Content Sequencing and its Impact on Student Learning in Electromagnetism: Theory and Experiment

Benjamin J. Dringoli,1 Ksenia Kolosova,1 Thomas J. Rademaker,1 Juliann Wray,1 Jeremie Choquette,1,2 and Michael Hilke1

1Department of Physics, McGill University, 3600 rue University, Montréal, Québec H3A2T8, Canada
2Dawson College, 3040 Sherbrooke St W, Montréal, Québec H3Z 1A4, Canada

(Dated: October 2, 2019)

We investigate the impact of content sequencing on student learning outcomes in a first-year university electromagnetism course. Using a custom-built online system, the McGill Learning Platform (McLEAP), we test student problem-solving performance as a function of the sequence in which the students are presented aspects of new material. New material was divided into the three categories of conceptual, theoretical and example-based content. Here, we present findings from a two-year study with over 1000 students participating. We find that content sequencing has a significant impact on learning outcomes in our study: students presented with conceptual content first perform significantly better on our assessment than those presented with theoretical content. To explain these results, we propose the Content Cube as an extension to the mental model frameworks. Additionally, we find that instructors preferences for content sequencing differ significantly from that of students. We discuss how this information can be used to improve course instruction and student learning, and motivate future work building upon our presented results to study the impact of additional factors on student performance.

I. INTRODUCTION

For physics instructors, a fundamental question underlies the evaluation of lectures, lesson plans, and other learning materials: “Is there a presentation of educational content that results in optimal learning for students, and what factors affect this?” While this question is simple, its investigation incorporates psychology, neurobiology, education research, and subject-specific considerations [1–5]. Additionally, to understand which teaching material is best for student learning, one must study how students, individually and as a collective, learn [6], as well as what other factors besides the educational content provided can modify their learning outcome. These issues motivate the writing process of new textbooks within physics, as authors may feel that other texts or guidelines do not suit the perceived needs of their students. The resulting works, which if popular serve as teaching guides or templates for additional professors and students, thus inherit the writers personal teaching goals and pedagogical methods. The investigation of how different approaches impact student learning is an active area in STEM education research due in part to the increased focus on proficiency in science and math related subjects in the past decades [7], especially at the university level. In addition, investigation of this kind is a unique opportunity to probe the subtle differences between the preferences of new and experienced learners, such as populations of students and professors.

While teaching an introductory Electricity and Magnetism course for first-year University STEM students, we began to investigate which characteristics of educational content impact its teaching effectiveness. Even though there are many proposed theories for how students learn best [8–10], we find that within physics there are specific educational content types that introduce additional complexity not addressed in these prescribed frameworks, and thus should be explored in more detail. We aim to measure the effectiveness of certain presentations of physics content in improving student learning, and to find what factors (content type, content order, student preferences, etc.) are the most important to consider when designing or improving lesson plans.

In this work, we probe how a fundamental aspect of content presentation, the sequence in which the content types underlying physics material are presented, contribute to a students learning outcome. This is achieved through an experimental assignment which tests students problem solving skills after being presented with different content types and sequences. To accomplish this we introduce our McGill Learning Platform (McLEAP) tool, which offers a flexible online space for investigating pedagogical questions such as these, and explain our findings through our “Content Cube” model as a new way of visualizing the learning process in physics and evaluating content presentation schemes.

II. BACKGROUND/THEORY

A. Common textbooks

To discern some common qualitative approaches taken by the wider physics education community, we first explore content presentation within popular textbooks
aimed at teaching electromagnetism to university students. It is instructive to examine texts of different levels to see how authors treat different populations of learners with different needs, and it can inform us about how physics educators break down topics for student learning. A general physics textbook used in first-year courses (for example, *University Physics* by Young and Freedman) commonly starts with a broad explanation of the topics being introduced, setting the stage before introducing mathematics or theoretical arguments, and finishes with practice problems. This is less common in more advanced treatments: *Griffiths’ Introduction to Electrodynamics* is a common 2nd-3rd year undergraduate text and offers little conceptual explanation before presenting mathematical background, formulae, and derivations. These upper level texts also often lack the large number of worked out example problems that frequently follow the presentation of new information in introductory texts or books that favor applications over deep understanding.

We therefore see that physics texts uniquely partition educational content used in instruction, with authors prioritizing different content presentations depending on the intended audience of the text. To underscore the different types of content prioritized for different approaches, we give an additional example of each (with content type distinctions in bold):

- *Conceptual Physics* by Hewitt [11], uses concept-heavy lessons to introduce new students at the high school or undergraduate level, as concepts consist of general information and relatable analogies allowing for broad connection and recognition for students less accustomed to science.
- *Classical Electrodynamics* by Jackson [12] is heavily theoretical to prepare graduate students for advanced analysis, as theories represent the mechanics behind the way nature works in full detail, often accompanied or assisted by mathematical equations.
- *Introductory Electromagnetics* by Popovic and Popovic [13] is largely example-based to cater to its application-centric engineering audience, since examples serve as a road map which shows the correct setup and solution method of a simple and commonly encountered problem.

