Using the Immediate Feedback Assessment Technique (IFAT) for non-assessments: Student perceptions and performance

Lynne N. Kennette & Dawn McGuckin

This paper examines an experiment conducted at a 2-year college with non-majors. The Immediate Feedback Assessment Technique (IFAT) is a cross between a lottery scratch ticket and a scantron form which was designed to replace scantrons (EpsteinEducation.com). Using the IFAT in an unconventional way (reviewing class material), we expected to find significant differences in students' final grades with the IFAT group out-performing the control group (who reviewed class material by re-reading their notes). Our hypothesis was confirmed. Additionally, formal student feedback on the use of this weekly activity was overwhelmingly positive, suggesting the activity is seen by students as a worthwhile undertaking.
Keywords: *immediate feedback, assessment, retrieval, collaboration*

CRAIK & LOCKHART first proposed their levels of processing framework in 1972. Since then, it has been used to examine memory performance in many contexts including the classroom. When students engage with course material in a way that is more involved than simply reading (or re-reading) the content, they are engaging in deep processing. Deeper levels of processing (which are, by definition, not passive) have been shown to make it easier to transfer information to long-term memory and makes that information remain more easily accessible for later retrieval (Baddeley, 1997; Craik & Lockhart, 1972). Without deeper linguistic processing, Ferreira, Bailey, and Ferraro (2002) have proposed that we engage in Good Enough processing, whereby we only process language as much as is needed for the task at hand. In a classic study by Hyde and Jenkins (1969), students were asked to either check for the letter 'e' in the word or rate how pleasant the word was (shallow and deep processing, respectively) in a list of words. Some were also told they would be asked to recall these words later, while others were not. They found that deeper processing led to better recall of the word list compared to shallow processing (whether they knew about the recall task and therefore were actually trying to learn the words or not did not significantly affect performance). So, when the task is for students to read their textbook or their notes, this likely means processing the text just enough to be able to understand the word meanings (i.e. to read) and results in them engaging in shallow processing; consequently, they do not make any connections between concepts or synthesise the information, which would lead to better recall in the future. Deeper processing has also been shown to allow us to more easily identify nuances that we may have otherwise missed (Kennette et al., 2010), which, in the classroom, may also help to explain the benefit to students' comprehension and/or recall.

Advantages for student learning or retention have been demonstrated for a number of classroom techniques such as collaborating with others, immediate and adequate feedback on performance, and retrieving information from memory.

Recognition memory for words and pictures benefited from collaboration in a study by Rajaram and Pereira-Pasarin (2007). In one condition, participants were able to discuss their answer in a group of 3 before responding individually to a recognition test; in the other condition, there

was no opportunity for group discussion. Those who discussed their answers outperformed the non-collaboration condition. This benefit was found when tested after 1 hour, 48 hours, and 1 week, demonstrating the long-term memory benefits to collaboration. Another paradigm which uses collaboration is the two-stage collaborative testing (e.g. Cortright et al., 2003). Here, students first take the test alone, and then take the test a second time while collaborating with a classmate. The grades on the individual and collaborative test are then combined to produce the student's grade. Cortright et al. (2003) demonstrated better retention after four weeks on the material that was tested collaboratively compared to the material that was tested individually. Bowman et al. (2013) argued that the benefit of many forms of peer-to-peer collaboration is that students can better bridge the knowledge gap which exists between the expert instructor and the novice learner. That is, students are uniquely placed (in terms of their knowledge networks) to explain concepts to their peers in a way that they can understand.

Immediate feedback also appears to benefit students' long-term learning. For example, Dihoff et al. (2004) administered questions with a number of feedback delays: feedback could be received after 24 hours, at the end of the entire test, or after each question (or no feedback for the control condition). They found that feedback after each question was linked to improved retention on the final exam (this finding held even when question was re-worded) and that no feedback at all resulted in the worst retention.

