x, A_y, B_x, B_y, k are unknown constants.

b) Use your expression from part a) together with the three boundary conditions for the grounded plates at $x = 0$ and $x = a$ and $y = \infty$ and determine as many of the constants A_x, A_y, B_x, B_y, k as possible.

(Note: I do not ask you to use the fourth boundary conditions $V(x, y = 0) = V_o(x)$. Thus, you won't be able to solve for all of the constants.)

F. Exam Questions Targeting Spherical Separation of Variables

Below are the nine distinct versions of spherical SoV questions used on midterm and final exams in the E&M 1 course at CU during the 9 semesters in which data were collected.

Version 1

The potential on the surface of a sphere with radius R is set as $V = V_o P_9(\cos \theta)$

Find the potential inside/outside the sphere. Hint: keep all your answers in terms of Legendre polynomials.

Version 2

The potential on the surface of a sphere with radius R depends only on the (polar) angle θ and is given by $V_R(r = R, \theta)$. Inside and outside the sphere there is vacuum and we assume, as usual, that $V(r \rightarrow \infty) = 0$.

a) Starting from the expression for the general formula for the separation of variables in spherical coordinates:

$$V(r, \theta) = \sum_{l=0}^{\infty} (A_l r^l + \frac{B_l}{r^{l+1}}) P_l(\cos \theta)$$

obtain expressions for the unknowns A_l and B_l for the potential **outside** the sphere.

b) Now and for this part only, suppose that the potential at the surface has the special form $V_R(r = R, \theta) = V_o \cos^2(\theta)$, where V_o is a constant with appropriate units. Calculate the potential inside the sphere.

Version 3 - Used twice

A non-conducting spherical shell of radius R has a complicated potential on its surface $V(r = R, \theta) = 3\cos^2\theta + 1$ (in Volts). Assume $V(r \rightarrow \infty) = 0$, and that all charges reside only on the surface of the shell at $r = R$.

a) Find the potential $V(r, \theta)$ everywhere (inside and outside this sphere)

b) Why might you expect the leading term in the potential $V(r > R)$ to depend on $1/r$?

Version 4

A hollow spherical shell of radius R and negligible thickness is centered on the origin and has voltage $V = V_o \cos \theta$ on its surface. (θ is the usual spherical coordinate)

Solve for the voltage in the interior of the shell ($r < R$).

Version 5

A spherical shell (radius R) has a known voltage at its surface: $V(r = R, \theta) = V_o(1 + \cos \theta)$

There are no charges outside the shell and we assume as usual that $V(\infty) = 0$.

Find the voltage $V(r, \theta)$ for all points outside this shell. (Very briefly, show your work).

Version 6

Consider a non-conducting sphere of radius R centered on the origin. The voltage on its surface is fixed to be $V(r = R, \theta) = V_o \sin^2 \theta$. We can also assume $V(r \rightarrow \infty) = 0$

- a) Write the potential on the sphere in terms of Legendre polynomials. (given P_l)
- b) Solve for the voltage $V(r, \theta)$ everywhere outside the sphere: give your answer as an explicit function of r and θ . Indicate each BC used.

Version 7

A non-conducting sphere of radius R has a voltage on its surface given by $V = V_o(1/2 + 3/2 \cos^2 \theta)$.

- a) Solve for the voltage V everywhere outside the shell
- b) By considering the form of your solution V in the limit of $r \gg R$, solve for the total charge on the sphere and the dipole moment of the sphere. Explain your reasoning.

Version 8 - Used twice

A thin spherical shell of radius R has the potential $V = 3V_o/2 \cos^2 \theta$ on the surface. Find the voltage outside the sphere. If your answer requires an infinite number of terms, describe how the coefficients of each term would be determined; if the solution only has a finite number of terms, proceed to determine the coefficients.

Version 9

The potential on the surface of a sphere with radius R has the special form

$V_R(r = R, \theta) = V_o(\cos \theta + \cos^2 \theta)$. Inside and outside the sphere there is vacuum and we assume, as usual, that $V(r \rightarrow \infty) = 0$. Using separation of variables in spherical coordinates we can write the general solution for the potential in the form:

$$V(r, \theta) = \sum_{l=0}^{\infty} (A_l r^l + \frac{B_l}{r^{l+1}}) P_l(\cos \theta)$$

a) We want to determine the potential inside the sphere for $r < R$. For this case you can conclude that (circle your answer, you do not have to explain your reasoning)

- (i) All A_l 's have to be zero due to the boundary condition: $V(r = \infty) = 0$
- (ii) All A_l 's have to be zero due to the boundary condition: V is finite at $r = 0$
- (iii) All B_l 's have to be zero due to the boundary condition: $V(r = \infty) = 0$
- (iv) All B_l 's have to be zero due to the boundary condition: V is finite at $r = 0$

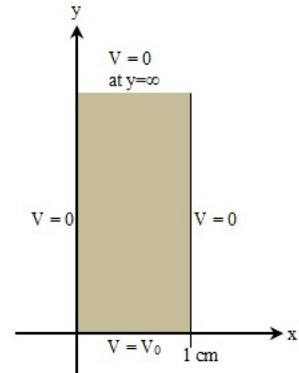
b) Based on your answer in part a) calculate all the remaining unknowns (i.e., either A_l 's or B_l 's) in terms of R and V_o .

G. Interview Questions Targeting Separation of Variables

Below are the two interview protocols targeting separation of variables in both spherical and Cartesian coordinates.

Interview Protocol 1

Q1. A semi-infinite plane with width 1cm is shown below. The bottom of the plate is held at some constant voltage V_o . The other three sides are held at $V = 0$. Assuming the temperature inside the plate obeys Laplace's equation, find an expression for the steady-state voltage inside the plate.



Interview Protocol 2

Q1. A hollow spherical shell of radius R and negligible thickness is centered at the origin and has a voltage on its surface, $V(R, \theta) = V_o(1 + \cos \theta)$. Solve for the voltage in all space

Q2. An infinite rectangular pipe running parallel to the z -axis has three grounded metal sides at $x = 0$, $y = b$, and $y = 0$. The fourth side $x = a$ is held at constant potential V_o . Find the potential inside the pipe.

H. The Coupled Multiple-Response CUE Instrument

This copy of the CMR CUE is provided for reference only.

Please visit per.colorado.edu/Electrostatics for the most recent version of the assessment along with all scoring materials.

Junior-level Electrostatics Content Review

Please fill out the following exam to the best of your ability. This will not count towards your final grade in the course. Do your best to get all the questions on the test.

When asked to provide your reasoning, **select the responses that *best* explain your choice.**

We are still interested in incorrect answers if they are accompanied by consistent reasoning.

If you have no idea how to answer a question, leave it blank. **Do not guess.**

Version 9

Name: _____

For questions 1-7, select the **EASIEST** method that you would use to solve the problem and the reason for your selection.

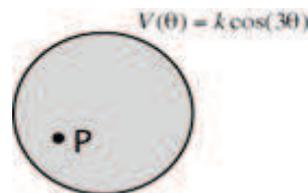
You may select as many of the ‘because ...’ options as you would like.

Q1 - A hollow, insulating spherical shell with radius R , with a voltage on its surface, $V(\theta) = k\cos(3\theta)$.

Find \vec{E} or V inside the sphere at point P.

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. symmetry allows you to calculate \vec{E} using a spherical Gaussian surface
- c. the boundary conditions is azimuthally symmetric (i.e., symmetric in ϕ)
- d. there is not appropriate symmetry to use other methods
- e. you can use $\vec{E}(\vec{r}) = -\vec{\nabla}(k\cos 3\theta)$ and evaluate this at point P
- f. $\nabla^2 V = 0$ inside the sphere and you can solve for V using Legendre Polynomials

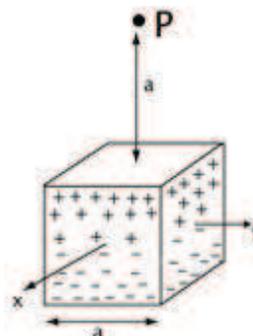
(Optional) If you would like to elaborate further, please do so in the box below

Q2 - A solid, neutral, non-conducting cube, centered on the origin, with side length a . It has a charge density that depends on the distance z from the origin, $\rho(z) = kz$, so that the top of the cube is strongly positive and the bottom is strongly negative, as in the figure.

Find \vec{E} or V at point P, **on the z-axis**, at a distance a from the cube.

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. the cube will look like a dipole; approximate with \vec{E} or V for an ideal dipole
- c. symmetry allows you to calculate \vec{E} using a cubical Gaussian surface
- d. symmetry allows you to calculate \vec{E} using a spherical Gaussian surface
- e. the observation point is **close** to the cube
- f. there is not appropriate symmetry to use other methods
- g. $\nabla^2 V = 0$ outside the cube and you can solve for V using Fourier Series

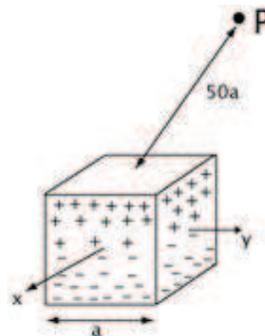
(Optional) If you would like to elaborate further, please do so in the box below

Q3 - The same neutral, non-conducting cube as above, with $\rho(z) = kz$.

Find \vec{E} or V at point P, where P is **off-axis**, at a distance **50a** from the cube.

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. the cube will look like a dipole; approximate with \vec{E} or V for an ideal dipole
- c. symmetry allows you to calculate \vec{E} using a cubical Gaussian surface
- d. symmetry allows you to calculate \vec{E} using a spherical Gaussian surface
- e. the observation point is **far** from the cube
- f. there is not appropriate symmetry to use other methods
- g. $\nabla^2 V = 0$ outside the cube and you can solve for V using Fourier Series

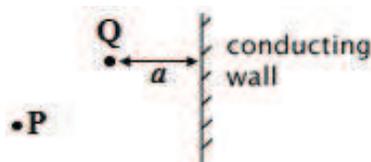
(Optional) If you would like to elaborate further, please do so in the box below

Q4 - A grounded conducting plane with a positive point charge Q at a distance a to the left of the plane.

Find \vec{E} or V at point P .

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. charges on the wall create the same field at P as a point charge a distance a behind the surface
- c. symmetry allows you to calculate \vec{E} using a Gaussian pillbox
- d. this method exploits the boundary condition
- e. there is not appropriate symmetry to use other methods
- f. $\nabla^2 V = 0$ on the left side of the wall and you can solve for V using Fourier Series

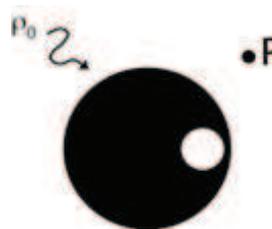
(Optional) If you would like to elaborate further, please do so in the box below

Q5 - A charged insulating solid sphere of radius R with a uniform volume charge density ρ_o , with an off-center spherical cavity of radius r carved out of it (see figure).

Find \vec{E} or V at point P, a distance $4R$ from the sphere.

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. the sphere will look like a dipole; approximate with \vec{E} or V for an ideal dipole
- c. \vec{E} or V outside a uniform sphere is the same as from a point charge at the center
- d. the location of the cavity doesn't matter, you just need $Q_{enclosed}$ to calculate \vec{E}
- e. you can treat this as two uniform spheres, one with charge density ρ_o and one with charge density $-\rho_o$
- f. this will be the same as a uniform sphere with total charge $4/3\pi(R^3 - r^3)\rho_o$
- g. electric fields from multiple sources can be combined through a vector sum
- h. $\nabla^2 V = 0$ inside the sphere and you can solve for V using Legendre Polynomials

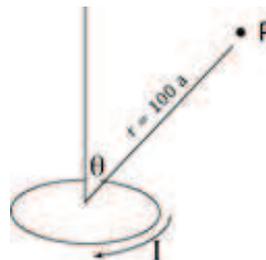
(Optional) If you would like to elaborate further, please do so in the box below

Q6 - A current loop of radius a that carries a constant current I .

Find \vec{B} at point P, where P is **off-axis**, at a distance $r=100a$.

Select only one: The easiest method would be ...

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{B} using the Biot-Savart Law
- b. the ring will look like a dipole; approximate with \vec{B} for an ideal dipole
- c. symmetry allows you to calculate \vec{B} using an Amperian loop centered on the origin
- d. symmetry allows you to calculate \vec{B} using an Amperian loop centered on the line of current
- e. the observation point is **far** from the ring
- f. there is not appropriate symmetry to use other methods
- g. it is straightforward to calculate \vec{A} and use $\vec{B} = \vec{\nabla} \times \vec{A} = 0$

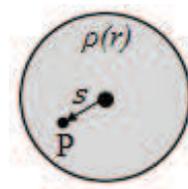
(Optional) If you would like to elaborate further, please do so in the box below

Q7 - A solid non-conducting sphere, centered on the origin, with a non-uniform charge density that depends on the distance from the origin, $\rho(r) = \rho_0 e^{-r^2/a^2}$ where a is a constant.

Find \vec{E} or V inside at point P a distance s from the center.