We find the classification of physics content into Example, Concept, or Theory types representative of our own experience and from these textbook archetypes and tested these distinctions. This was accomplished by a survey of our local physics department and students enrolled in physics classes, supporting our classification. The full description is explained in the Results section. Before investigating how these different content types affect student learning, though, we must also consider different frameworks of how students learn. Different theories of learning could dramatically affect how certain content types and their sequences are integrated by students, and therefore they must be considered so the results of our sequencing studies can be placed within an appropriate model of the learning process.

### B. Frameworks for Student Learning

To understand how the presentation of different types of educational content could affect student learning, one must explore how to present educational content and how students learn new topics. Both facets have been open and interesting questions in the field of education research for many years [6, 15–17], as improving learning outcomes through understanding the student mind is far from a novel approach. To develop the context for our investigations into the impact of content type and sequencing on student performance, we review some previously proposed theories of student learning. Throughout this explanation we will connect existing frameworks to the three content types within physics, and describe how they relate to content sequencing and our current study. We will then introduce our Content Cube learning framework to visualize the impact of sequencing of our content type divisions on student learning.

#### 1. Early Frameworks

Two historical educational frameworks that explore educational content presentation are the Revised Bloom’s Taxonomy and the Theory of Multiple Intelligences. Bloom’s Taxonomy (BT) posits that there exists a hierarchy of objectives, that certain teaching material can be more effective in teaching, and that there can exist an optimal order for the presentation of content that maximizes student understanding. The revised version modernizes its vocabulary to emphasize the dynamic nature of learning and the cognitive processes required. ‘Bloom’s revised Taxonomy’ is commonly visualized, like many hierarchical systems, as a sectioned pyramid, shown in the right of Fig. 1. This interpretation ranks the desired outcomes of teaching, with more surface-level or basic understanding types representing conceptual understanding making up the foundation. These support the higher-level learning outcomes which require more expertise and familiarity [18, 19]. The higher-level outcomes, in turn, are similar to those expected from theoretical treatments in later lessons or advanced courses, including advanced textbooks as discussed in Section 1.1A

This system is contrasted by another framework that seeks to address long-standing questions about the impact of individual learner diversity on teaching effectiveness: the Multiple Intelligences model (MIM). Shown on the left of Fig. 1 the MIM does away with the hierarchy of prerequisites and rigid structure which are the foundation of BT, instead suggesting that there exist a variety of learning ‘styles’ that are effective for particular learn-
ers depending on the individuals predispositions or skills. These styles are wide-ranging in their origin: they represent both external (interpersonal/social/verbal) and internal (logical/intrapersonal/bodily) interactions and can create complex systems of learning preferences in individuals. The MIM resists definite classification within our content type framework, and suggests educational content should be tailored in a very different way than that prescribed by BT. It states that individuals will learn better if the content is tailored to their learning style, rather than when a rigid structure is imposed that presents information in the same logical progression for all learners. In addition to its applications for understanding student learning in physics, it has influenced later theories that can help us make predictions about content order effectiveness in physics instruction.

2. Mental Models, Scaffolding, and Sequencing

While the BT and MIM frameworks of student learning break up teaching content into distinct categories based on depth of information or mode of learning, many recent frameworks instead focus on how certain modes of instruction build up a particular idea or view within the student, introducing the idea of a ‘mental model’. These mental model-based frameworks forgo individual knowledge types for a more central theme based on holistic understanding. Many of these theories propose opening learning environments so that students are allowed to interact with facets of the topic that may not be touched upon with more prescribed, lecture-style instruction, combining the advantages of both flexible and hierarchical models. For example, Schwartz and colleagues’ work in education research focused on activities that promoted exploration, so that students developed an understanding of the greater structure of the topic through self-discovery. This ‘greater structure’ echoes more conceptual treatments of physics, where the big picture must first be established before being bolstered by theoretical justification and applications. Additionally, their work stresses that having a sound grasp on a topic requires not only content knowledge, but also the ability to adapt and frame new information within previously learned schemes. This treatment of learning as a more organic process leads to different teaching goals. Instead of expecting certain presented content to be absorbed by learners, mental-model based instruction seeks to guide students through the topic so that they discover the important details and become familiar with how facets of the topic are connected. This allows for new information to be incorporated such that a network of new and previous knowledge is created, not just added to a growing pile of isolated facts.

These goals both support the view that learning is not strictly about accumulating information until one has enough to understand the topic, but about creating a framework within the learners mind, which can actively accept, process, evaluate, and adapt to new information or challenges. This view is similar to other works prioritizing student mental models. These mental frameworks, especially within physics, can consist of previous instruction, the individual’s lived experience or other factors, and they influence the student’s interaction with new material meant to teach or advance their knowledge in that topic; this is very much in line with our definition...
of conceptual understanding. It is noted that if a model is already present (and one that is usually due to previous exposure, even if brief), it is difficult to change if incorrect and must be explained in a way that avoids grafting correct content onto an incorrect interpretation instead of correct model adoption. This connects to theories of conceptual change, a topics which has been heavily studied and has multiple competing theories within the field [23–25]. These theories posit different ways that the conceptual model is created and advanced within the student as they learn [26]. These interpretations are interesting and present new ways of thinking about the process of learning, but are also loosely defined and still debated. When we relate the conceptual change interpretation to previous theories of learning, we realize teachers must carefully design teaching method so that students can form the correct mental model and build upon it. Using conceptual change to form a correct mental model therefore requires a more prescribed teaching sequence like what is posited in BT. That being said, this more individualized view also feeds back into the notion of an individual learning style: because each learner has unique previous experiences incorporated into their mental model, their interaction with new teaching material will be different, like in the MIM. This interpretation also stresses the notion of change within the student going from naivety or an incorrect view to a sound mental model which can be built upon. This occurrence is extremely important for effective teaching, and thus has been a topic of study in itself that we wish to incorporate into the design and analysis of our study of content sequencing.

