PROCEEDINGS OF THE
2023 AAU TEACHING SHOWCASE
VOL. 26

ASSOCIATION OF ATLANTIC UNIVERSITIES
ASSOCIATION DES UNIVERSITÉS DE L'ATLANTIQUE

Evidence of Effectiveness of Mastery Tasks in Introductory Physics Courses

Geoffrey Lee-Dadswell, Cape Breton University
Fraser Turner, Mount Allison University

keywords: formative assessment, feedback, scaffolding

Abstract
We briefly describe a one-on-one, iterative, formative assessment technique which we call "mastery tasks". We examine the effect that completion of a mastery task has on student performance on exam and assignment questions in introductory physics courses. Data collected across five courses show that mastery tasks have a strong effect on student performance on midterm exams and assignments ($p < 0.0001$), but that the effect is much less significant on final exam performance.

Introduction

A standard weakness of many forms of assessment is that they do not have an effective feedback loop (Sadler, 1989). Students often do not read feedback when assessments are handed back to them, and even if they do read it, they have little opportunity to act upon the feedback (Mensink & King, 2020; Sinclair & Cleland, 2007). Many traditional assessments, such as weekly problem sets, essays that are handed in with no subsequent requirement for revision, and especially exams, do not give students any encouragement to act upon feedback. In some thinking about formative assessment design there has been a focus on the nature, frequency, and quality of feedback given, as well as on the context that the feedback occurs within and how this can direct students to reflect upon and act on feedback (Black & Wiliam, 1998). Some researchers have proposed frameworks of criteria to guide the design of formative assessments with effective feedback (Gibbs & Simpson, 2005; Hattie & Timperley, 2007).

In this work we describe an assessment technique which we call "mastery tasks". Mastery tasks are designed around a very fast feedback loop. In the process of completing a mastery task, the student receives feedback immediately after they complete the task, and they are asked to act upon that feedback very quickly following the feedback being given. The feedback is provided one-on-one in a low-stakes conversation with the instructor. As such, the feedback is rooted "in the vitality and authenticity of the relationships between teachers and students," (Torrance, 2012 p. 339) which may improve its impact on student learning.

What Are Mastery Tasks?

Mastery tasks are an iterative, one-on-one, formative assessment. As used in our courses, they take place outside of class time, but this is not fundamental to their design and one could imagine using them during class time, especially with very small classes.

Each task targets a key skill in the course which students must do repeatedly, usually as part of a larger, complex task. In physics, good examples of such skills are drawing free body diagrams, drawing energy bar charts, and deciding whether a described system is isolated. Examples in other courses could include such things as writing correctly formatted quotations and references, carrying out a specific form of textual analysis on a short passage, examining a graph to identify which of a number of scenarios is depicted, etc. If the instructor can identify a core skill that a student can demonstrate in a few minutes then it might be appropriate to design a mastery task around that skill.

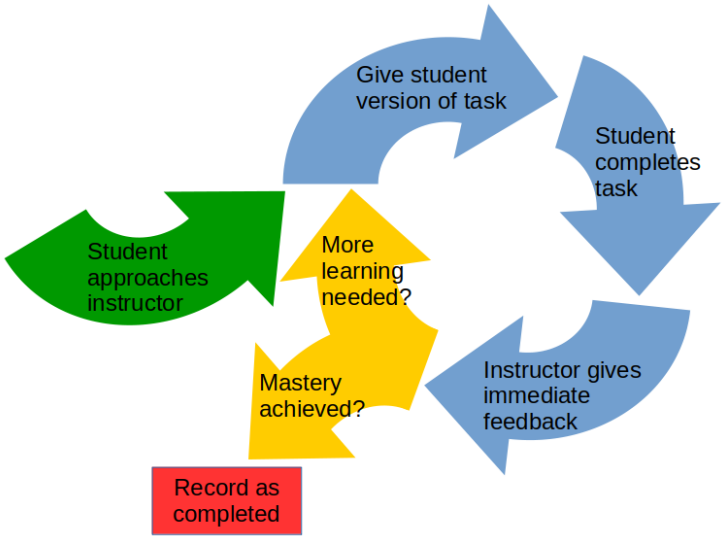

Figure 1: A flowchart of how a mastery task is administered.