Select only one: **The easiest method would be ...**

- A. Direct Integration
- B. Gauss's Law
- C. Separation of Variables
- D. Multipole Expansion
- E. Ampere's Law
- F. Method of Images
- G. Superposition
- H. None of these



because ... (select **ALL** that support your choice of method)

- a. you can calculate \vec{E} or V using the integral form of Coulomb's Law
- b. $|\vec{E}|$ is constant at all points where $r = s$
- c. \vec{E} is the same as a point charge at the origin with total charge $Q = \frac{4}{3}\pi s^3 \rho(r)$
- d. there is a non-uniform charge distribution
- e. \vec{E} is perpendicular to a spherical surface of radius s centered on the origin
- f. there is not appropriate symmetry to use other methods
- g. you can use Legendre polynomials to express \vec{E} or V inside and evaluate this at the point P

(Optional) If you would like to elaborate further, please do so in the box below

Q8 - A mass density is given by $\rho(\vec{r}) = m_1\delta^3(\vec{r} - \vec{r}_1) + m_2\delta^3(\vec{r} - \vec{r}_2)$, where m_1 and m_2 are constants.

i. What is the value of $\int_{all\ space} \rho(\vec{r})d\tau$?

- a) $m_1\delta^3(\vec{r} - \vec{r}_1) + m_2\delta^3(\vec{r} - \vec{r}_2)$
- b) $m_1\delta^3(\vec{r}_1) + m_2\delta^3(\vec{r}_2)$
- c) $m_1 + m_2$
- d) $m_1\vec{r}_1 + m_2\vec{r}_2$
- e) $m_1r_1 + m_2r_2$
- f) $m_1r_1^2 + m_2r_2^2$

(Optional) If you would like to elaborate further, please do so in the box below

ii. What physical situation does this mass density represent?

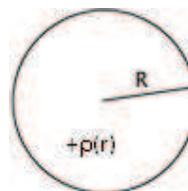
- a) Two solid spheres of radius r_1 and r_2
- b) Two concentric spherical shells of radius r_1 and r_2
- c) Two nested cylindrical shells with radius r_1 and r_2 (i.e., a coaxial cable)
- d) Two point masses located at \vec{r}_1 and \vec{r}_2
- e) Two spheres of non-zero radius located at \vec{r}_1 and \vec{r}_2 (i.e., a dumbbell)

(Optional) If you would like to elaborate further, please do so in the box below

Q9 - You are given a problem involving a non-conducting sphere centered at the origin. The sphere has a non-uniform, positive and finite volume charge density $\rho(r)$. You notice that a classmate has set the reference point for the voltage (V) such that $V=0$ at the center of the sphere $V(r=0)=0$.

What would $V=0$ at $r=0$ imply about the sign of the potential at $r \rightarrow \infty$?

- A. $V(r \rightarrow \infty)$ is positive (+)
- B. $V(r \rightarrow \infty)$ is negative (-)
- C. $V(r \rightarrow \infty)$ is zero
- D. It depends



Select ALL of the following statements which support your choice

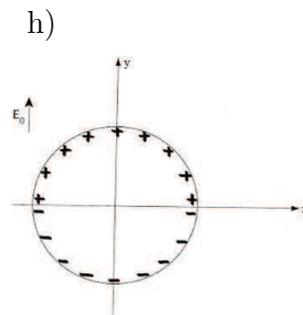
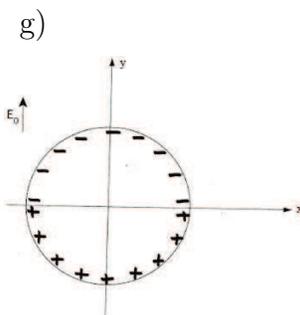
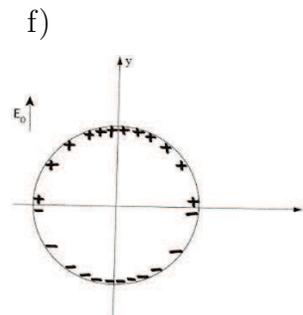
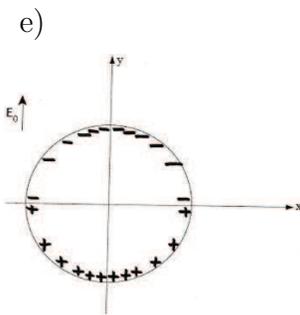
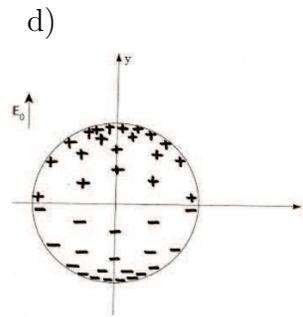
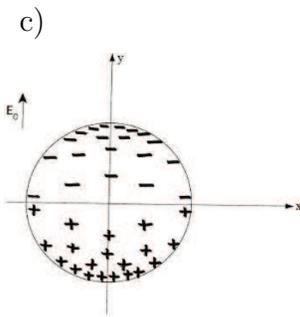
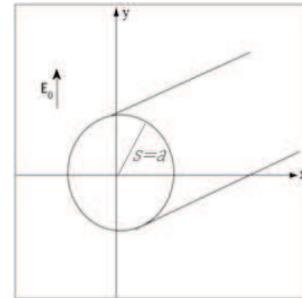
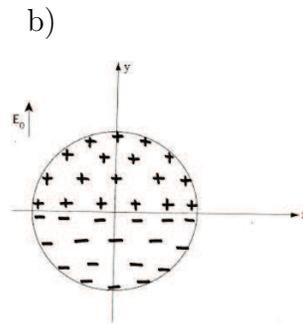
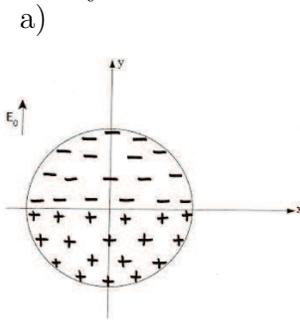
- a. It takes positive external work to bring a positive test charge in from infinity and $W = q\Delta V$
- b. It takes negative external work to bring a positive test charge in from infinity and $W = q\Delta V$
- c. V goes as $\frac{1}{r}$ for positive charges and must go to zero as $r \rightarrow \infty$
- d. V decreases as you follow \vec{E} field lines and for positive charges the field lines point outward
- e. V increases as you follow \vec{E} field lines and for positive charges the field lines point outward
- f. The potential is arbitrary up to the addition of a constant
- g. We need to know more about the functional dependence of $\rho(r)$

(Optional) If you would like to elaborate further, please do so in the box below

Q10 - You are given an infinite solid, **conducting** cylinder whose vertical axis runs along the z direction, that is placed in an external electric field, $E_0\hat{y}$, as in the figure to the right. The cylinder extends infinitely in the $+z$ and $-z$ directions.

i. Which of the following sketches best represents the resulting charge distribution?

Select only one.



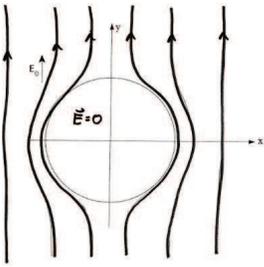
i) None of these
(please elaborate
in the space below)

Recall, we are considering a solid, conducting cylinder in the presence of a uniform external electric field, $\vec{E}_{external} = E_0 \hat{y}$.

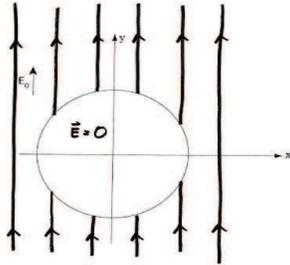
ii. Which of the following sketches best represents the electric field everywhere?

Select only one. ($\vec{E} = 0$ only where explicitly stated)

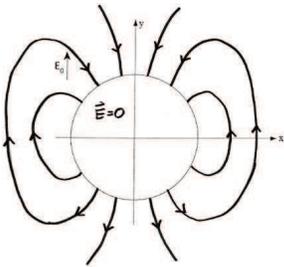
a)



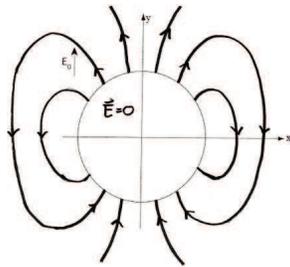
b)



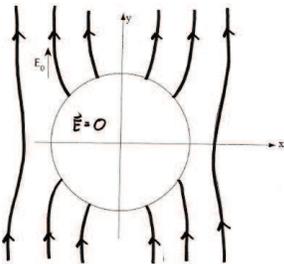
c)



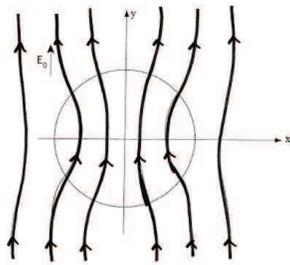
d)



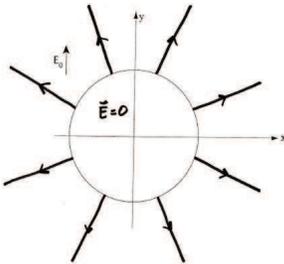
e)



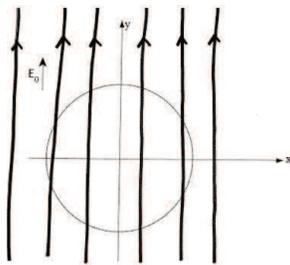
f)



g)



h)



i) None of these
(please elaborate
in the space below)

Q11 - For the conducting cylinder with radius **a** given in the previous problem (also shown below), we want to use the method of separation variables to solve for:

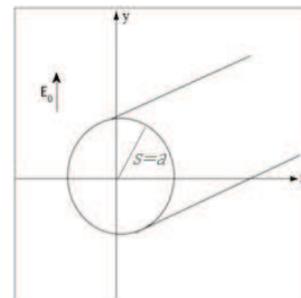
(a) the potential everywhere **AND** (b) the surface charge density, σ

i. What are the boundary conditions on V *at the surface* ($s = a$)

needed to do this? Here, s and ϕ are the usual cylindrical variables.

Select **ALL** that are suitable.

- a) $V_{in} = V_{out}$ at $s = a$
- b) $V_{out} \rightarrow \infty$ at $s = a$
- c) $V_{out} = 0$ (or a non-zero constant) at $s = a$
- d) $\frac{\partial V_{out}}{\partial s} \Big|_a - \frac{\partial V_{in}}{\partial s} \Big|_a = -\frac{\sigma}{\epsilon_0}$
- e) $\frac{\partial V_{out}}{\partial s} \Big|_a - \frac{\partial V_{in}}{\partial s} \Big|_a = 0$
- f) $\frac{1}{s} \frac{\partial V_{out}}{\partial \phi} \Big|_a - \frac{1}{s} \frac{\partial V_{in}}{\partial \phi} \Big|_a = -\frac{\sigma}{\epsilon_0}$



(Optional) If you would like to elaborate further, please do so in the box below

ii. What are the boundary conditions on E *at the surface* ($s = a$) needed to do this?

Select **ALL** that are suitable.

- a) $\vec{E}_{in} = \vec{E}_{out}$ at $s = a$
- b) $\vec{E}_{out} \rightarrow \infty$ at $s = a$
- c) $\vec{E}_{out} = 0$ at $s = a$
- d) $E_{out}^{\parallel} - E_{in}^{\parallel} = \frac{\sigma}{\epsilon_0}$
- e) $E_{out}^{\parallel} = E_{in}^{\parallel}$
- f) $E_{out}^{\perp} - E_{in}^{\perp} = \frac{\sigma}{\epsilon_0}$
- g) $E_{out}^{\perp} = E_{in}^{\perp}$

(Optional) If you would like to elaborate further, please do so in the box below

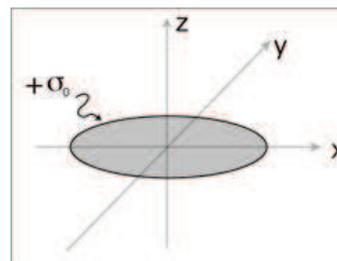
Q12 - The following set of problems refer to the uniform flat, infinitely thin disk of radius **R** carrying uniform positive surface charge density $+\sigma_o$ as in the figure.

i. How does the z-component of the electric field along the z-axis

(E_z) behave as you get **very far from the disk** ($z \gg R$).

Select only one.