The frameworks of scaffolding and sequencing are also closely related to mental models and conceptual change. Scaffolding refers to setting up learning such that students have sufficient prior knowledge and correct mental model formation to learn complex topics and complete evaluation tasks [27]. Sequencing is also formalized in a broad literature [28, 29], which suggest that the sequence in which students are presented information will have an impact on their learning quality. Podolefsky et al. have conducted studies in electromagnetism classes that have found that scaffolding by teaching students new topics using analogies improves student learning outcomes [30, 31]. This type of analogical scaffolding relates closely to our definition of the concept content type, as it links student knowledge to information that they are already comfortable with, such as scenarios from the natural world, and uses diagrammatic representations to cement new ideas. Results from this literature thus motivate the use of content sequences that begin by providing students with a strong conceptual base, which we intend to further investigate and describe with our sequencing tests and models.

FIG. 2. “Content Cube” of learning. Each axis represents one of the content types in our model (Example, Concept, or Theory). Traversing a set of three edges from the origin to the far vertex corresponds to learning via a given content order. Each outward face of the cube represents the available “content-space” after seeing the first content type (e.g. the top face would be populated by students who were presented with Concept first).

3. Content Cube

Building upon the previously discussed literature, we have developed a framework of student learning that incorporates our content type division and illustrates the process of its order of presentation on student learning. To visualize our content type division and its order of presentation, we have created a representation of the three content types and their application. Our framework takes on a three-dimensional form which we call the ‘Content Cube’, where each of the three spatial axes represent the presentation of a certain content type. The different content presentation sequences manifest as paths along the edges of the cube from the origin to the (1,1,1) corner, shown as colored arrows in Fig. 2. Each path represents the students full journey from minimal prior knowledge (origin), through all three content types presentations (edge traversal), and to the desired learning outcome (far corner). This places each content order as a unique path in space and thus allows for each order to be represented uniquely, instead of relying on overlapping populations to describe the results of different content orders. It can also help visualize populations at different stages in their journey to the learning outcome: for example, the outer faces of the cube represent populations of students that saw the same content type first (here the top face would represent seeing Concept first, etc). This division of the total student population is useful in evaluating the impact of particular content types and type ordering, as will be shown in the results sec-
tion. This population division and subdivision can assist in visualizing the time-dependent changes that can occur for students throughout the learning process, which would be much less clear in the previous models. The stratified view of all the possible paths to the learning outcome also lends itself well to evaluating the effectiveness of these paths and to exposing subtleties due to other environmental factors. This will create a clear but nuanced and flexible framework for examining the effects of specific content sequences and other potential variables on student learning outcome.

With the ability to express and evaluate teaching methods not only by what they present, but also in their order, we can start to infer the best course of action for planning a lesson that attempts to maximize student learning. Using the three content types previously defined, is there an obvious choice for the optimal content order? Based on the results from the survey and how others have used the types in previous works like textbooks, it seems like starting with concepts is a generally accepted strategy, since this is the first material presented in introductory works. Examples, on the other hand, provide students with a means of self-assessment, mirrored in their placement at the end of chapters after conceptual and theoretical arguments have been established as well as their use in examinations. Thus, it is a fair guess to assume that a content order like CTE would perform better than TEC, since this presentation aligns with how a student naturally progresses through learning. TEC, on the other hand, effectively tries to deepen understanding before introducing it, and checks for facility before completing the teaching process. While this order may have some unseen merits, from a simple thought-process framework it would seem to produce a less effective path for student learning than one which follows the student progression as well as what has been generally adopted by the field. These inferences are the basis for the more specific questions that will be investigated using the results of our student study.

### III. METHODS

#### A. Survey

To validate our educational content classification, and also to compare the content and order preferences of the introductory course students to those more familiar with physics topics, we issued an online survey within the local STEM academic community. Participants were asked to self-identify their field (e.g. physics, chemistry) and role (e.g. undergraduate student, faculty).

111 total participants consented to sharing their survey results for research purposes; we first asked them to assess the effectiveness of our content groupings, and then the effectiveness of the content in teaching. For the first 9 content samples, participants were asked to label content type between Example, Concept, and Theory. For the following 9 content samples, participants were asked to rate the accuracy of a given label, and the contents effectiveness in teaching the topic contained, both on a scale of 1 (Not Accurate/Effective) to 5 (Very Accurate/Effective). Three of those samples were mislabeled, serving as a control for response bias. The participants were lastly asked if they have a preferred content type between Example, Concept, and Theory, if there is an order for presenting educational material they think is best between all permutations of Examples, Concept, and Theory, and for any comments and concerns about the content organization.

