21

An experiment in guided learning of ethics for medical students

Isaac Wagner

Department of Public Health
University of Copenhagen

Background and Goals

I conducted my teaching experiment in the course 'Medical Ethics and Philosophy' in Spring 2022. This course is for 4th semester medicine students, and consists of large lectures followed by small-group work in 2-hour SAUs ('Studenter Aktiverende Undervisning'). I taught three SAUs for the course (there are 12 in all), and I conducted my experiment in each of these SAUs.

One of the biggest challenges the medicine students face in this course is lack of familiarity: most of them have never taken a course in ethics, nor a course that requires them to argue for their opinions in a rigorous philosophical way. This gives rise to uncertainty and anxiety among them about whether they truly understand the material and will be able to handle the exam at the end of the course. Thus, **Goal 1** of my teaching experiment was *to help students effectively prepare for the course exam.*

Of course, it is possible to help students prepare for an exam in a way that does not promote their long-term learning of course material (e.g., by focusing on reviewing technical aspects of the exam, focusing on general 'exam tips', or the like). I want all in-class activities that I spend significant time on to help my students *actually learn course material* (whatever else those activities may accomplish). Thus, **Goal 2** of my experiment was *to promote students' long-term learning of course material*. **Goal 2** is related to **Goal 1**: promoting student's learning of course material also helps to prepare them for the exam. Thus, achieving **Goal 1** would contribute to achieving **Goal 2**.

In the next section, I describe the experiment and draw on literature to support the expectation that it would help to achieve these two goals.

Experiment and Justification

My experiment concerned exam-preparation, more specifically preparation for the *essay* portion of the exam for the course. Essay questions on the exam are open-ended questions designed to test students' understanding of fundamental theories and concepts from the course as well as their ability to synthesize those theories and concepts into a thorough, clear answer to the question. A typical essay question presents students with a moral dilemma in medical practice and asks them to carefully argue for a resolution to the dilemma using course material.

Historically, students have not received much official help in preparing for the essay portion of the exam. The most the course leader has done is share with students essay questions from previous exams, *without providing any model answers*. While this is better than nothing for test preparation, it is not particularly effective for that purpose. It is essentially a form of what is called "discovery learning": students are given a problem and asked to discover a solution to it entirely by themselves. Discovery learning is akin to trying to teach someone to swim by throwing them out into the deep end of the pool. Discovery learning has been contrasted with "guided learning", where the instructor *actively assists* students in arriving at an adequate answer (while still challenging them). Research strongly suggests that guided learning is superior to discovery learning in achieving positive learning outcomes (Mayer, 2004). Thus, my teaching experiment aimed to test out a particular form of guided learning in the present context: *in-class collaborating with students to construct model answers to previous exam questions.*

Let me describe the experiment. In a given SAU (20 students), students were first formed into small groups (3-4). An essay question from a previous exam was then displayed for all students using powerpoint, and they were asked to discuss in their groups for 5 minutes how best to approach answering that question (e.g., what theories or principles or concepts would be most relevant to include in an answer). After that, the groups were broken apart, and students were given 15 minutes to individually attempt to write an answer to the question. I asked students to focus on producing an *outline* of an answer to the question, in which they lay out how they would

structure an answer. (They were encouraged to fill in this outline with specific details as much as they could, with any time they had remaining). I had students focus on writing an outline of an answer (rather than attempting a complete answer in full detail) because (1) 15 minutes is not sufficient for writing a satisfactory complete answer to an essay question in this course, and (2) given this, the best use of the limited time is for students to think primarily about how to best structure their answer, since good structure is the foundation of a good essay answer.

The rationale behind including a 5-minute group discussion prior to the 15-minute individual writing period was to assist students who would have had difficulty thinking up a good approach to answering the question entirely on their own. Instead of having these students struggle to come up with ideas by themselves for the entire experiment, they could—via the opening 5-minute 'brainstorming' small-group discussion—simply 'borrow' an idea or two from the group discussion in order to get themselves going. From there they would be in a position to construct an outline on their own, during the 15-minute writing period. Thus, this initial 5-minute group discussion increased the chance that all students could fully participate in the experiment.