- a) E_z goes to an arbitrary constant
- b) E_z goes to $\sigma_o/2\epsilon_o$
- c) E_z goes to σ_o/ϵ_o
- d) E_z goes to ∞
- e) E_z varies as $1/z$
- f) E_z varies as $1/z^2$
- g) E_z varies as $1/z^n$ where n is a positive integer other than 1 or 2
- h) $E_z = 0$



(Optional) If you would like to elaborate further, please do so in the box below

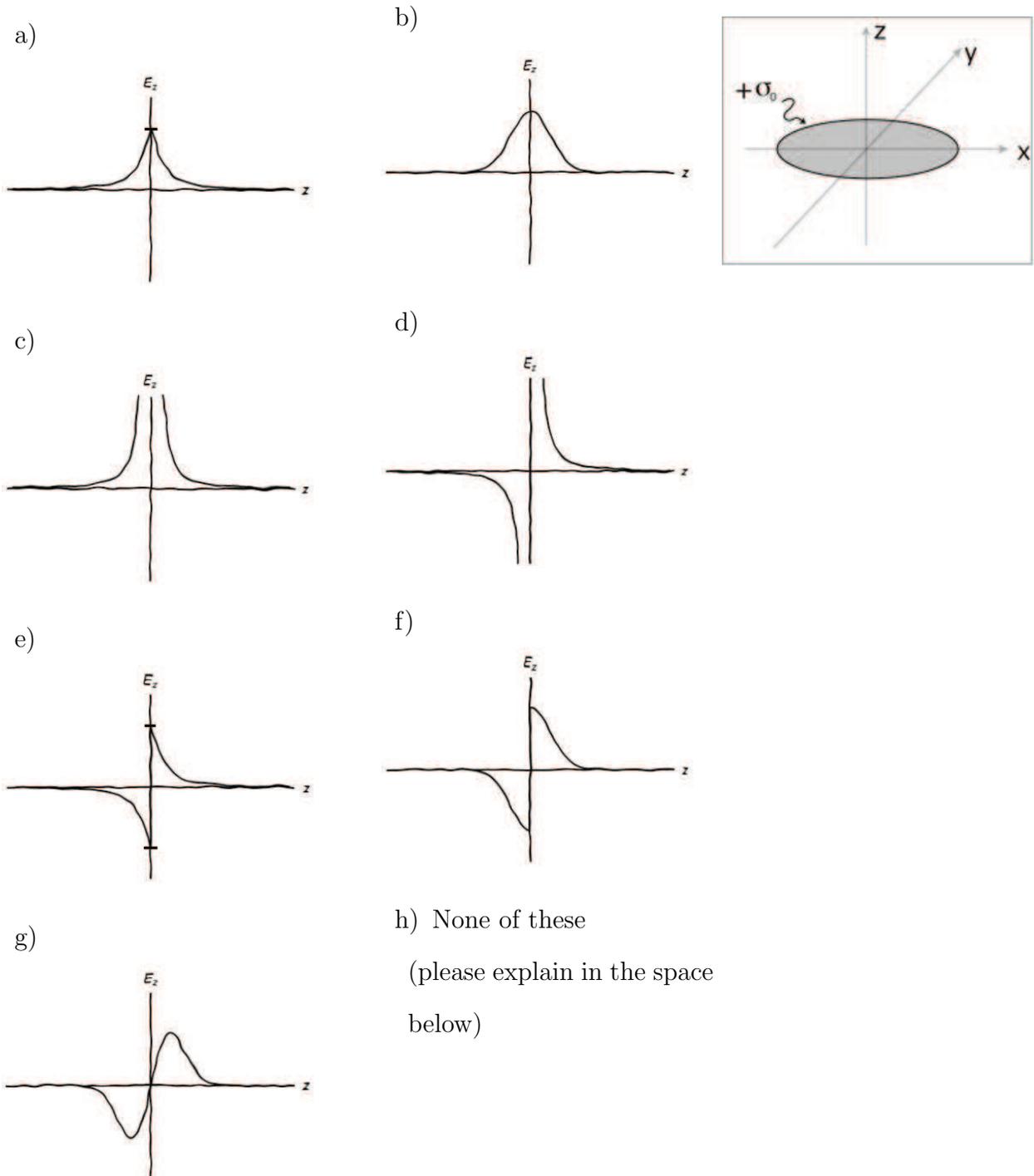
ii. How does E_z behave **very near the origin** ($R \gg z > 0$)?

Select only one.

- a) E_z goes to an arbitrary constant
- b) E_z goes to $\sigma_o/2\epsilon_o$
- c) E_z goes to σ_o/ϵ_o
- d) E_z goes to ∞
- e) E_z varies as $1/z$
- f) E_z varies as $1/z^2$
- g) E_z varies as $1/z^n$ where n is a positive integer other than 1 or 2
- h) $E_z = 0$

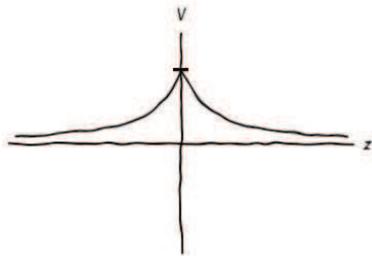
(Optional) If you would like to elaborate further, please do so in the box below

iii. Which one of the following qualitative graphs best represents the **relative sign and magnitude** of E_z as you move away from the disk along the z-axis?

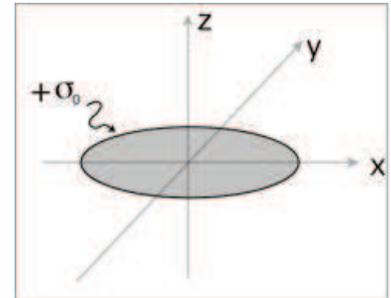
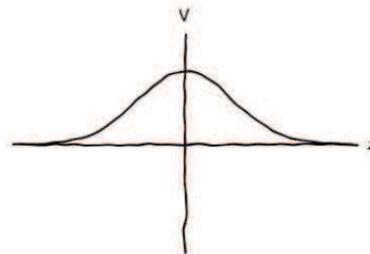


iv. Which one of the following qualitative graphs best represents the **relative sign and magnitude of V** as you move away from the disk along the z -axis?

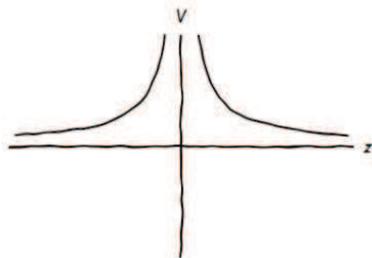
a)



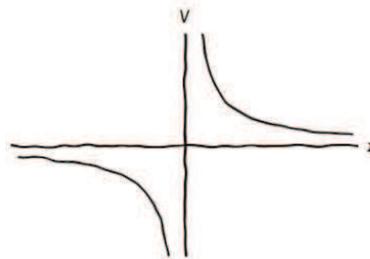
b)



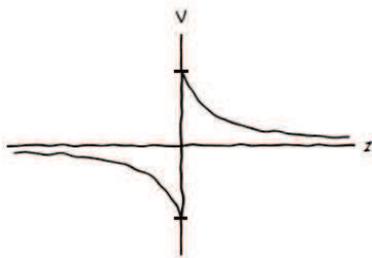
c)



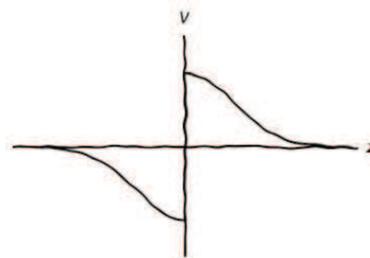
d)



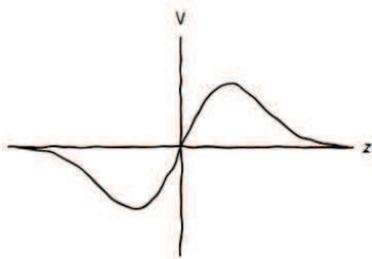
e)



f)



g)



h) None of these

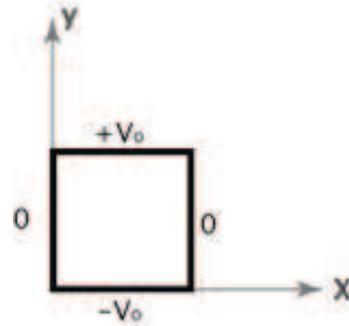
(please explain in the space below)

Q13 - You are given a 2-D box with potentials specified on the boundary as indicated in the figure below. Inside the box, the voltage obeys Laplace's equation, $\nabla^2 V = 0$.

To solve for V inside the box by separation of variables, which form of the solution should you choose? Here k is a positive and real constant.

Select only one.

- A. $V(x,y) = (A e^{ky} + B e^{-ky})(C \sin kx + D \cos kx)$
- B. $V(x,y) = (A \sin ky + B \cos ky)(C e^{kx} + D e^{-kx})$
- C. $V(x,y) = (A \sin ky + B \cos ky)(C \sin kx + D \cos kx)$
- D. $V(x,y) = (A e^{ky} + B e^{-ky})(C e^{kx} + D e^{-kx})$
- E. More than one of these could be used

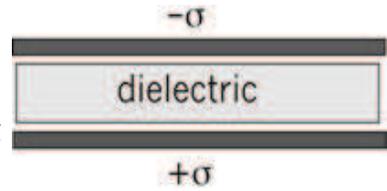


because ... (select ALL that support your choice of method)

- a. it is harder (but not impossible) to fit the boundary conditions on x with exponentials
- b. it is harder (but not impossible) to fit the boundary conditions on y with exponentials
- c. oscillatory solutions *cannot* fit the boundary conditions on x
- d. oscillatory solutions *cannot* fit the boundary conditions on y
- e. exponential solutions *cannot* fit the boundary conditions on x
- f. exponential solutions *cannot* fit the boundary conditions on y

(Optional) If you would like to elaborate further, please do so in the box below

Q14 - A linear, neutral, and isotropic dielectric is inserted into an isolated but charged infinite parallel plate capacitor, as shown. The dielectric fills the space without quite touching the plates, which are fixed in position.



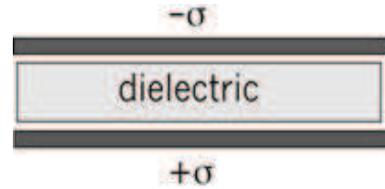
i. What happens *to the dielectric* (both in the bulk and at the surfaces) when it is inserted into the capacitor?

Select **ALL** that apply.

- a) Nothing, since it is non-conducting the dielectric is not affected by the capacitor
- b) A polarization, P , is induced in the dielectric resulting in a $+\sigma_b$ on the top surface and $-\sigma_b$ on the bottom surface
- c) A polarization, P , is induced in the dielectric resulting in a $-\sigma_b$ on the top surface and $+\sigma_b$ on the bottom surface
- d) A polarization, P , is induced in the dielectric resulting in a non-zero ρ_b in the bulk
- e) Free $+$ and $-$ charges in the dielectric move to the surfaces resulting in a $+\sigma_f$ on the top surface and $-\sigma_f$ on the bottom surface
- f) Free $+$ and $-$ charges in the dielectric move to the surfaces resulting in a $-\sigma_f$ on the top surface and $+\sigma_f$ on the bottom surface

(Optional) If you would like to elaborate further, please do so in the box below

iii. In the limit that the dielectric is infinitely polarizable (i.e., the electric susceptibility, $\chi_e \rightarrow \infty$) what would be the limiting value of the **magnitude of the charge densities in the dielectric?**



Select **ALL** that apply.

- a. $|\sigma_b|$ on the surfaces $\rightarrow 0$
- b. $|\sigma_b|$ on the surfaces $\rightarrow \infty$
- c. $|\sigma_b|$ on the surfaces $\rightarrow \sigma$
- d. $|\rho_b|$ on the surfaces $\rightarrow 0$
- e. $|\rho_b|$ on the surfaces $\rightarrow \infty$

(Optional) If you would like to elaborate further, please do so in the box below

iv. In the limit that the dielectric is infinitely polarizable (i.e., the electric susceptibility, $\chi_e \rightarrow \infty$) what would be the limiting value of the **net electric field** in the dielectric, $\vec{E}_{NET} = \vec{E}_{plates} + \vec{E}_{dielectric}$?

Select only one.

- a) $\vec{E}_{NET} \rightarrow 0$
- b) $\vec{E}_{NET} \rightarrow \infty$ pointing upwards
- c) $\vec{E}_{NET} \rightarrow \infty$ pointing downwards
- d) $\vec{E}_{NET} \rightarrow \sigma/\epsilon_o$ pointing upwards
- e) $\vec{E}_{NET} \rightarrow \sigma/\epsilon_o$ pointing downwards
- f) $\vec{E}_{NET} \rightarrow \sigma/2\epsilon_o$ pointing upwards
- g) $\vec{E}_{NET} \rightarrow \sigma/2\epsilon_o$ pointing downwards

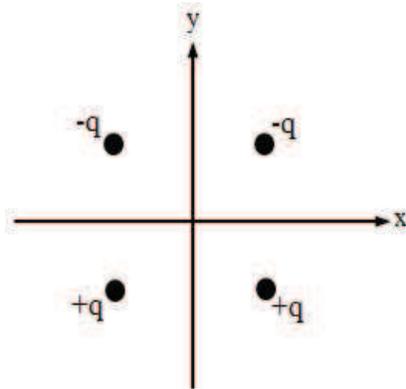
(Optional) If you would like to elaborate further, please do so in the box below

Q15 - You are given the following charge distributions made of point charges; each located a distance **a** from the x- and/or y-axis.

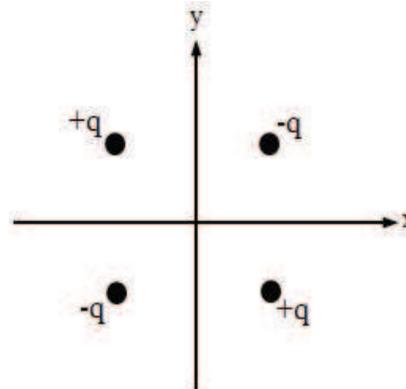
Given our choice of origin, which of the following charge distributions have a *non-zero* dipole moment?