This survey allowed for differences in the classification and preference to be quantified across multiple disciplines as well as level of expertise with the topic and with teaching. The participants within physics were contacted via departmental email, and further information about other disciplines was gained through distributing the survey to the undergraduates who had taken the course, who have concentrations across the sciences. Additional responses were solicited through social media posts, attracting some responses from outside the McGill community, but a majority (> 95%) came from within the faculties of science and engineering of McGill University.

#### B. Experimental assignment

Once the goals of the study were identified, an assignment was designed to test them and collect student information. This was done through an online system where the students filled out five evaluations: a pre-test, three intermediate tests (quiz 1 to 3), and a post-test. In Fig. [Supplemental Material we show a screen shot of McLEAP to illustrating one intermediate test (quiz 1). An intermediate test has between 3 and 4 questions.

The full methodological format is shown in Fig. [5] Two experimental assignments were given in two subsequent years (2017 and 2018), yielding 4 datasets with 426, 437, 346 and 399 participants respectively. There were only 2 students that participated in experiments from both years. The assignment was designed to introduce a topic within the scope of the course that had not been previously taught (in this case reflection and refraction, which normally are taught toward the end of an introductory Electricity and Magnetism course). The pre-test is given directly after the instructions and consent form, and is meant to test the students prior knowledge as well as familiarize them with the assignment format. After the pre-test, the students fill out a short questionnaire, asking them what type of content they would prefer to see as well as how prepared they feel going into the next assessment. After the first questionnaire, the students are shown a randomized piece of content from one of the
Internal reflection (type C)
The figure shows rays of monochromatic light from a point source \( S \) in glass incident on the interface between the glass and air. For ray \( a \), which is perpendicular to the interface, part of the light reflects at the interface and the rest travels through it with no change in direction.

For rays \( b \) through \( d \), which have progressively larger angles of incidence at the interface, there are also both reflection and refraction at the interface. As the angle of incidence increases, the angle of refraction increases; for ray \( d \) it is 90º, which means that the refracted ray points directly along the interface. The angle of incidence giving this situation is called the critical angle \( \theta_c \). For angles of incidence larger than \( \theta_c \), such as for rays \( e \) and \( f \), there is no refracted ray and all the light is reflected; this effect is called total internal reflection.

Total internal reflection of light from a point source \( S \) in glass occurs for all angles of incidence greater than the critical angle \( \theta_c \). At the critical angle, the refracted ray points along the air-glass interface.

Internal reflection (type E)

Calculate the index of refraction of the material in the figure, assuming this is the smallest angle for which total internal reflection occurs.

Internal reflection (type T)

Internal reflection occurs when passing from a medium of higher refractive index \( (n_2) \) to lower \( (n_1) \). When the angle is greater than the critical angle, given by

\[
\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)
\]

all the light is reflected rather than transmitted, and we have total internal reflection.

Solution: The angle of incidence appears to be approximately 50º. The critical angle should be

\[
\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)
\]

so that

\[
n_1 = \frac{n_2}{\sin(\theta_c)} = \frac{1.0}{\sin(50^\circ)} = 1.31
\]
three categories (theory, concept, or example), asked to solve two to four problems regarding the material they were just introduced to, and then fill out a similar questionnaire after to see if their preference or preparedness have changed. This was repeated twice more to cover all three types of content, then a longer (five questions in 2017 or ten in 2018) post-test was given. Student learning outcome was defined by their performance in quantitative problem-solving tasks as used in previous studies [33]. Upon completion of the entire assignment, a final survey was given to gather information on student preferences, including which learning methods (for example use of resources) they found helpful.

The web interface used for this study allowed for the random selection of test questions and random content type pages, and could also generate content or questions based on previous answers. It tracked all of the answers as well as their timing. Students were required to complete one part of McLEAP in five hours, within a 5 day period. McLEAP consisted of two parts, each part testing different course content. The online structure of part 1 and 2 were the same. The test questions (which are automatically graded and give immediate feedback to the students) are standard assignment questions (see Fig. 11) and the content pages are typical lecture content pages based on the three different content types (theory, example, and concept), which are illustrated in Fig. 1. No timing differences were observed with the different content types. However, students who started close to the end of the five day period did worse than the average. The entire McLEAP platform was written in PHP (a website programming language) and is stored on a McGill server. McLEAP stores all the answers in an encrypted data file (DATA) that can only be accessed by the PI (Hille). From DATA we use a decoding program which generates two different files: FILE A, which contains the student IDs and their assignment grade. FILE B, which extracts the data from the students who have answered yes to the consent question, and assigns a random number to the remaining students and includes all the data saved by McLEAP. This ensures that analysis on FILE B can be done while keeping the participants anonymous.

IV. RESULTS

The results section is divided into three parts: First, we analyzed the student assignment results. Second, we considered the influence of content type order on student learning outcome. The main quantity of interest is the relative deviation from the mean grade of a subpopulation for a given test, which is defined as

\[
G_{\text{subgroup}} - \langle G \rangle
\]

\[
\langle G \rangle
\]

(1)

where \( G \) is the mean grade of the test overall and \( G_{\text{subgroup}} \) is the mean grade of a subgroup. Finally, we will discuss results from the student questionnaires during the assignment and from the independent survey on content type preference.