In an innovative study demonstrating the powerful effect of retrieval on memory performance, Roediger and Karpicke (2006) asked participants to read a short passage about a topic (e.g. the sun) and then answer questions to assess how much of the information they remembered. They combined 'study' blocks (reading the passage over and over until 7 minutes had elapsed) and 'test' blocks (writing down everything they could remember from the passage) across two experiments. In the first experiment, they compared performance in a Study-Study condition to the Study-Test condition. Students who studied during both blocks did out-perform the Study-Test condition during immediate testing (5 minute delay), but the Study-Test group remembered significantly more after 2-day and 1-week delays. In a second experiment, they included Study-Study-Study-Study (SSSS), Study-Study-Study-Test (SSST), and Study-Test-Test-Test (STTT) groups. Important to note here is that participants read the passage many more times in the first two conditions (on average 14, and 10 times, respectively) compared to the STTT which averaged 3 readings of the text. The number of times they had read the passage predicted performance when tested immediately (5 minute delay; SSSS had the best performance, followed by SSST, and STTT showed the worst performance). However, for delayed recall (1 week), the number of tests predicted performance (STTT had the best recall, followed by SSST, and then SSSS). So, the benefits of retrieval on memory may not be immediately detectable, but the effect appears to be robust. More recently, Adesope, Trevisan, and Sundararajan (2017), conducted a meta-analysis with 272 effect sizes which confirmed the importance of retrieval (i.e. repeated testing) in long-term memory retention. They specifically concluded that repeated testing protects against forgetting and improves performance much more than repeated rehearsal.

Hypotheses

Given the ample research in the fields of cognitive science and education (e.g. Baddeley, 1997; Craik & Lockhart, 1972) demonstrating that (1) retrieving information from memory is one of the best ways to learn it; (2) having the opportunity to discuss and collaborate while deciding on a correct answer leads to greater long-term retention; and (3) students benefit from knowing what the correct answer is immediately; it stands to

reason that the Immediate Feedback Assessment Technique (IFAT; EpsteinEducation.com) could be a valuable tool to use collaboratively in the classroom for non-assessment activities. In the case of this study, students would be working in small groups to review previously presented materials (retrieval and collaboration). Through this IFAT activity, they would receive immediate feedback and be able to further collaborate to determine the correct answer. This should lead to the best student learning outcomes. Of specific interest were the following research questions:

1. Do students enjoy engaging in the IFAT collaborative review activity?
2. Do students perceive that the IFAT activity helps them learn and/or remember the information covered in the course?
3. Do student grades demonstrate a benefit to using the IFAT to review course material?

Methods

Participants

Participants were students enrolled in 2 sections of an introductory psychology course. In the section serving as the control condition, 17 students elected to participate in the research study. In the experimental section, 34 students volunteered for the study and 24 filled out the survey at the end of the semester. No other demographics were collected as there were no related hypotheses.

Procedure

Students in both the control and experimental groups were engaged in their normal classroom activities for most of the semester and both had the same tests, assignments, and were taught by the same instructor. Both groups reviewed the previous week's content at the beginning of every class, but did so differently: the experimental group used the collaborative IFAT activity to do so (described below), while the control group reviewed their notes individually. At the end of the semester, the students who had been experiencing the IFAT activity were invited to complete an anonymous online survey about their experience.

Materials

Immediate Feedback Assessment Technique. The IFAT is a cross between a scantron form and a scratch lottery ticket; it has response boxes (like a scantron) covered with a metallic substance which is scratched off for students to indicate their response to a question (like a lottery ticket). Developed by Epstein Education to use in lieu of scantron forms during tests, it provides students with immediate feedback on their performance by indicating whether their response is correct (a star appears when the correct box is scratched). The IFAT also allows students to earn partial credit for obtaining the correct response on a subsequently scratched square.

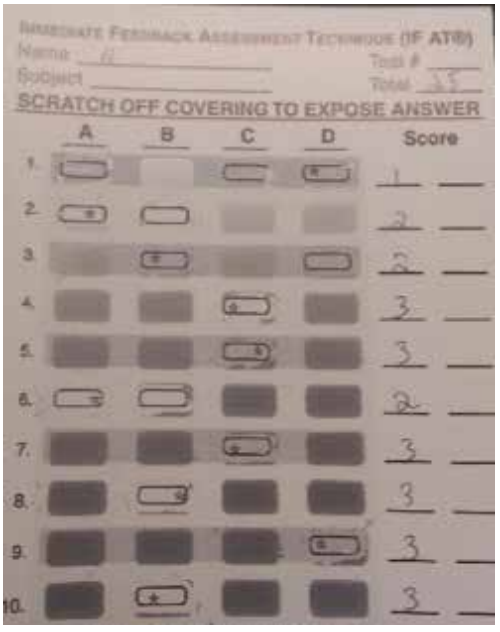

Figure 1: The Immediate Feedback Assessment Technique (already scratched)

The IFAT review activity asked students to form small groups and answer 10 multiple choice questions related to the previous week's content. These questions were similar

to the questions you might find on a test in an introductory psychology course.