Select **ALL** that apply.

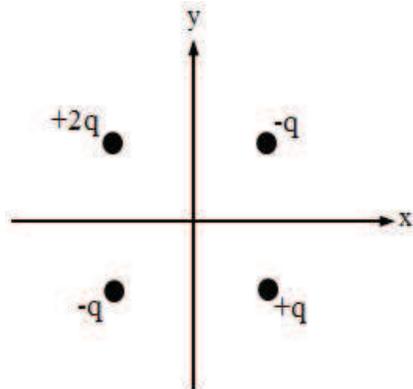
a.



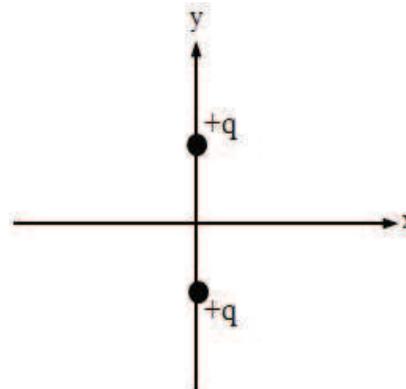
b.



c.



d.

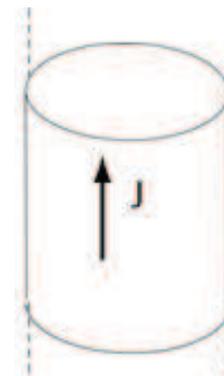


e. None of these

Q16 - Consider an infinite, non-magnetizable cylinder with a uniform volume current density \mathbf{J} (as in the figure).

Select only one: **Where is the B field maximum?**

- A. At the center of the cylinder
- B. Somewhere between the center and edge of the cylinder
- C. At the edge of the cylinder
- D. Somewhere between the edge of the cylinder and $r \rightarrow \infty$
- E. At $r \rightarrow \infty$
- F. B is maximum at more than one radius



because ... (select **ALL** that support your choice)

- a. B is constant outside because the cylinder is infinite
- b. you can use the right-hand rule
- c. you can use Ampere's law
- d. the current enclosed in maximum
- e. the circumference of a circular loop increases as r
- f. B is proportional to $\frac{1}{r}$ for all $r > 0$

(Optional) If you would like to elaborate further, please do so in the box below

How seriously did you just take this diagnostic exam?

- (a) I pretty much blew it off, didn't think much about a lot of the answers.
- (b) I took it sort of seriously, but when I didn't know an answer I didn't think very hard about it
- (c) I took it seriously, and thought about my answers

If you imagine getting a letter grade on the portion of this test that you were able to complete within the time limit, what do you think that grade would be? _____

Any other comments?

I. Detailed R2 Grading Scheme for the CMR CUE

Full nuanced grading scheme for CMR CUE (v9)

Calculation of Scores

Total score on questions 1-7 is broken into 3pts for Method and 2pts for Reasoning. The scores on the Reasoning section are conditional on the selection in the Method section.

For the Reasoning sections, add together the points for each reasoning element given the selection of method. If the total is negative, give no points for Reasoning (0/2) and if the total is greater than 2pts, give full credit for Reasoning (2/2). Any reasoning element (RE) not listed gets + 0pts.

Q1. Separation of Variables (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
C	+3	c	+1.0
		f	+1.0
		a,b, or e	-2.0
D,H	+0	c&f	+2.0
A,B,E,F, or G	+0	Any	+0

Q2. Cube Close (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
A	+3	a	+2.0
		e	+0.5
		f	+0.5
		b,c,d, or g	-2.0
D	+0.5 (with RE b)	b	+0.5
	+0 (without RE b)	a,c,d,e or g	-3.0
B,C,E,F,G or H	+0	Any	+0

Q3. Cube Far (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
A	+1	a	+0.5
		b,c,d or g	-0.5
D	+3 (with RE b) +2 (without RE b)	b	+1.0
		e	+1.5
		f	+0.5
		a,c,d or g	-3.0
H	+0	b	+1.0
		b& e	+2.0
B,C,E,F or G	+0	Any	+0

Q4. Method of Images (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
F	+3	b	+1.0
		d	+1.0
		a,c or f	-2.0
G	+2 (with RE b) +0 (without RE b)	b	+1.0
		d	+1.0
A,B,C,D,E or H	+0	Any	+0

Q5. Superposition (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
A	+1	a	+0.5
		e	+1.0
		g	+1.0
		b,d,f or h	-2.5
B or BG	+3 (with RE e&g)	ceg	+2.0
	+0 (without RE e&g)	eg	+2.0
		a,b,d,f or h	-2.0
D	+0.5 (with RE b)	b	+0.5
	+0 (without RE b)	a or h	-2.0
G	+3 (without RE d or f) +1 (with RE d or f)	c	+0.5
		e	+1.0
		g	+1.0
		a,b or h	-2.5
C,E,F or H	+0	Any	+0

Q6. Loop (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
A	+1	a	+0.5
		b,c,d or g	-0.5
D	+3 (with RE b) +2 (without RE b)	b	+1.0
		e	+1.5
		f	+0.5
		a,c,d or g	-3.0
H	+0	b	+1.0
		b& e	+2.0
B,C,E,F or G	+0	Any	+0

Q7. Gauss (5pts)

Method Selection	Points (3pt max)	Reasoning Selection	Points (2pt max, 0pt min)
A	+1	a	+0.5
		c or g	-0.5
B	+3	b	+1.0
		e	+1.0
		a,c or g	-2.0
C,D,E,F,G or H	+0	Any	+0

Q8. Delta (5pts)

Part	Selection	Points
i (3pt max)	c	+3
	a,b,c or d	-0
ii (2pt max)	e	+2
	a,b,c or d	+0

Q9. Reference point for V (5pts)

Method Selection	Points (1pt max)	Reasoning Selection	Points (4pt max, 0pt min)
A	+0	a (without b)	+1.0
		a (with b)	+0
		d (without e)	+1.0
		d (with e)	+0
		b,c,e or g	-0.5
B	+1	a (without b)	+4.0
		a (with b)	+0
		d (without e)	+4.0
		d (with e)	+0
		b,c,e or g	-3.0
C	+0	a (without b)	+1.0
		a (with b)	+0
		d (without e)	+1.0
		d (with e)	+0
		b,c,e or g	-0.5
D	+0	Any	+0

Q10. Cylinder (8pts)

Part	Selection	Points	
		Choice	Consistency
i (3pt max)	f	+3	-
	h	+2.5	-
	d or e	+1.0	-
	b,c or g	+0.5	-
ii (5pt max- 4pts choice, 1pt consistency)	a	+1	+0
	b	+0	+1 (with part i b,d,f or h)
	c	+1	+1 (with part i a,c,e or g)
	d	+1	+1 (with part i b,d,f or h)
	e	+4	+1 (with part i b,d,f or h)
	f	+1	+1 (with part i b,d,f or h)
	g	+0	+0
	h	+0	+0

Q11. BC's on E and V (6pts)

Part	Selection	Points
i (3pt max)	a (without b)	+1
	a (with b)	+0
	c (without b)	+0.5
	c (with b)	+0
	d (without e or f)	+1.5
	d (with e or f)	+0
ii (3pt max)	e (without d)	+1.5
	e (with d)	+0
	f (without g)	+1.5
	f (with g)	+0
	a,b or c	-1.5

Q12. Disk (14pts)

Part	Selection	Points	
		Choice	Consistency
i (2pt max)	f	+2	-
ii (2pt max)	a	+0.5	-
	b	+2.0	-
	c	+1.0	-
iii (5pt max- 4pts choice, 1pt consistency)	a	+0.5	+1 (with part ii a,b or c)
	b	+0	+1 (with part ii a,b or c)
	c	+0	+1 (with part ii d,e,f or g)
	d	+0.5	+1 (with part ii d,e,f or g)
	e	+4	+1 (with part ii a,b or c)
	f	+3.5	+1 (with part ii a,b or c)
	g	+0.5	+1 (with part ii h)
iv (5pt max- 4pts choice, 1pt consistency)	a	+4	+1 (with part iii e))
	b	+3.5	+1 (with part iii g)
	c	+0.5	+1 (with part iii d)
	d	+0	+1 (with part iii c)
	e	+0.5	+1 (with part iii b or a)
	f	+0.5	+0
	g	+0	+0

Q13. Cartesian SoV (5pts)

Answer	Points	Reasoning	Points
Selection	(2pt max)	Selection	(3pt max, 0pt min)
A	+2	e	+3.0
		d	-1.0
		a,c or f	-3.0
B,C,D or E	+0	Any	+0

Q14. Polarization (6pts)

Part	Selection	Points
i (2pt max)	b (without e)	+2
	b (with e)	+0.5
	e	+0.5
	d	-1.0
	a,c or f	-2.0
ii (2pt max)	c	+2
	a,b or e	-2
iii (2pt max)	a	+2

Q15. Dipole (4pts)

Selection	Points
b	+3
c	+1
a or d	-3
e	-4

Q16. Ampere (5pts)

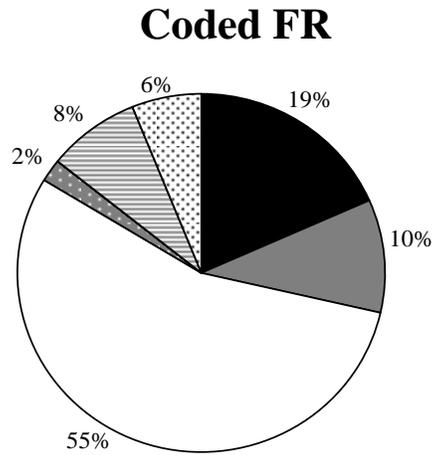
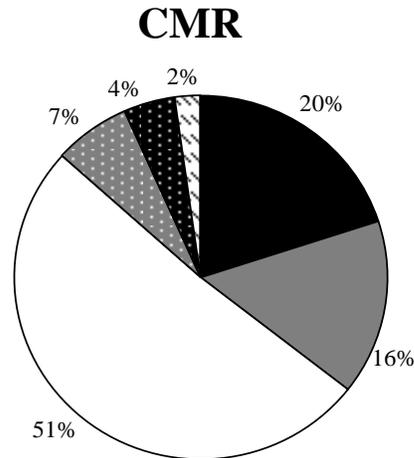
Answer Selection	Points (1pt max)	Reasoning Selection	Points (4pt max, 0pt min)
C	+1	c	+1.0
		d	+1.5
		e	+1.5
		a	-4.0
		f	-3.0
A,B,D,E or F	+0	Any	+0

J. Method Selections for Q1-7 on the CMR and FR CUE

Side-by-side comparison of students' method selections on the CMR and FR versions of the CUE. FR selections were coded from students' responses to the FR version during the 2-semester comparison study (Sec. V A 3).

Question 1 - Separation of Variables

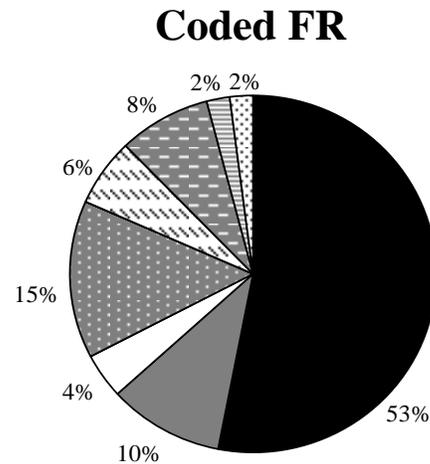
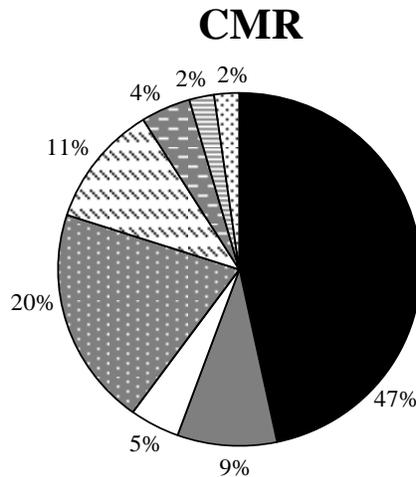
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q1 distributions is not statistically significant, $p = 0.1$.