A. Example optimizes students learning outcome

Fig. 5 depicts the performance of the student population per content type per test relative to the mean grade per test, averaged over all data (2017-2018). The whole population is distributed across the three content types with the constraint that every individual is subject to a particular ECT ordering; in other words, they cannot receive the same content type twice. We note that Example first enhances Test 1 results by \((+2.2\% \pm 2.2)\) relative to the overall mean. Similarly, Theory first decreases Test 1 results with \((-2.8\% \pm 2.1)\). This means that if students are limited to material from only one content type, it is most beneficial for their learning outcome to provide them with an example, while giving them only theory is the least beneficial. In terms of the content cube model, in the Example-Theory plane before the Concept direction has been touched, we conclude that the Example will result in a higher learning outcome as Theory. This suggests that with an Example content type, despite not necessarily understanding the fundamentals, students are able to solve exercises best by merely repeating the procedure outlined in provided Examples. We note here that only one particular example was provided. The Theory plane provides the weakest learning outcome, corresponding to the idea that students without any prior knowledge will get confused by the Theory content type, unaware of how it supports the general concept or how to apply the knowledge of the Theory to problem-solving exercises.

B. Theory reinforces Example/Concept content type

Students who were assigned Theory for Test 2 (sequence CTE and ECT) had seen Concept or Example first. If we compare them to students assigned sequence CET and ECT, who have also seen Concept and Example first, but are given not-Theory second, we see that the grade relative to the mean grade for the group with Theory second \((+3\% \pm 2)\) is higher as the for the group with not-Theory second \((0\% \pm 2)\). This difference is not quite a significant effect due to the large standard deviation in the quiz scores. Yet, it shows that students with Theory second perform significantly better as with Theory first, and moreover, it indicates that Theory reinforces the Concept/Example content type and positively aids the learning outcome this way.

The net positive Theory second effect is mainly due to the ETC content order, as can be seen in Fig. 13 with
FIG. 5. Relative change from the mean grade grouped per test and content seen (Data 2017-2018). Example first (left orange bar) (+2% ± 2) significantly enhances test results over theory first (left green bar) (−3% ± 2). Results for students who see Theory at test 2 (center green bar) (+3% ± 2) are also significantly enhanced over students who see Theory at test 1.

Test 1 and Test 2 outcomes of (+4%±4) and (+4%±4.3), but the standard errors are such that we cannot be conclusive, despite the high number of students. The large standard error is due to only two (2017) or three (2018) questions being asked for each intermediate test, giving rise to a wide distribution with large standard deviation. Only from the post-test results can we ultimately conclude how content type ordering affects learning outcome. This is because students get to see all of the content types again, normalizing the population and removing memory effects (students may benefit while still having Example in mind), and also because the number of questions increases to five (2017) and ten (2018), allowing for more precise measurements of the students learning outcome.

C. Post-test: content order significantly influences learning outcome

Fig. 6 depicts the learning outcome (defined as the post-test score) for each content type order by the relative change from the mean grade. Starting with Concept leads to the best post-test grade, while starting with Theory leads to the worst learning outcome (Data 2017-2018). The difference in learning outcomes between different content type orders hints on the existence of hierarchical learning, which supports a mental model framework. Here, students have to adapt the correct mental model, which will only happen when the new mental model is intelligible, plausible, useful, and causes a large conflict with a previously established, potentially incorrect mental model [6, 16]. Such a conflict is instigated conceptually, and not by example or theory, whose implications can more easily be added as an extension to an incorrect mental model.

We can explain hierarchical learning in more detail as follows. The Concept content type introduces the named concepts such that students can place them within an existing mental framework of physical laws, which every human being has naturally built up over the course of their lifetime (ex: apples fall downwards, similarly charged objects repel one another, etc.). This is the vertical step in the Content Cube model. The Theory content type then adds structure to the concept that has been newly placed into an existing framework through a formal language (in this case mathematics). Eventually, the Example content type consolidates the learning by repeated and explicit usage of the new concept in various scenarios, providing exploration in the topic and stimulating memory. As it turns out, the order of Example and Theory can be reversed but the first content type is vital for effective learning. In other words, upon having reached the top of the Concept plane in the Content Cube, one can move around freely along the Example and Theory axes until eventually reaching the desired learning outcome at the far vertex (point of convergence of the three arrows in Fig. 2). Our data suggests that presenting Theory first brings confusion, from which students not fully recover (Fig. 6) resulting in the lowest learning outcome. We also find that presenting Example first gives the best immediate learning outcome (“cheap learning”), but this effect is short-lived, and for the post-test, students who
saw Example have average performance. Thus, Example and Theory first are inferior to Concept first as measured by post-test grades. We therefore conclude that first impressions carry significant weight, at least in presenting new physics material, and that the most efficient learning takes place in the Concept plane of the Content Cube model (Fig. 2).

To be certain that content type ordering is the true variable underlying the variation in learning outcome between students in Concept first and Theory first populations, we have considered various environmental factors that could confound our results. The four factors we chose to investigate are student preference for content types, whether or not they worked in groups, whether or not they used online resources, and pre-test results.