Once the 15 minutes were up and each student had constructed an outline of an answer to the practice essay question, it was time for the entire class to collectively construct a single model answer to the question, based on their individual writing. I facilitated this process by taking suggestions from students and, in collaboration with them, meshing these suggestions into single coherent model answer. I typed the answer into a Word document that I projected onto the big screen, so that the students could observe, in real time, the construction of a proper answer to the essay question (based on their input). The model answer contained quite a bit of detail—it was more than just an outline. Thus, the process of constructing the answer with students involved not only soliciting their views about how best to structure an answer based on their individual work during the 15 minute writing period (in which they focused on answer-structure), but also filling in that structure with significant detail via further class discussion. The completion of the model answer marked the end of the experiment, and I posted the answer to Absalon for reference.

Let me now provide some support for my hypothesis that this experiment would fulfill the two goals I had for it. Above, I mentioned that my experiment constitutes a kind of guided learning. To better appreciate its merits in relation to my two goals, we should note that it is a particular *kind*

of guided learning, namely, a kind involving *active engagement* combined with *error feedback*.

Consider first active engagement. It has been established that students who are actively engaged in the classroom learn better than those who are not (Freeman et al., 2014). My experiment clearly required active engagement of the students. Students were asked to engage in the active task of creating—first individually and then collectively—an answer to a question, as opposed to passively listening to the teacher describe what a good answer looks like. More specifically, my experiment made use of a particular form of active engagement called "retrieval practice." During the exercise, students did not have time to constantly look up course material. They instead had to rely heavily on retrieving (recalling) what they had learned previously during the course in order to create an answer outline individually and then to participate in co-creation of the single model answer as a class. Having students strive to retrieve (recall) previously learned information has been shown to be highly effective for long-term learning of that material (Larsen et al., 2008). In short, through use of active learning (specifically retrieval practice), my experiment could be expected to increase student learning, i.e., **Goal 2**. Moreover, the students actively engaged with course material *in a way that was directly relevant for the exam* (as we were working with a previous exam question), hence the experiment could be expected to fulfill **Goal 1**, that of preparing students for the exam.

Now consider error feedback. Error feedback involves students receiving high-quality, quick feedback on their work, and it has been shown to promote learning significantly (Hattie & Timperley, 2007), especially when paired with active engagement (Roediger & Butler, 2011). My experiment made use of error feedback. As I constructed the model answer using student inputs, I would regularly correct mistakes that students made. For instance, a student might incorrectly or imprecisely describe a course concept or would suggest a structure for the model answer that would not be especially beneficial; I would gently correct these errors for the class (while still trying to incorporate some insightful aspect of the student's input). Thus, various students received feedback on their specific ideas that was both quick and of high quality (i.e. from the teacher). This could be expected to promote the learning of the particular student who directly received the feedback, as well as other students who had the same or similar (incorrect) thoughts as that particular student had. For this reason, my experiment could again be expected to fulfill the goal of promoting student learning of the course material, i.e., **Goal 2**.

As noted in the previous section, promoting students' learning of course material (**Goal 2**) serves at the same time to help students prepare for the exam (**Goal 1**). Thus, all the evidence above suggesting that my experiment will fulfill **Goal 2** is thereby also evidence that it will help to fulfill **Goal 1** as well.

In summary, evidence suggests that my experiment could reasonably be expected to fulfill both goals I had for it.

Results

The ideal assessment of this experiment would involve seeing if the exam performance of my SAU students who received the model answer was better than the performance of my SAU students from previous semesters of the course who did not receive a model answer. However, such a comparison was not possible. Student exams in the course are anonymous, so when I grade the exams, I cannot tell whether a given exam was written by one of my SAU students or by a student with a different SAU teacher. Thus, I was not in a position to compare the exam performance of my current students with that of my past students.

Instead, I had to rely on student feedback. First, I asked students about their impressions of the experiment immediately after we completed it. They responded (verbally) that they found the experiment helpful. They said that it was useful to see the formulation of an answer in real time, by someone who knows what a good answer looks like. According to them, seeing so clearly how to construct a good essay answer "from the ground up" made the essay-portion of the exam seem less daunting to them and something they could handle. They also said that the experiment was an effective test of their understanding of key course material and thus was useful for their reviewing and learning of the material.