Question 2 - Cube Close

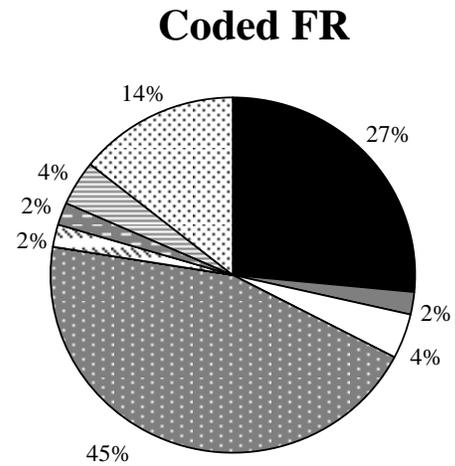
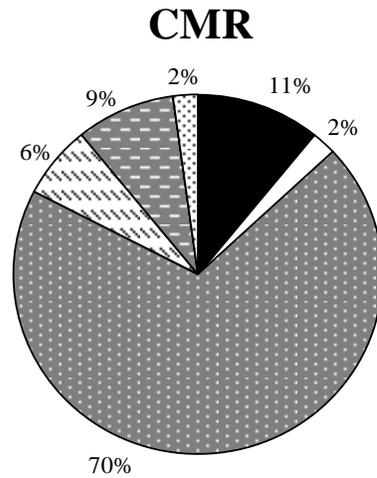
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q2 distributions is not statistically significant, $p = 0.9$.

Question 3 - Cube far

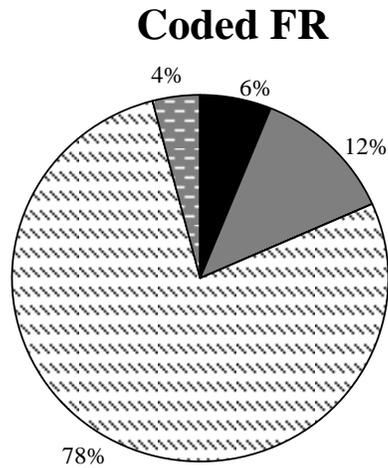
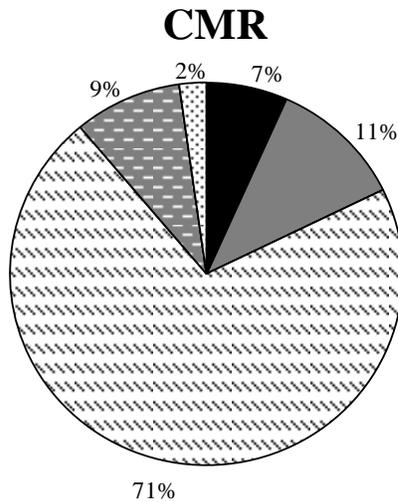
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q3 distributions is statistically significant, $p = 0.02$.

Question 4 - Method of Images

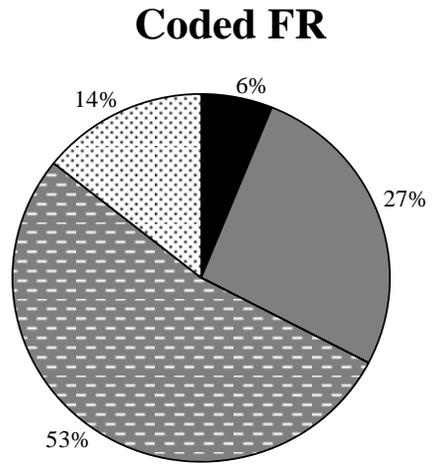
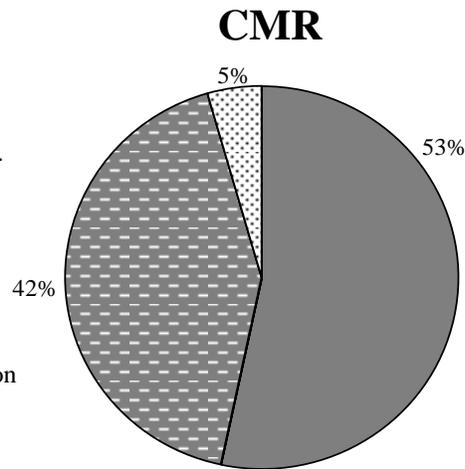
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q4 distributions is not statistically significant, $p = 0.8$.

Question 5 - Superposition

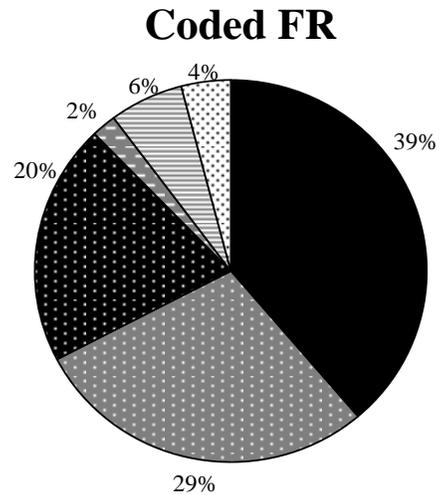
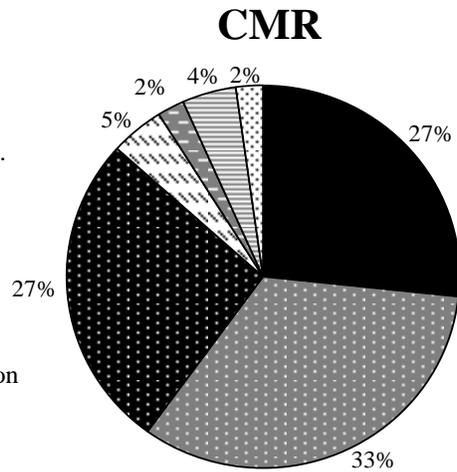
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q5 distributions is statistically significant, $p = 0.02$.

Question 6 - Loop

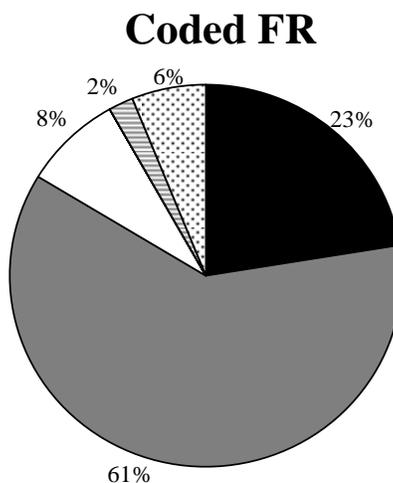
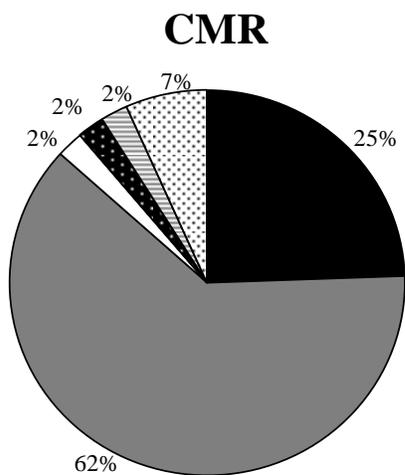
- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



The difference between Q6 distributions is not statistically significant, $p = 0.7$.

Question 7 - Gauss

- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- Images
- Superposition
- None
- Blank



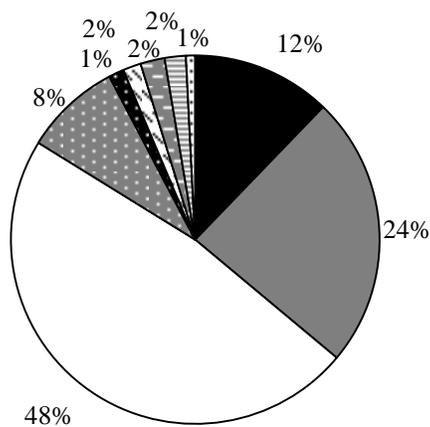
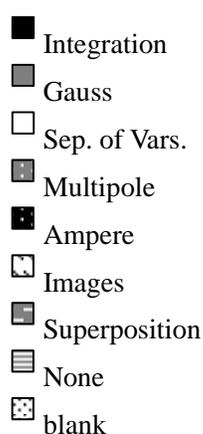
The difference between Q7 distributions is not statistically significant, $p = 0.8$.

K. Student Responses to All Questions on the CMR CUE

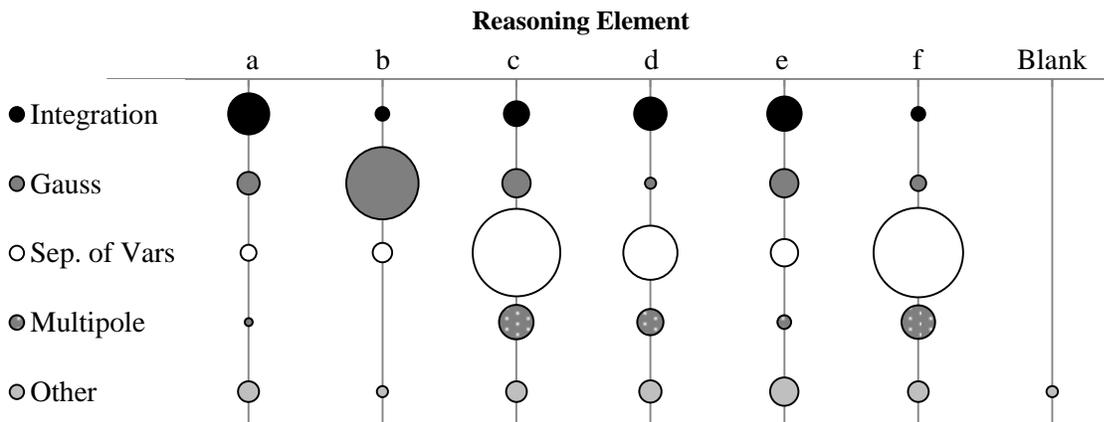
Distributions of N=373 student responses to questions 1-16 on the CMR CUE. Data represent the full pool of available CMR CUE data from all institutions and courses described in Sec. V B. Only methods/answers that received greater than 5% of the overall votes are given reasoning element breakdowns.

Question 1 - Separation of Variables

(a) Method Selection



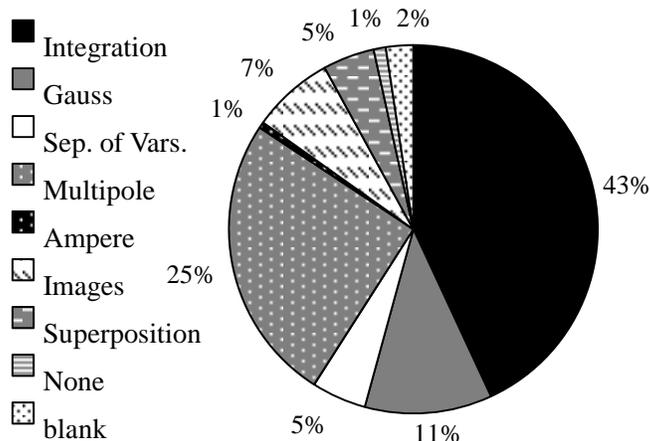
(b) Reasoning Selection



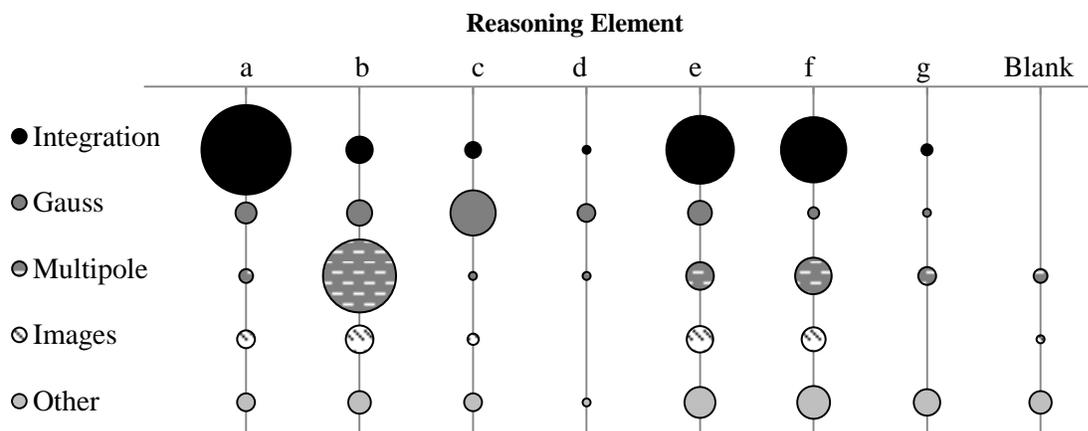
(a) Distribution of method selections for Q1. Separation of variables is the easiest and only possible method. (b) Reasoning selections for the most common method selections: Direct Integration (N=46), Gauss's law (N=90), Separation of Variables (N=180), Multipole Expansion (N=32), and other (N=25). Number of responses for each reasoning element is proportional to the area of the circle.

Question 2 - Cube Close

(a) Method Selection



(b) Reasoning Selection

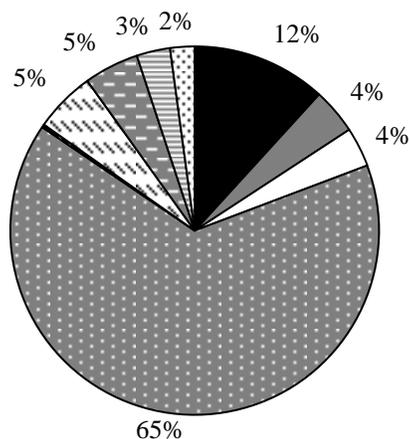


(a) Distribution of method selections for Q2. Direct integration is the easiest and only completely correct method. (b) Reasoning selections for the most common method selections: Direct Integration (N=162), Gauss's law (N=42), Multipole Expansion (N=95), Method of Images (N=27), and other (N=47). Number of responses for each reasoning element is proportional to the area of the circle.