Survey. Students in the experimental group also had the option of participating in an online survey used to gage their personal experience with the IFAT and the perceived effect that the tool had on their learning and their grades. The online survey was available to students during that final two weeks of the semester. The 12-question survey included 11 Likert-rated questions (5-point scale, ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*)) such as 'I feel the scratch ticket activity helped me to better learn the material we covered in class' and 'I would like it if this type of activity occurred in more of my courses' as well as one open-ended question, 'Do you have anything to add about the scratch card review activity?' to allow students to share their experiences if they were not elicited from the other questions.

Results and discussion

Outliers were removed (-3SD) and we confirmed that no assumptions of normality were violated. Although we were primarily interested in final grade outcomes, we compared each assignment and each test across the two conditions (IFAT and control) and only two comparisons were significant: Final grades ($t(19)$, = 2.44, p <.05, (d = .80) and Assignment 2 ($t(19)$, = 2.28, p <.05, d = .68). Final grades showed the most obvious difference (control group 66.18% (23.96); experimental group 81% (11.02)). It is important to note here that participants in the control condition would likely also have been retrieving content (which Adesope et al., 2017, showed was key to long-term retention), but it is impossible to determine exactly how students in the control condition engaged with their notes during the classroom review time (e.g. whether they were passively reading their notes or more actively engaged).

Our measure of learning, which was the three repeated questions from the first two tests on the final test (this measured learning in contrast to the more immediate measure of retention which tests typically tap into), did not provide us with a clear finding. Because each question was from a different topic area, we compared the two conditions for each question separately. The question asking about dependent variables showed a significant difference in the expected direction with the IFAT group out-performing the control group ($t(48)$, = 2.40, p <.05), and the question asking about classical conditioning was marginally significant ($t(48)$, = 1.50, (p = .07), but the question about operant conditioning was not significant ($t(48)$, = .43, (p = .34). However, there was no significant difference overall (with performance on these three questions combined) between the control and experimental conditions ($t(48)$, = .19, p = .42). So, it may be the case that the IFAT activity helped students retain some of the more difficult topics, but it is challenging to make a strong statement to that effect given that there were very few data points to consider and this, paired with a small sample size, would make it difficult to detect any real differences. There is reason to believe this possibility, however, as there are many empirical results where students of various abilities are differentially affected by an intervention. For example, Agarwal et al. (2017) found that students with lower working memory capacity showed significantly greater benefits to their performance (after a two-day delay) following a retrieval practice intervention than students with a higher working memory capacity. Therefore, it stands to reason that students in our study could be demonstrating greater benefits from the IFAT activity for the more difficult topics in the course and show no significant differences from the control condition for less complex topics compared to the control condition. In fact, the research on topic difficulty doesn't appear to break down broad, overarching chapter topics (e.g. learning) into its sub-topics (e.g. classical conditioning, operant conditioning) when reporting their ranking results, so it is difficult to determine the relative difficulty of these three ques-

tions (see McNamara et al., 2011; Peck & Ali 2006).

Given the small sample size, the authors were concerned that the control group may not have been a representative sample. To this end, data from previous semesters were also analysed, serving as an additional control condition. An ANOVA for final grades comparing these data to the two conditions reported above indicated a significant difference, which is not surprising, given the significant difference already reported between the control and experimental conditions ($F(2,297), = 5.15, p < .01$). Pairwise comparisons showed that the experimental group was also significantly different (in the expected direction) from the 'old' control data gathered from previous semesters ($t(71), = 2.65, p < .001, d = .65$).