I had also planned to get systematic, written student feedback on the experiment on the final day of the course, as part of a general student assessment of my teaching. Unfortunately, I was ill on the final course day and could not hold any of my three SAUs. Thus, I was unable to get this written feedback.

Did my experiment fulfill the two goals I had for it? That students said the exercise put them more 'at ease' regarding the exam suggests that it helped them with exam preparation - **Goal 1** - at least in the sense of making them feel more optimistic about the exam (which can translate into

improved performance). On the other hand, regarding helping students to learn course material - **Goal 2** - very little can be concluded from this experiment. That students found the experiment useful for their reviewing and learning of course material does not show that it actually helped them learn anything, given that students (like anyone else) regularly suffer from 'illusions of learning' whereby they take themselves to have learned things they did not in fact learn (Bjork et al., 2013). And, as noted earlier, ideal data—i.e., the performance of these particular students on the course exam compared with the performance of my previous students—was inaccessible to me. All this leaves me virtually in the dark about whether **Goal 2** was achieved.

Let me conclude this section by noting a different dimension of the 'results' of this experiment, namely, my own *perception* of how the experiment went. It is interesting that students' verbal feedback on the experiment was positive. This is because I did not *feel* comfortable conducting the experiment in any of my three SAUs. First, the experiment involved giving up a lot of control to students. It was fundamentally student inputs—not teacher inputs—that the model answer was generated from. I have never before given up so much control over my teaching to students, and this made me uneasy. Second, this experiment required me to think "on my feet" in a way I had not done before. I had to quickly process student inputs and formulate them into precise writing in a Word document while all the students watched. While I excel at leading spontaneous group discussions, the extra element of having to formulate spontaneous thoughts *in writing* for the students made me quite uncomfortable. These feelings of discomfort with the experiment led me to immediately conclude that it must have failed to achieve anything of value for the students (despite their claims to the contrary).

Concluding Thoughts

I learned a few important things from conducting this teaching experiment. I learned, first, that teaching in a way that involves giving up significant control to students makes me uncomfortable. This is valuable to know. If I am aware that a particular teaching exercise challenges me, then I can take action to try to rise to the challenge (if the exercise is worth doing) or be sure to avoid that exercise in the future (if it is not worth doing). Second, I learned that there can be a mismatch between how *I* feel when

conducting a class exercise and how the *students* feel about that exercise. As I said above, students seemed to like the experiment, even though I found it uncomfortable and likely pointless. Clearly, it is possible for students to benefit from an exercise even if the teacher feels as though they did not. Given this, it is important as a teacher not to let one's own subjective assessment of a teaching activity dominate one's overall assessment of that activity.

I plan to try some version of this exercise again, for two reasons. First, students responded positively to it, and exercises that students enjoy tend to be worth further exploration and development. Second, there is good evidence that the exercise's core elements—active engagement with error feedback—are pedagogically powerful, and thus that the exercise is effective for learning (though, sadly, I was unable to add to this evidence through my own experiment). However, I might conduct the exercise slightly differently in the future. For instance, when beginning to construct the model answer with students, I might start by asking students to use their ideas to generate a word cloud (instead of immediately launching into a group discussion with them, as I did this time around). This would allow each student to see the inputs of *all* other students and thus appreciate the broad range of ideas available for answering the question. Moreover, it would give me the opportunity to select from this broad range of ideas those particular ideas that I would be most comfortable basing a model answer on, thus hopefully reducing my level of discomfort with the exercise. Finally, a word cloud would help me to see *patterns* in students' inputs, so that I would have the option of constructing a model answer that, e.g., included elements that a majority of students offered as input.

References

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual review of psychology*, *64*, 417–444.

Bovill, C. (2020). Co-creation in learning and teaching: The case for a whole-class approach in higher education. *Higher Education*, *79*(6), 1023–1037.

Dehaene, S. (2020). *How we learn: The new science of education and the brain*. Penguin UK.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences 111*, (23), 8410–8415.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Rev. Educ. Res*, *77*, 81–112.

Larsen, D. P., Butler, A. C., & Roediger, H. L., III. (2008). Test-enhanced learning in medical education. *Medical education*, *42*(10), 959–966.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American psychologist*, *59*(1), 14.

Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in longterm retention. *Trends in cognitive sciences*, *15*(1), 20–27.