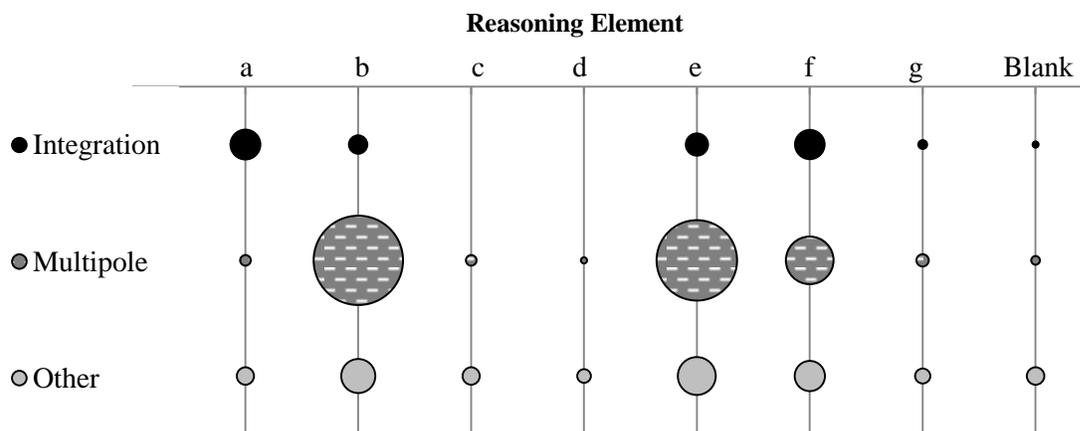
Question 3 - Cube Far

(a) Method Selection

- Integration
- Gauss
- Sep. of Vars.
- Multipole
- Ampere
- ▨ Images
- Superposition
- ▨ None
- ▨ blank



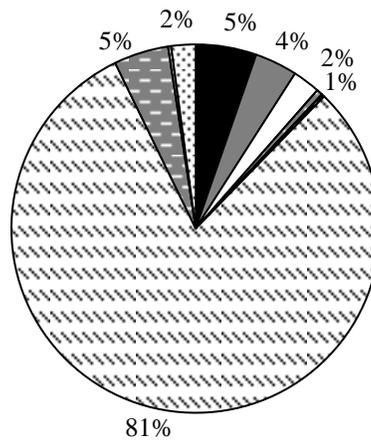
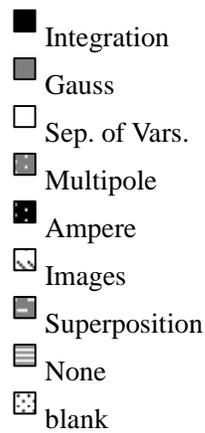
(b) Reasoning Selection



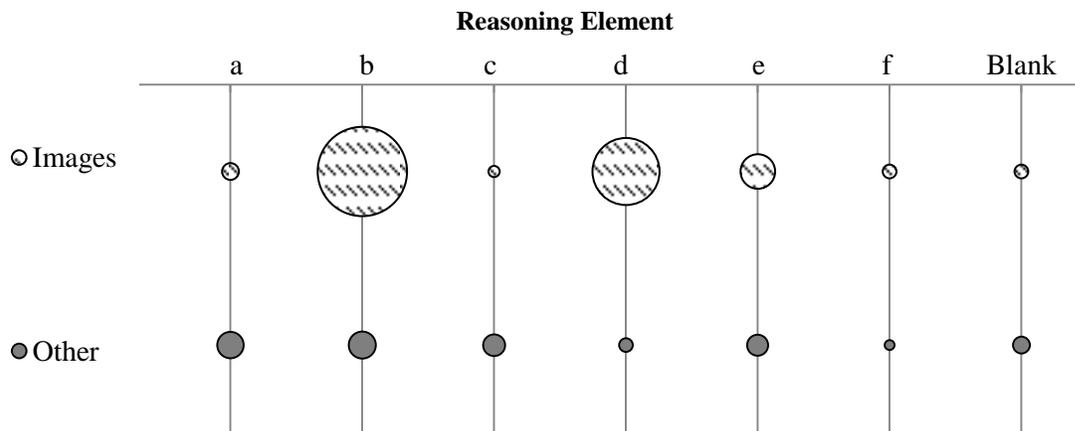
(a) Distribution of method selections for Q3. Multipole expansion is the correct method; however, Direct integration is also possible. (b) Reasoning selections for the most common method selections: Direct Integration (N=44), Multipole Expansion (N=244), and other (N=85). Number of responses for each reasoning element is proportional to the area of the circle.

Question 4 - Images

(a) Method Selection



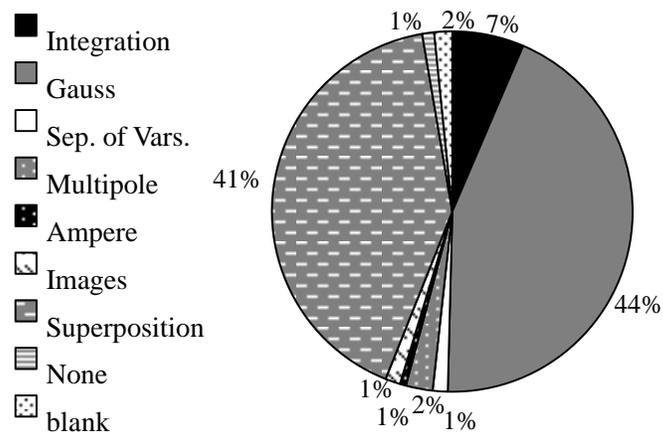
(b) Reasoning Selection



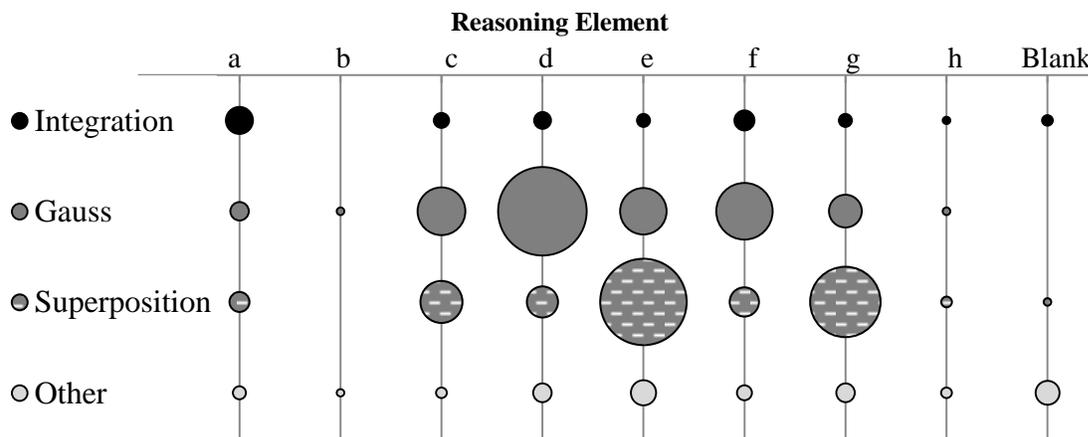
(a) Distribution of method selections for Q4. Method of images is the easiest and only possible method. (b) Reasoning selections for the most common method selections: Method of images (N=302), and other (N=71). Number of responses for each reasoning element is proportional to the area of the circle.

Question 5 - Superposition

(a) Method Selection



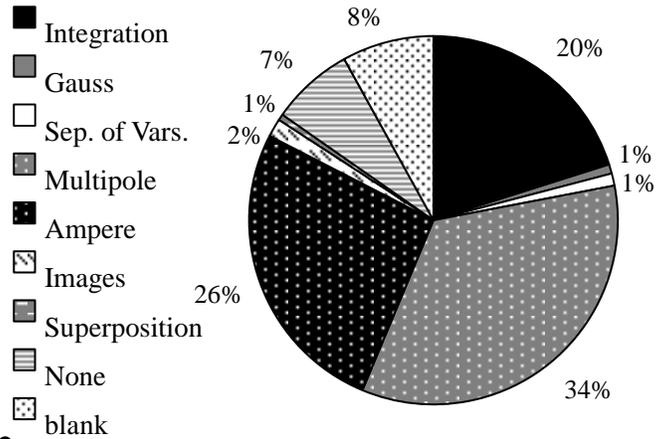
(b) Reasoning Selection



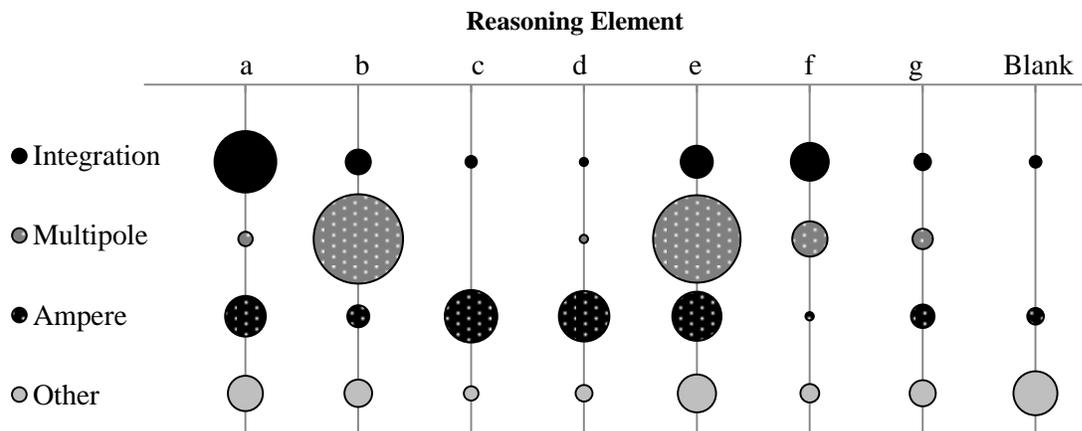
(a) Distribution of method selections for Q5. Superposition is the easiest method; however, Direct integration is also possible. (b) Reasoning selections for the most common method selections: Direct integration (N=24), Gauss's Law (N=164), Superposition (N=154), and other (N=31). Number of responses for each reasoning element is proportional to the area of the circle.

Question 6 - Loop

(a) Method Selection



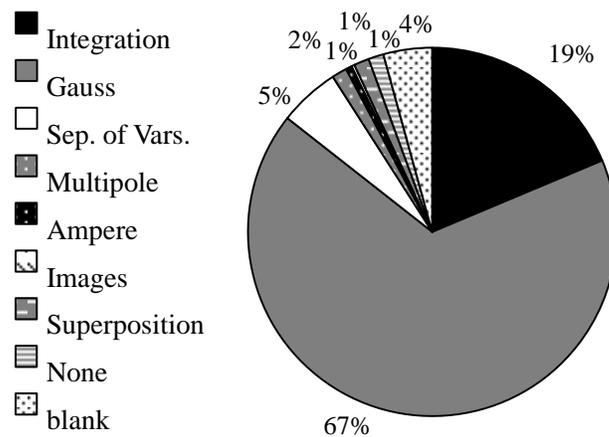
(b) Reasoning Selection



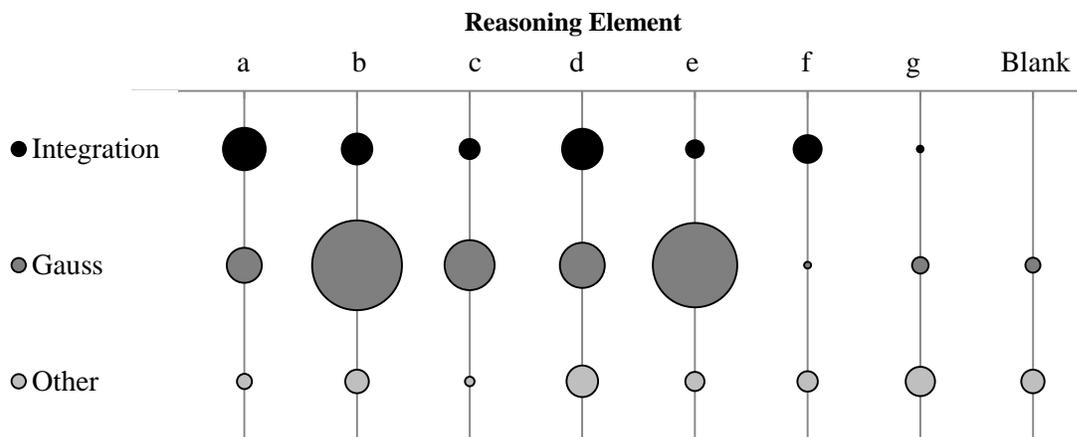
(a) Distribution of method selections for Q6. Multipole expansion is the correct method; however, Direct integration is also possible. (b) Reasoning selections for the most common method selections: Direct integration (N=75), Multipole Expansion (N=128), Ampere (N=98), and other (N=72). Number of responses for each reasoning element is proportional to the area of the circle.