The qualitative data gathered from the survey was quite rich. The majority of the students in the experimental group chose to complete this optional survey and the feedback was overwhelming positive. On the question, 'I feel the scratch ticket activity helped me to better learn the material we covered in class,' 83 per cent of respondents agreed with this statement. Showing that while there may not be strong statistical evidence in the test grades, students truly believed this activity was a beneficial one which had a positive effect on their grades. For the question, 'I would like it if this type of activity occurred in more of my courses,' 92 per cent of students answered affirmatively. We believe this is a strong testament to both the enjoyment that students derived from using this learning tool as well as the widespread applicability of this activity to a variety of courses. One participant's response to the open-ended question nicely summarises the feedback we received: 'I found that it really helped me understand the course content a lot better, it helped to discuss the answers which further helped me understand them and remember them […] Having the scratch ticket also made it a lot more fun to do and I know for me that if I find something I'm studying to be fun it helps me retain the information so much more!' This supports the claim discussed earlier whereby peers are in a unique position to explain concepts to their fellow students (Bowman et al., 2013). It also leads us to speculate that students in the IFAT condition may have been more motivated to attend class, and therefore earn more in-class participating activity points than their control counterparts. This would also provide them with more opportunities to be exposed to and discuss the final assignment content (Assignment 2), which may have lead to their significantly higher scores. Future research is needed to confirm this, as attendance was not recorded as part of this research project.

Taken together, these results suggest definite benefits to using the IFAT. The quantitative data do seem to suggest that final grades benefitted from the IFAT activity, and that the benefits on other assessments were primarily seen on Assignment 2, which could be driving this overall difference in final grades. Further probing is required, but since the assignment in question is a more global assessment of multiple topics during the semester (whereby students take pictures of psychology concepts in their everyday lives and explain the link between the picture and the concept), it is conceivable that students exposed to the IFAT encoded the course material more deeply and were therefore better able to make the required links and explain course concepts for this assignment. Future research should replicate these findings and investigate the alternative possibilities discussed here.

Acknowledgement

This research was funded by the Durham College Internal Research Fund awarded by the Office of Research Services, Innovation, and Entrepreneurship

Authors

Lynne N. Kennette & Dawn McGuckin

School of Interdisciplinary Studies, Durham College, Oshawa, Ontario, Canada

References

Adesope, O.O., Trevisan, D.A. & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*(3), 659–701.

Agarwal, P.K., Finley, J.R., Rose, N.S. & Roediger, H.L. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*(6), 764–771.

Baddeley, A.D. (1997). Human Memory: Theory and Practice. Hove, UK: Psychology Press.

Bowman, M., Frame, D.L. & Kennette, L.N. (2013). Enhancing teaching and learning: How cognitive research can help. *Journal on Excellence in College Teaching: Brain-Based Learning (Special Issue), 24*(3), 7–28.

Cortright, R.N., Collins, H.L., Rodenbaugh, D.W. & DiCarlo, S.E. (2003). Student retention of course content is improved by collaborative-group testing. *Advanced Physiological Education, 27*(3), 102–108.

Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Dihoff, R.E., Brosvic, G.M., Epstein, M.L. & Cook, M.J. (2004) Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback, *The Psychological Record, 54*(2), 207–231.

Ferreira, F., Bailey, K.G.D. & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11–15.

Hyde T.S. & Jenkins, J.J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology, 82*, 472–491.

Kennette, L.N., Wurm, L.H. & Van Havermaet, L.R. (2010) Change detection: The effects of linguistic focus, hierarchical word level and proficiency. *The Mental Lexicon, 5*(1), 47–86.

McNamara, C.L., Williamson, A.L. & Jorgensen, T.D. (2011). Assessment of differential learning by topic in introductory psychology. *Psychology Learning and Teaching, 10*(3), 253–260.

Peck, A.C. & Ali, R.S. (2006). Introductory psychology topics and student performance: Where's the challenge? *Teaching of Psychology, 33*(3), 167–170.

Rajaram, S. & Pereira-Pasarin, L.P. (2007). Collaboration can improve individual recognition memory: Evidence from immediate and delayed tests. *Psychonomic Bulletin & Review, 14*(1), 95–100.

Roediger, H.L. & Karpicke, J.D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255.