To complete a mastery task, the student approaches the instructor and asks for the task. The instructor gives the student a version of the task, and the student sits down nearby to complete it immediately. Completion by the student should take no more than a few minutes. The student then brings the task back to the instructor, and the instructor gives immediate feedback, hopefully in less than a minute. If the instructor is satisfied that the student has demonstrated "mastery" or "fluency" on the task, then they mark it as complete and the student receives a completion mark. If the instructor is not satisfied then they give the student a new version of the task and the process is repeated. This is repeated as many times as it takes for the student to demonstrate satisfactory mastery or fluency. The mark recorded is simply for completing the task, and does not depend on how many attempts the student needed to complete the task. If the instructor feels that the student needs to do more learning on their own before reattempting the task, this can also be an option. So, while mastery tasks are designed around the idea of a rapid feedback loop, they do not always need to lead to such a fast loop.

Mastery tasks could play many roles within a course. In our courses they provide a bridge to help students transition from in-class activities, to assignments. In-class activities are usually relatively simple, and students have the support of their peers as they work on them. Assignments are complex, and are to be completed individually. Mastery tasks provide a safe, low-stakes way for students to verify that they have developed skills which they will need as a foundation for completing the work on assignments. An ideal mastery task would place the student in their zone of proximal development (Chaiklin, 2003), and would give them support from the instructor to expand the zone of development. Clearly, not every mastery task will achieve this with every student. But upon completion of a task, the student and the instructor should both have confidence that the student is prepared to apply the skill assessed on the mastery task in more sophisticated contexts.

Study Design

The effectiveness of mastery tasks was examined in five introductory physics courses at Cape Breton University. These course were:

- PHYS 1101 - algebra-based, first-semester introductory course
- PHYS 1104 - calculus-based first-semester introductory course
- PHYS 1204 - calculus-based second-semester introductory course focused on electricity and magnetism

Two offerings of PHYS 1101 were examined, in fall 2021 and fall 2022 which will be called Alg F2021 and Alg F2022. Two offerings of PHYS 1104 were examined in fall 2021 and fall 2022, which will be called Calc F2021 and Calc F2022. One offering of PHYS 1204 was examined in winter 2022 which will be called E+M W2022.

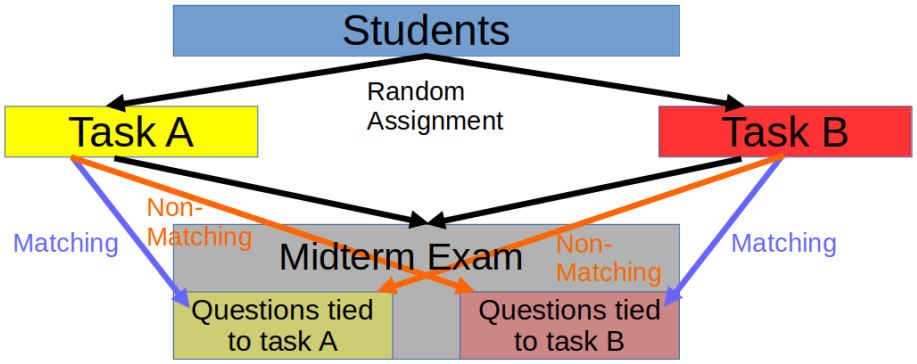

Figure 2: Flowchart of the study design. Students are randomly assigned one of task A or B which are associated with a topic unit. The midterm exam has questions directly connected with each of tasks A and B. Students who are assigned to A are the control group for students who are assigned to task B and vice versa.

In each of these courses, students were randomly assigned to two mastery tasks, A and B, associated with a single topic unit. A midterm exam occurred approximately 2 weeks after the deadlines for these mastery tasks. This midterm contained questions directly related to mastery task A and other questions directly related to mastery task B. We say that students who did mastery task A are "matching" with respect to questions related to A, and are "non-matching" with respect to questions related to B. Thus, students assigned to A are the control group for students assigned to B, and vice versa.