Question 7 - Gauss

(a) Method Selection

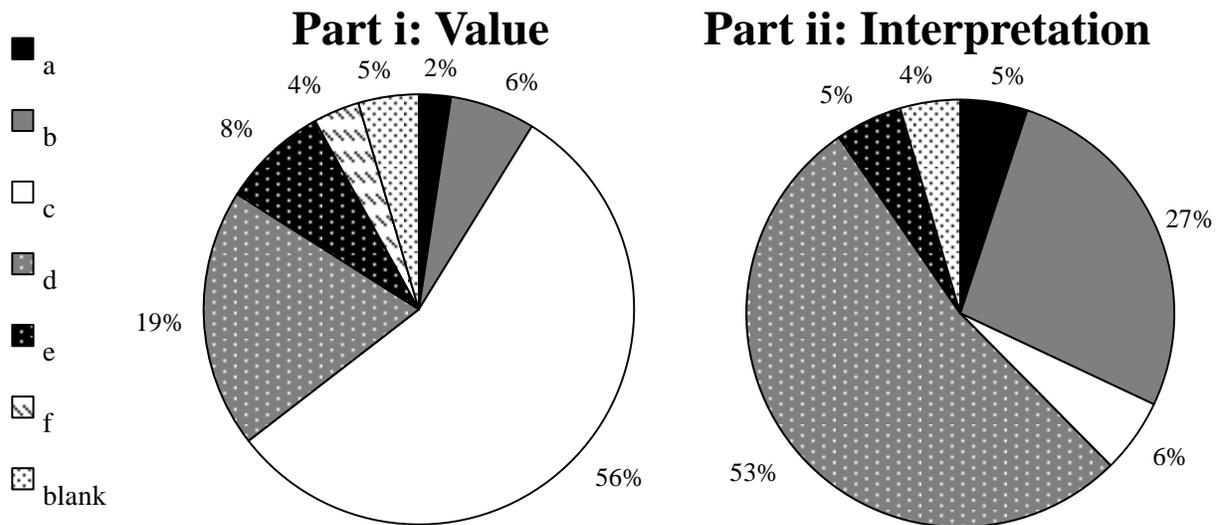


(b) Reasoning Selection



(a) Distribution of method selections for Q7. Gauss's Law is the correct method; however, Direct integration is also possible. (b) Reasoning selections for the most common method selections: Direct integration (N=70), Gauss's Law (N=251), and other (N=52). Number of responses for each reasoning element is proportional to the area of the circle.

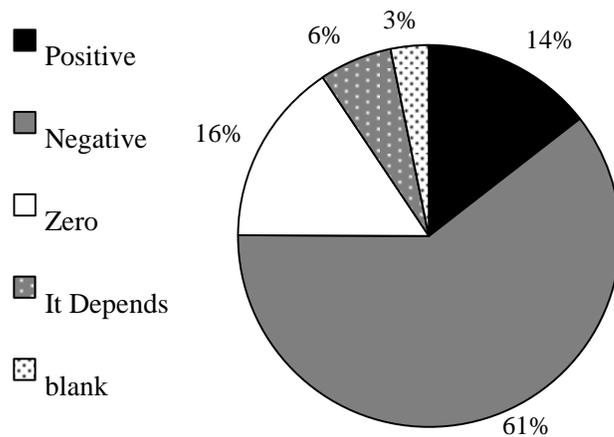
Question 8 - Delta



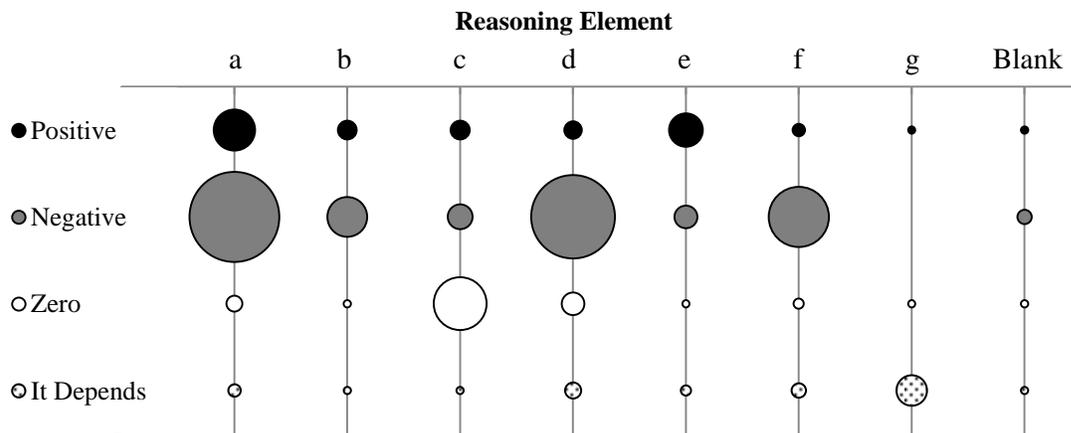
Distribution of N=373 responses to subparts i and ii on Q8; correct responses are c and d respectively.

Question 9 - Reference point for V

(i) Value

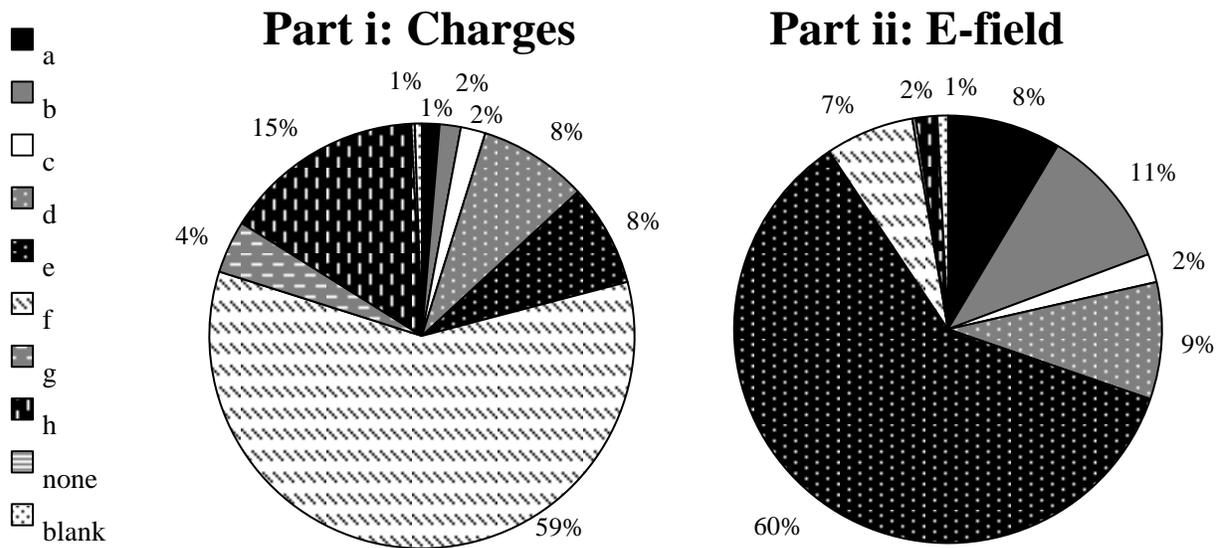


(ii) Interpretation



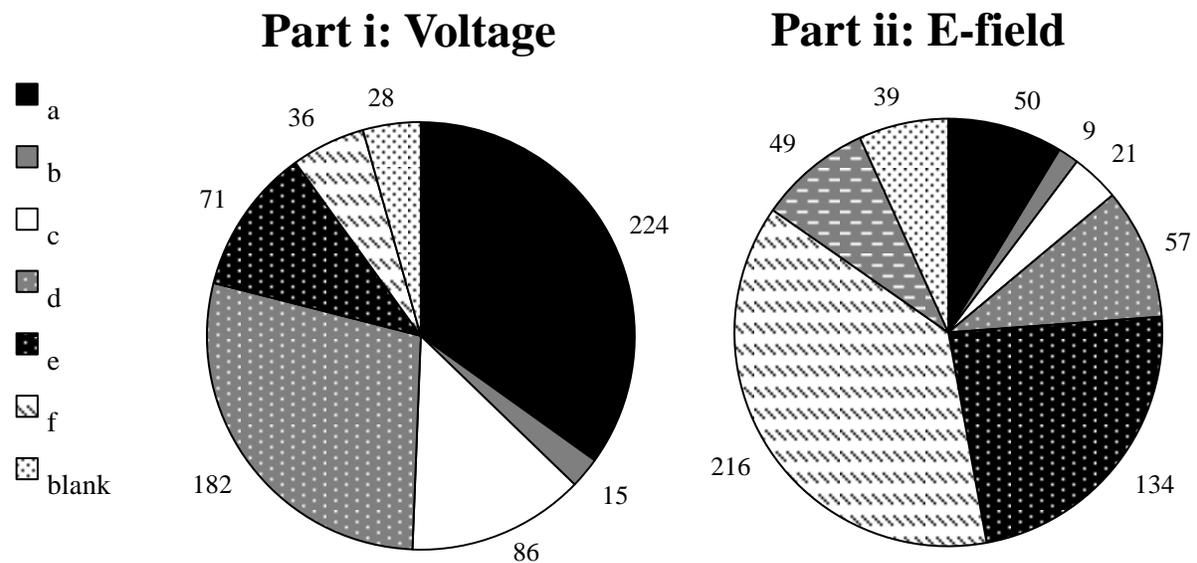
(i) Distribution of answer selections for Q9. The correct value for the potential at infinity is negative. (ii) Reasoning selections for the most commonly selected values: positive (N=54), negative (N=226), zero (N=58), and other (N=35). Number of responses for each reasoning element is proportional to the area of the circle.

Question 10 - Cylinder



Distribution of N=373 responses to subparts i and ii on Q10; correct responses are f and e respectively.

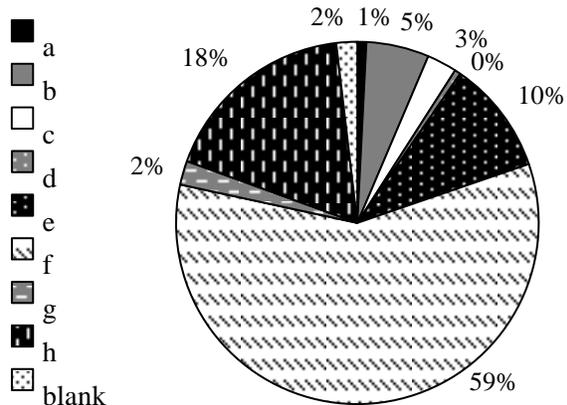
Question 11 - Boundary Conditions on E and V



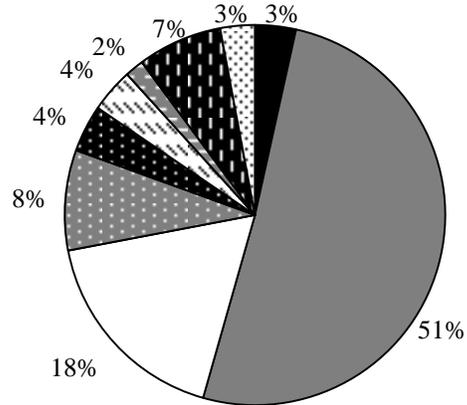
Distribution of N=373 responses to subparts i and ii on Q11; correct responses are a, c, & d on part i, e & f on part ii. Values for each option are number of responses rather than percent as students can select more than one option.