With an Ordinary Least Square Multivariate Regression (computed via Python's StatsModels library), we computed the correlation of the factors with the content order as well as post-test results. In Tables I and II in the Supplemental Material, we summarize the results for the combined data and for the 2018 data separately. While working in groups ($p < 1e-7$) and pre-test ($p < 1e-5$) results significantly affected the learning outcome, they did not correlate with the content order, as expected, because students were randomly assigned one of the six orders. Preferences did not correlate with post-test results ($p = 0.58$), but to our surprise significantly correlates with content order ($p = 0.05$). We hypothesize that the assigned order must have biased the students in choosing their preferred content type. We conclude that content type ordering is the true variable underlying the variation in the learning outcome in these initial investigations. For the remainder of the section, we will discuss how the ordering affected the students preparedness and preferences.

### D. Questionnaire and Survey Results: Preparedness

In the 2017 data and part I of 2018 data, we observed the impact of content order on learning outcome. We decided to add student preparedness and preference to the fourth experiment to understand how content order affects learning outcome in more detail. The results from the students regarding their preparedness show that students felt more prepared when they were presented with concept first, compared to example and theory first. We use the student self-reported preparedness as a marker for student confidence and self-efficacy.

As shown in Fig. 7, we find that the preparedness of the students in the concept first groups undergoes an overall increase over time, from 2.33 to 2.69 (on a scale of 1 to 5). In the example first group, the rating decreases from 2.54 to 2.33. For the theory first group, there is an increase from 2.80 to 2.85. The absolute self-rating of preparedness is lower for the concept first group. This is however the only group that significantly increases in preparedness over time. The students preparedness beginning with concept first agrees with the previously reported results, where the order that yields the greatest learning outcome is also concept first. Moreover, the results of this analysis on preparedness suggest that students are aware of what type of content, and which content order, is allowing them to acquire enough knowledge properly to understand the material at each step in the assignment. This will be reinforced by the results of the survey, discussed at the end of the Results section.

#### E. Content type preference over time

We believe that the change in content type preference over time may stem from the fact that during the assignment, the students become aware of how they are learning physical concepts and how well that particular strategy is working for them, and thus discover that they may benefit from a different content type or order to continue to further their knowledge acquisition. Fig. 8 shows this result through a plot of content type preference over time. A majority (73%) of the students changed their initial preference over time. This agrees with our Content Cube theory; students only attain a certain stage of understanding with each content type, and in order to reach the desired learning outcome they require a change of content type.

We find that the percentage of students who prefer Example rises from 42% before Test 1 to 66% after the post-test, the preference of Theory falls from 39% before Test 1 to 22% after post-test, and the preference of Concept falls from 19% before Test 1 to 12% after the post-test. As shown in Fig. 8 the Concept and Theory preferences mostly become Example preferences. We believe that this change in content type preference may demonstrate a form of conceptual change, and we plan to further investigate the student ability to remodel previous knowledge, as well as their awareness of how
FIG. 8. Details of the preference change by student. The student preferences Example (E), Concept (C) and Theory (T) are represented by the colors orange, blue and green, respectively. White is used to represent a lack of answer, while black is used for more than one preference. The top row is the initial pre-test preference, while the lower rows show subsequent tests. The lowest row represents the post-test preference distribution based on the initial pre-test student preference population. The right column shows the relative preference evolution over the course of the different tests. Overall, the preference evolves towards example for all sub-groups, regardless of initial preference. The initial concept preference population changes almost entirely to example or theory preference.

they learn (see conclusion for further details). Students’ change in content type preference over time (i.e. their preference of content order) is developmental; they are improving and building upon their previous conceptions throughout the assignment.

We also considered how students’ preferences relate to the optimal content order for their problem-solving performance, when probed at individual time points. We find that the percentage of students preferring to see the Example content type increased toward the end. The percentage of students preferring theory and concept decreased. There was however a large percentage of students who opted consistently to receive Example. Moreover, within the content order preference results of the survey in the following section, we come across an important distinction between what students and teachers believe to be an ideal learning order.

Additionally, the survey data collected concerning the participants preferences of both content type and content order show similar results. We observe that though the content type preference appears to be evenly distributed when considering the participant population as a whole, there exist discrepancies among the different classes of participants, shown in Fig. 9.

Notably, 53% of undergraduate participants prefer Example, and only 21% of graduate students (TAs) share this preference. Moreover, 0% of faculty members prefer Example. 60% of graduate students and 55% of faculty members prefer Concept. In other words, the type of content that teachers and teaching assistants believe yield their greatest learning outcome differs from that of the undergraduate students. This difference between student and teacher preferences is an important factor for the study of educational content type and sequencing, as

F. Content order preference varies across academic population

We justify our partitioning of physics content through a survey asking others in the sciences if the labels of concept, example, and theory accurately describe content excerpts used in teaching electromagnetism. We find that the majority of participants, from undergraduates to faculty members across multiple disciplines in STEM, agree with the labels we placed on our educational content (see Fig. 13). Examples are particularly clear, since they pose a question and a solution. Theory and concept are the most likely to be confused (see Fig. 14); we believe the inclusion of mathematical formulae in theory often acts as a marker that can make it distinctive as well.

FIG. 9. Survey results for preferred content type used for self learning for undergraduate student, graduate students, faculty members and overall survey responders.
this discrepancy could have a direct impact on the fit of teaching materials for optimal student learning outcome.