Each course had two midterm exams and this process was carried out with each. All students were required to complete 5 mastery tasks as part of the course. However, only 2 tasks completed by each student were included in the data for this study. For the remaining 3 tasks, students were given choices of which tasks they wished to complete rather than random assignment, and so these additional tasks were not suitable for inclusion in the study.

All of these courses also had approximately weekly problem assignments. In the case of Alg F2021 some questions from assignments were included along with midterm exam questions. This was not done for the other courses.

Finally, on the final exams there were questions related to each of the mastery tasks. This allowed us to examine the persistence of any effect associated with completion of mastery

tasks. The midterm exams occur only about 2 weeks after the mastery tasks, but the final exams occur over 2 months after some of the mastery tasks. So we expected to see a much smaller effect. In the fall 2021 semester, an outbreak of COVID-19 forced Cape Breton University to switch to online exams. The questions on these online exams were not suitable for inclusion in this study. So, we do not have final exam data for Alg F2021 or Calc F2021.

Results

For every exam or assignment question included in the study, we rescale its grade to a mark out of 5 for easy comparison. For every student, each question is either "matching" or "non-matching". We can, thus, compile averages across an entire course of how well students performed on matching questions and compare those with averages of how well students performed on non-matching questions. Data is discarded for students who did not complete either of the randomly assigned mastery tasks, or who did not write the relevant midterm exam or assignment. We make the null hypothesis that mastery tasks caused no improvement in performance on exam and assignment questions.

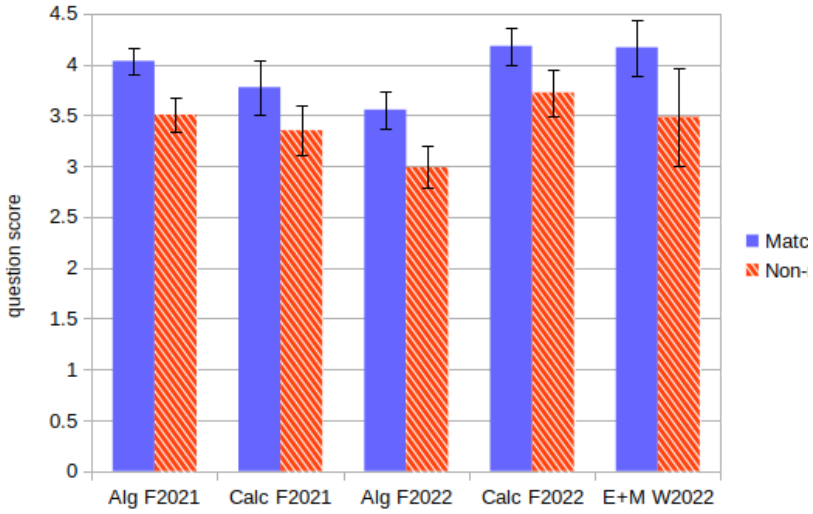

Figure 3: Comparison of student performance on matching and non-matching questions on midterm exams and assignments in each of the courses included in the study. Error bars show standard deviations.

The results of this analysis on a course-by-course basis, for midterm exam and assignment questions is shown in Figure 3. We see an improvement associated with mastery task completion in all cases. The statistical significance of these improvements are summarized in the Table 1. We have calculated student's t-value in each case. The only case where there is a discernible difference between Z and t is E+M W2022 because of the relatively small number of questions included in that data. We also show the compiled statistics for all five courses. We are able to reject the null hypothesis for Alg F2021 and Alg F2022 at $p < 0.01$ and $p < 0.05$ respectively. The significance is not as great for the other courses. However, for the combined data of all five courses we are able to reject the null hypothesis at $p < 0.0001$.