Question 12 - Finite Disk

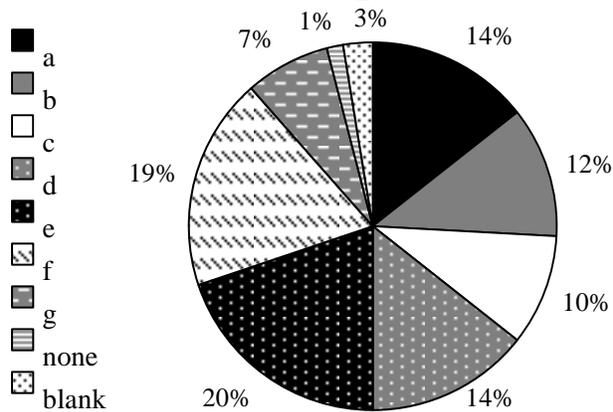
Part i: Disk Far



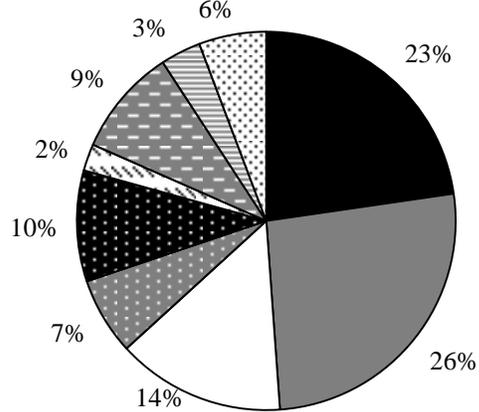
Part ii: Disk Close



Part iii: Graph E_z



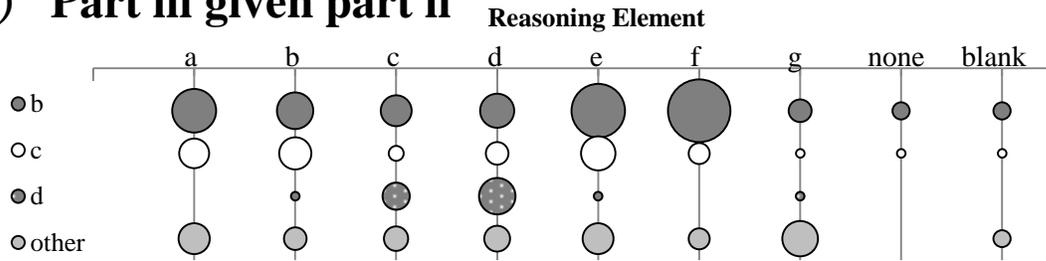
Part iv: Graph V



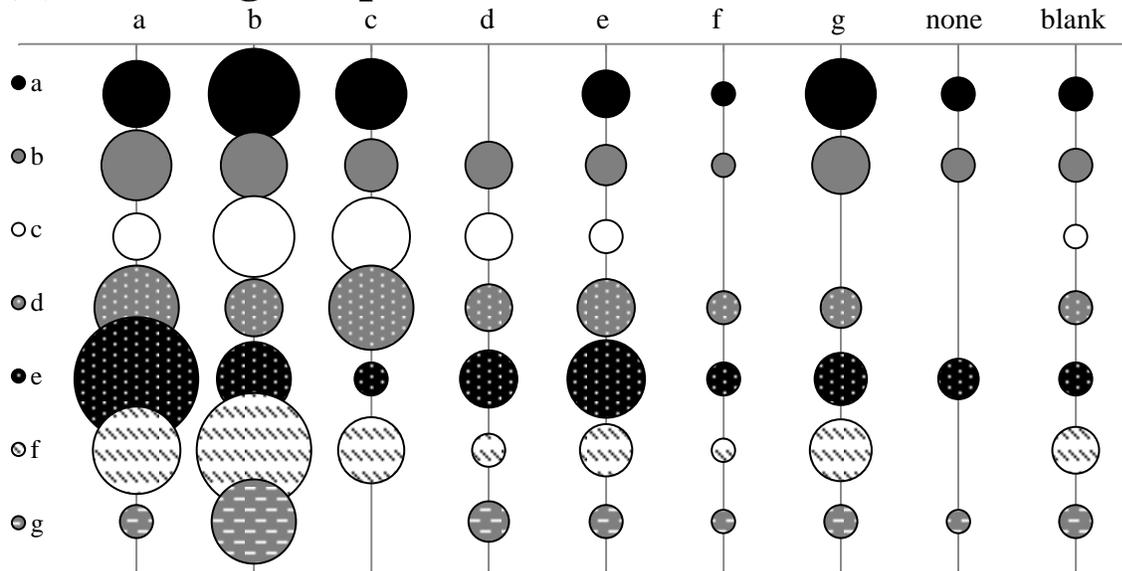
Distribution of responses to subparts i-iv on Q12; correct responses are f, b, e, and a respectively. Parts i and ii represent the full N=373 responses; however, the version of the CMR CUE used during the first semester of data collection included different response options that following semesters. Thus the distributions for parts iii and iv represent only a N=348 student subset.

Question 12 - Finite Disk

(a) Part iii given part ii



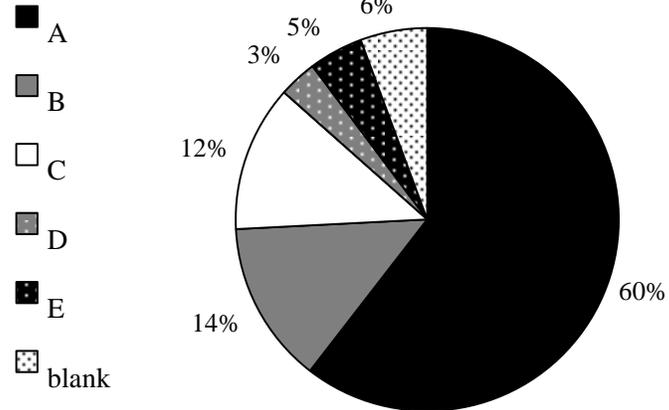
(b) Part iv given part iii



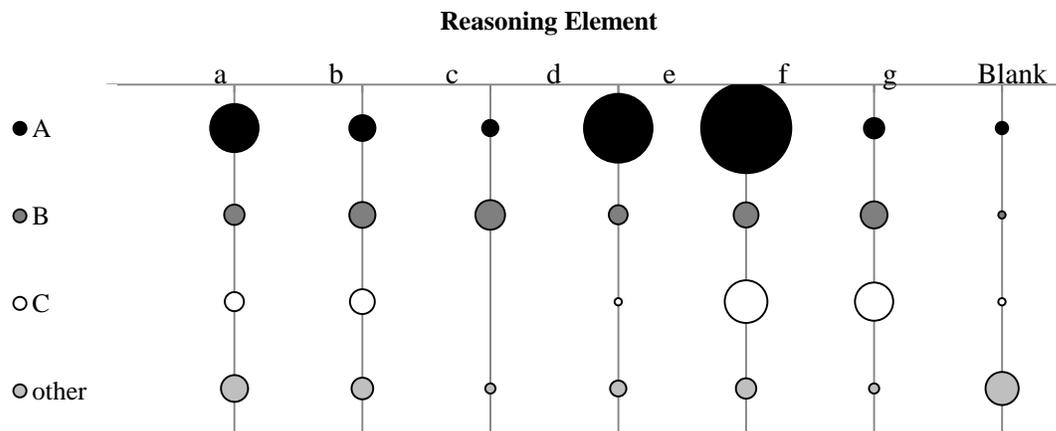
(a) Part iii selections for the most commonly selected responses to part ii: b (N=191), c (N=66), d (N=32), and other (N=84). (b) Part iv selections for the most commonly selected responses to part iii: a (N=50), b (N=40), c (N=34), d (N=50), e (N=69), f (N=65), and g (N=26). Number of responses for each reasoning element is proportional to the area of the circle. Responses represent the N=348 students who took the final version of Q12.

Question 13 - Cartesian SoV

(a) Answer Selection



(b) Reasoning Selection

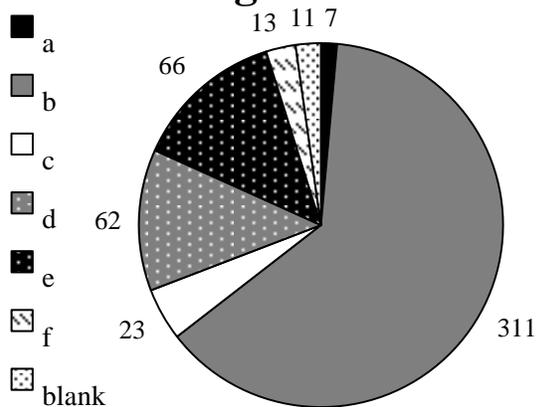


(a) Distribution of answer selections for Q13; the correct response is option A. (b)

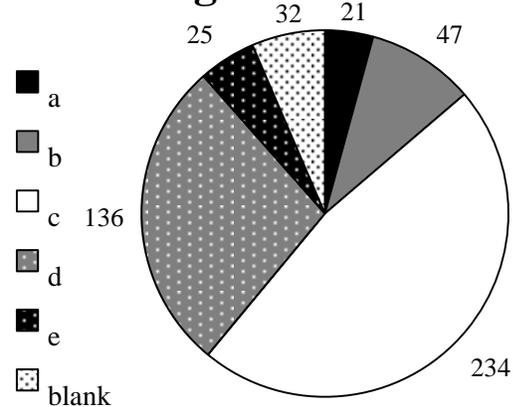
Reasoning selections for the most commonly selected values: A (N=225), B (N=51), C (N=46), and other (N=51). Number of responses for each reasoning element is proportional to the area of the circle.

Question 14 - Polarization

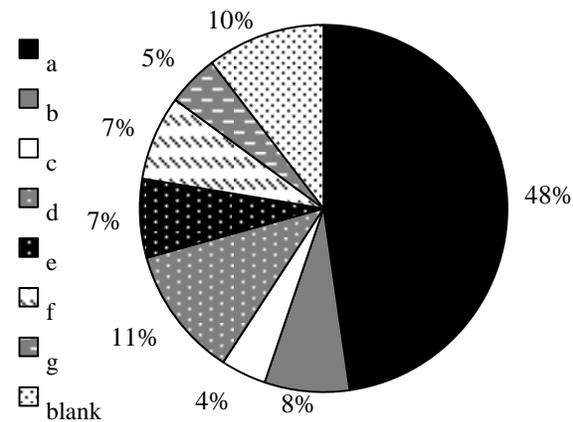
Part i: Charges



Part ii: Charges Limit



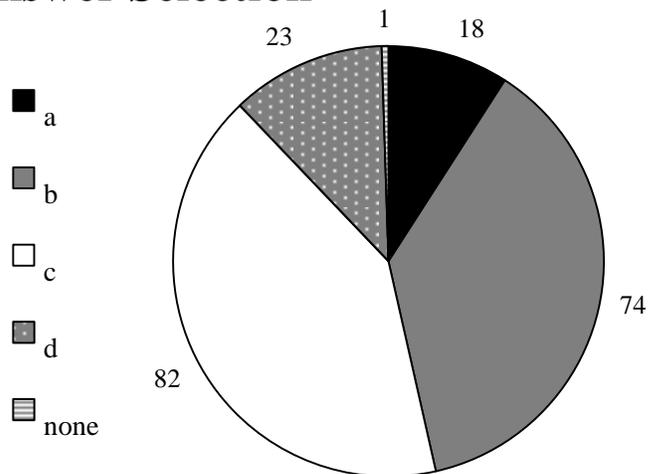
Part iii: E-field Limit



Distribution of responses to subparts i-iii on Q14; correct responses are b on part i, c on part ii, and a on part iii. Values for each option on parts i and ii are number of responses rather than percent as students can select more than one option. Values part iii are given in percent. Responses for parts i and iii are of N=373, while the distributions for part ii represent only the N=348 students who took the final version of the CMR CUE.

Question 15 - Dipole

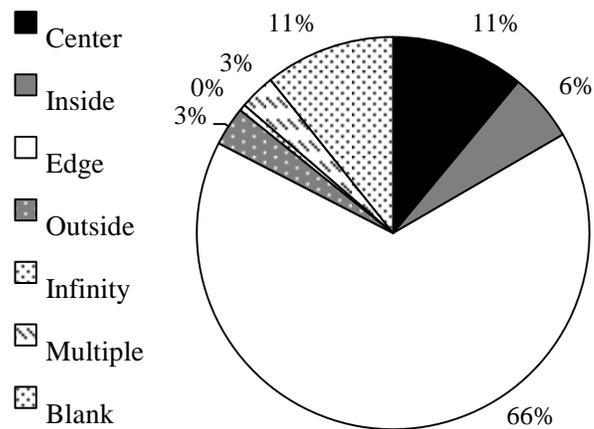
Answer Selection



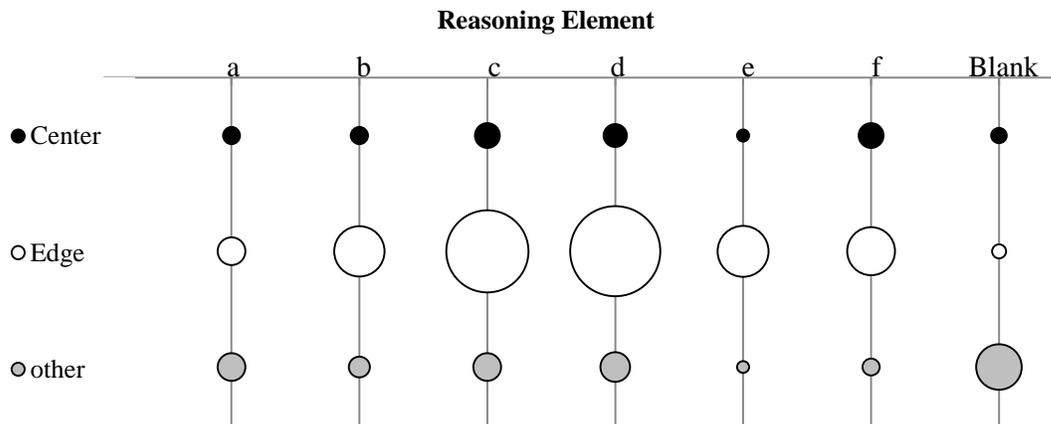
Distribution of N=373 responses to Q15; correct responses are b & c. Values for each option are number of responses rather than percent as students can select more than one option.

Question 16 - Ampere

(a) Answer Selection



(b) Reasoning Selection



(a) Distribution of answer selections for Q16; the correct response is that the B-field is maximum at the Edge. (b) Reasoning selections for the most commonly selected locations: Center (N=41), Edge (N=246), and other (N=86). Number of responses for each reasoning element is proportional to the area of the circle.