The results of the survey pertaining to content order preference, shown in Fig. 10, support our primary findings about content effectiveness and order. A majority of participants preferred to be presented conceptual material first when confronted with new topics. We believe that the Concept is preferred as the first content type because it bridges the gap between knowledge and experience that the students may already have with the more specific physics that is described within a theory or an example [35]. Also, the TCE content order was preferred by 27% of undergraduates, 12% of graduate students and 11% of faculty members, in contrast with the results of the students best learning outcome. This preference may be due to the fact that the material can still be understood well when starting with Theory, given the mathematical description, but that sufficient computational practice has not been provided when seeing Example last. This observation uncovers a limitation in our method, as we are only able to test the students problem solving ability, a variable that factors into learning outcome, but may not be exactly the same and could be measured in a variety of ways [15, 36].

Similarly to the content type preferences shown in Fig. 9B, an important result presented within Fig. 10 is the discrepancy between content order preferences for self-learning expressed by faculty members and students. Notably, 51% of undergraduate students and 52% of graduate students prefer the content order CTE, the order which yields the greatest overall learning outcome in our study (Fig. 6). This suggests that a majority of the participating students can appropriately gauge how they learn best. However, of the faculty member participants, 33% preferred the content order CET, and 22% had no preference. Given that our data suggests that there is a content order that yields the greatest learning outcome, these differences in content order preferences indicate that instructors may not always be translating their knowledge of introductory physics concepts through a content order that yields the greatest overall learning outcome for students [37]. This in turn has potential to influence the effectiveness of teaching and the quality of learning within university courses.

To this end, the preference results of the survey further suggest that there exists a hierarchical model that best suits learning needs, and that it is majoritarily preferred. Additional survey results pertaining to the justifications of our choice of categorization of learning materials are available in the Supplemental Material (Figs. 15, 17).

V. CONCLUSIONS

The results of our two-year study with close to 1000 students completing the McLEAP assignment show that sequencing of content types had a significant impact on student learning outcome. We find that regardless of student sequencing preference, there is an optimal structuring of material presentation that led to better problem-solving performance. To maximize learning outcome, the optimal sequence of content presentation is Concept first. Our Content Cube model of student learning progression depicts this outcome and proposes a structured theory of the impact of sequencing on learning outcome, formalizing and building upon the mental model learning framework. Course instructors and students may benefit from this learning model, as our results suggest there are unifying aspects in the way students interact with new material, existing alongside heterogeneity in individual learning preferences. Additionally, we find that receiving theory first leads to the lowest learning outcome. Despite timescale differences, the learning process is comparable, so this can be compared to learning presented in textbooks. These results reflect that content type sequencing seen in textbooks is important at the stage where mental models still need to be formed, as introductory textbooks are concept-heavy in the beginning of chapters, and are heavy on examples at the end of chapters. More advanced treatments assume that readers have a strong conceptual foundation to build upon (are in the right mental model) and start directly with theory without expanding on examples too much.

While we find a significant difference between Concept first and Theory first, our results do not differentiate between the learning outcome from sequences CTE and CET with statistical significance. CTE results in slightly higher learning outcomes below statistical significance. Interestingly, student’s preference in the survey is also biased towards CTE, which could be a reflection of their experience in many introductory courses and text-
books, making this an interesting subject for further exploration. More detailed studies might show one of these content sequences to be more significant. However, the inherent student variability might also suggest adapting the sequence to individual or groups of students based on their respective optimal learning framework. Further studies may be able to segment the population and find which learning method works for different students.

These findings are supported by conclusions from our content type survey; within survey participants from STEM academia, a majority recognize the content type distinctions we put forth, and prefer to learn with the content sequence that matches our most effective order of content type survey; within survey participants from STEM-wide survey also suggest that there exist important discrepancies in preferences of content types and content sequences among undergraduate students, graduate students, and professors. Sharing this result with course instruction teams, including professors and teaching assistants, can be an important step to improving instructors’ awareness of how their students learn, leading to the application of learning tools and strategies optimized for student learning.

ACKNOWLEDGMENTS

We would like to thank Chris Roderick, Janette Barrington, Marcy Slapcoff, Véronique Brulé, and Rebecca Brosseau for useful discussion, editing, and literature recommendations. We would like to thank Zezhou Liu for programming support on the online learning platform. We acknowledge support from the Tomlinson Chair in University Science Teaching at McGill.

[1] A. Thambyah, European Journal of Engineering Education 36, 35 (2011). https://doi.org/10.1080/03043797.2010.528559.
[2] E. Brewe, J. E. Bartley, M. C. Riedel, V. Sawtelle, T. Salo, E. R. Boening, E. I. Bravo, R. Odean, A. Nazareth, K. L. Bottenhorn, R. W. Laird, M. T. Sutherland, S. M. Pruden, and A. R. Laird, in ICT 5, 10 (2018).
[3] R. J. Sternberg, The Educational Forum 59, 47 (1995). https://doi.org/10.1080/00131729409336362.
[4] B. G. Davis, Tools for teaching (John Wiley & Sons, 2009).
[5] B. F. Jones, A. S. Palincsar, D. S. Ogle, and E. G. Carr, Strategic teaching and learning: Cognitive instruction in the content areas. (ERIC, 1987).
[6] E. F. Redish, American Journal of Physics 62, 796 (1994).
[7] N. Council, D. Education, B. Assessment, B. Education, and C. Education, Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics (National Academies Press, 2011).
[8] L. W. Anderson, D. R. Krathwohl, and B. S. Bloom, A taxonomy for learning, teaching, and assessing: a revision of Bloom’s taxonomy of educational objectives: complete edition (Longman, 2001).
[9] S. Meleod, "Kolb’s learning styles and experiential learning cycle." (2017).
[10] H. Gardner, Frames of mind: the theory of multiple intelligences (Basic Books, 2011).
[11] P. G. Hewitt, Conceptual physics (Pearson Educación, 2002).
[12] J. D. Jackson, Classical electrodynamics (AAPT, 1999).
[13] Z. Popovic and B. Popovic, Introductory electromagnetics (Prentice Hall, 2000).
[14] L. O. Wilson, Five Basic Types of Questions (2016).
[15] R. J. Shavelson, Journal of educational psychology 63, 225 (1972).
[16] G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog, Science Education 66 (1982).
Special Topics-Physics Education Research 11, 010107 (2015).