Course	Alg F2021	Calc F2021	Alg F2022	Calc F2022	E+M W2022	Combined
t	2.4	1.2	2.1	1.6	1.1	3.8
p	0.008	0.13	0.02	0.06	0.15	0.0001

Table 1

Perhaps as expected, the observed effect of mastery tasks on final exam performance is considerably weaker. Reasons for this to be an expected outcome will be discussed below. For the three courses where we have final exam performance data the value of t in the combined data across all courses is 1.3 ($p = 0.18$). So, we cannot reject the null hypothesis with confidence on the basis of the final exam data.

Discussion

We can conclude that completion of mastery tasks had a very significant effect on student performance on assignments and midterm exams. It is worth noting that all five courses included in this study used a flipped classroom methodology with active learning methods with a wide variety of formative assessments carried out during the pre-class readings, and during class. Additionally, the courses are highly focused on weekly problem assignments with additional formative assessment part way through completion of the assignments. So, the mastery tasks occur within a "sea" of other assessments which might be expected to make contributions to student learning. In this context it is striking that mastery tasks have such a strong, measurable effect on student performance.

The effect of mastery tasks on final exam performance is much less clear. We argue that this is to be expected given the large time which elapses between the mastery tasks and the final exam. Nevertheless it is interesting to speculate about the cause of this reduced effect. Presumably it is a combination of students forgetting what they have learned on the mastery tasks during the intervening time ("decay"), and the effect of students who did not complete the relevant mastery task learning the same skill via other means ("catch-up"). It is difficult to see how we could determine how much of the reduced effect is due to (undesirable) decay and how much of it is due to (desirable) catch-up. Any examination of this question would require much more intensive research methods such as directly interviewing students.

We have used mastery tasks in courses with approximately 30-60 students. In courses of this size, administration of mastery tasks is a large, but manageable workload for the instructor. For larger courses it is likely not feasible unless graduate teaching assistants are available to take on a significant portion of the workload. However, we caution that training of the teaching assistants would be crucial in making this work.

The focus on this work has been on the relatively easily measurable and quantifiable effect of mastery tasks on student exam performance. In other words, our attention has been on how mastery tasks affect the learning of content related skills. It is at least as interesting to consider the effects of mastery tasks on less easily quantified factors such as the development of more

general academic skills, and learning in the affective domain (Krathwohl et al., 1973). In particular, one might expect the use of mastery tasks in a course to contribute towards the development of communication skills among students, especially oral communication skills. They could also help students to develop a willingness to try new things, be wrong, and try again. This improvement in confidence and resilience, if it occurs, would be valuable and is worth studying. As interesting as the effects that mastery tasks have on students are the effects that they may have on the instructor. Anecdotally, we have found that mastery tasks have allowed us to develop a greater appreciation of the difficulties that students have with course material and with presumed prior knowledge. They are also valuable for developing an instructor's confidence with one-on-one assessment, perhaps with an eye to moving towards oral exams or other more individualized assessment. Finally, mastery tasks probably serve to humanize students to the instructor, and the instructor to the students. Research into any of the above is likely to be more challenging than this current work, but could lead to valuable insights about this and similar assessment methods.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. *Vygotsky's Educational Theory in Cultural Context*, *1*(2), 39–64.

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *1*, 3–31.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1973). Taxonomy of educational objectives, the classification of educational goals. Handbook II: affective domain. David McKay Co. *Inc., New York*, *1*, 956.

Mensink, P. J., & King, K. (2020). Student access of online feedback is modified by the availability of assessment marks, gender and academic performance. *British Journal of Educational Technology*, *51*(1), 10–22.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sinclair, H. K., & Cleland, J. A. (2007). Undergraduate medical students: Who seeks formative feedback? *Medical Education*, *41*(6), 580–582.

Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*, *38*(3), 323–342. https://doi.org/10.1080/03054985.20

PROCEEDINGS OF THE
2023 AAU TEACHING SHOWCASE
VOL. 26

ASSOCIATION OF ATLANTIC UNIVERSITIES
ASSOCIATION DES UNIVERSITÉS DE L'ATLANTIQUE