[35] B. A. Brown and K. Ryoo, Journal of Research in Science Teaching 45, 529 (2008).

[36] J. Blackmore, Studies in Higher Education 34, 857 (2009).

[37] L. Halim and S. M. M. Meerah, Research in Science & Technological Education 20, 215 (2002).
VI. SUPPLEMENTARY FIGURES

This section contains Supplementary figures. Fig. 11 (quiz question) and Fig. 12 (order of problems and questions in an assignment) show more details on McLeap, Figs. 13 and 14 show test results per content order and dataset, and Figs. 15 - 17 show more details on the survey results.

Quiz 1

This quiz will count for 10% of your total assignment grade.

1) The glass core of an optical fiber (n=1.6) is surrounded by a plastic cladding (n=1.4). What is the maximum angle light can make with the wall of the core to avoid exiting the core? Keep 1 digit after decimal point. Your answer should be formatted like: 0.0
Answer: 

2) The palisade cells in a plant’s leaves make use of total internal reflection to maximize the chances that light passes through their chloroplasts. Here we treat the cells as the rectangular prism shown in the diagram (ignoring the 3rd dimension). If the interior of the cell is mostly water (n=1.33), what is the largest angle a light ray can have from the normal of the upper surface before it risks exiting the cell through the sides? Keep 1 digit after decimal point. Your answer should be formatted like: 0.0

Answer: 

FIG. 11. Student view of McLEAP showing Test 1.
FIG. 12. Pages diagram of the McLEAP assignment interface; shows student progression through the assignment with the order of presentation for both problems and questions. Each arrow represents a new webpage seen by the students.

FIG. 13. Relative change from the mean grade of the intermediate tests grouped per content sequence and test (Data 2017-2018) showing a detailed breakdown of Fig. 5 in the main text. Standard deviations are high, reflecting both variability in learning and the small number of questions per test (two in 2017, three in 2018).
FIG. 14. Relative change from the mean grade of the post-test grouped per content sequence and dataset (Data 2017-2018) showing a detailed breakdown of Fig. 6 in the main text. Standard deviations are higher in Data 2017, because only 5 instead of 10 questions were asked in the posttest.

FIG. 15. Response to the instruction *Choose the description you think best describes the educational content shown*, which provides the accuracy of different levels of people in physics at labelling content. Overall, a majority of people accurately label each content type with the same categorization that we assume in this experiment, though identifying an example is clearer than the distinction between concept and theory.
FIG. 16. Response to the question *How accurate does the label 'Content type' fit for the content above?* where Content type was either C,T or E, and the accuracy was measured on a scale from 1 to 5, 1 being very inaccurate and 5 being very accurate. This data yields the rating of the content labelling by different levels of people in physics. Overall positive results for each label, distinctly worse rating for the falsely labelled content other than theory, which is harder to distinguish from concept.

FIG. 17. Response to the question *How would you rate the quality of the content above in explaining a physical concept?* This data provides us with the rating of the content type at describing the actual content by the different classes of participants. Overall concept has the highest rating, and each content type has a relatively positive rating.
VII. SUPPLEMENTARY TABLES

This section contains Supplemental Tables on determining confounding variables in 2018 data (Table I) and 2017 data (Table II).

**TABLE I.** Results from multivariate regression on content order and posttest results for the 2018 Data. Preference is significantly correlated with content order, but does not affect the learning outcome, while pre-test, groups, and online (significantly) correlate with learning outcome, but not with content order. Every variable has been standardized with mean 0, standard deviation 1, so that the magnitude of coefficient can be compared across factors.

|                     | Content order | Posttest |
|---------------------|---------------|----------|
|                     | coefficient   | p-value  | coefficient | p-value |
| pre-test            | 0.01          | 0.71     | 0.17        | 0.00    |
| groups              | 0.06          | 0.13     | 0.16        | 0.00    |
| online              | 0.04          | 0.30     | -0.06       | 0.10    |
| preference          | 0.07          | 0.05     | 0.02        | 0.58    |

**TABLE II.** Results from multivariate regression on content order and posttest results for all data. pre-test is the only factor that is consistently measured, and it is not correlated with content order, while it correlates significantly with learning outcome.

|                     | Content order | Posttest |
|---------------------|---------------|----------|
|                     | coefficient   | p-value  | coefficient | p-value |
| pre-test            | -0.02         | 0.40     | 0.27        | 0.00    |