

BRIEF

Determining Whether Pharmacy Students Can Accurately Recall Items on an Examination Using Their Collective Memory

Adam M. Persky, PhD, Kathryn A. Fuller, PharmD

University of North Carolina at Chapel Hill, UNC Eshelman School of Pharmacy, Chapel Hill, North Carolina

Submitted March 15, 2021; accepted July 6, 2021; published March 2022.

Objective. To determine the extent to which students can recreate a recently completed examination from memory.

Methods. After two mid-term examinations, students were asked, as a class, to recreate recently completed examinations. Students were given 48 hours to recreate the examination, including details about the questions and answer choices. The results were compared to the original examination to determine the accuracy of students' recall and reproduction of content.

Results. The students were able to collectively recreate 90% of the questions on the two examinations. For the majority of questions (51%), students also recreated the question as well as the correct response and at least one incorrect response. The majority of questions that the students recreated were of medium to high accuracy as they contained detailed phrasing that aligned with the original question on the examination.

Conclusion. The collective memory of a group of students may allow them to accurately recreate the majority of a completed examination from memory. Based on the findings of this study and tenets of social psychology, faculty should consider the potential implications for examination security, whether to provide feedback to students on examinations, and whether completed examinations should be released to students following the examination.

Keywords: collective memory, examination, recall, academic dishonesty

INTRODUCTION

When it comes to releasing a completed examination to students, faculty generally fall into one of three categories: those that release the examination, those that do not release the examination but offer reviews/provide feedback, and those that do not release the examination and provide no review. Often, students want to see their completed examination so they can learn from their mistakes and focus their efforts on clear gaps in their knowledge, consistent with the *discrepancy reduction model*.¹⁻³ Additionally, given the time and effort students put into preparing for an examination, they naturally want to know the results of this effort.^{4,5} As educators, faculty should provide feedback as evidence suggests it is an essential part of the learning process.⁶⁻⁸ However, faculty often protect their examinations and do not release them because of the time and resources required to construct quality examinations. Faculty believe that by not releasing completed

examinations they can safely reuse questions and ensure that future examinations are secure and learning is accurately evaluated. Faculty often see students discussing answers to examination questions or learn that students recreate the examination in their study groups. Their inherent curiosity combined with the investment they have made in their learning inclines students to seek answers. The question then arises whether completed, graded examinations that are not released to students as a form of feedback are truly safe from reproduction.

In theory, a group of students should be able to recreate a shared experience such as an examination, because of *collective memory*. The social psychology literature is replete with examples of collective memory, which is the phenomenon in which a group of people can recall information about an event better than any one individual from the group can.⁹⁻¹² Brown and colleagues asked college aged students to individually recall advertisements seen on local mass transit.¹³ As a whole, the 175 participants recalled close to 900 items and over 200 different advertisements. However, the average person could only recall five advertisements. If this phenomenon were applied to recalling examination questions and each student could

Corresponding Author: Adam M. Persky, University of North Carolina at Chapel Hill, UNC Eshelman School of Pharmacy, 325 Beard Hall, Chapel Hill, NC 27599. Tel: 919-966-9104. Email: apersky@unc.edu

recall five questions, then it would only take approximately 20 students to recall all items on a 100-question examination, assuming no students recall the same items. Additionally, due to a phenomenon called *collaborative facilitation*, recall of one item by one person who then shares that item with a group may cue recall of other items by other members of the group, allowing students to remember more items as they discuss the examination questions.¹²

As mentioned earlier, students compare their responses after an examination to determine the “correctness” of their responses and to communicate. Humans communicate to remember and share experiences. Harber and colleagues conducted a study where 33 students completed a field trip to a morgue.¹⁴ In three days, over 800 people had heard about the field trip. The spread of the news of the field trip was attributed to the compulsion people have to disclose traumatic experiences. In other words, people tend to share emotional events with others with little prompting. Examinations may represent an emotional event for students, and discussing details about the event may be a natural human response.¹⁵ Thus, if students can recall information from an examination, they may naturally want to share it.

To date, there is negligible research on the ability of students to recreate assessments. Anderson and McDaniel examined students’ memory of a quiz they had recently completed.¹⁶ Students were asked to write down questions from the quiz that they remembered, and based on the results, the authors concluded that students had poor memory of the quiz. While individual students in the study did not have strong memories about the content of the quiz, that does not mean that their collective memory of the entire quiz would not have been better.

The objective of this study was to test a faculty observation that students can recreate a recently completed examination from memory. By answering this question, faculty members’ views about releasing examinations to students may change, which in turn may persuade them to focus more on using assessment to promote learning than to evaluate learning.

METHODS

Participants were 125 first-year, first semester student-pharmacists enrolled in a core course within a Doctor of Pharmacy program. This course had three mid-term examinations, each of which assessed new content, and a cumulative final. The three examinations included 40 to 70 single-answer, multiple-choice questions, ranging in difficulty from lower level (knowledge/comprehension) to middle level (application/analysis) on Bloom’s Taxonomy.¹⁷ Each question had one correct response and two or three distractors. Examinations were administered using

ExamSoft. Because of the COVID-19 pandemic, examinations were virtually proctored to minimize academic dishonesty.¹⁷ Immediately after completion of the second and third midterm examinations, the instructor emailed the students a link to a single GoogleDoc with instructions for them to work as a group to recreate the examination. For the second examination, students were instructed to be as accurate as possible with content, order, and phrasing. For the third examination, the students were asked to collaborate on recreating the examination in GoogleDocs, but were given even more explicit instructions directing them to include both correct and incorrect answer choices. Students were given 48 hours after each examination to complete the group activity and offered extra credit (one extra percentage point on their final examination grade) for participation.

In previous years, course examinations were not released to students. Students only received their examination score and a report of their performance on question categories. That is, they did not have access to examination questions from previous years. In the study year, the same procedures were used: students only received their score and overall performance within question categories.

The primary outcome measures were the number of questions students were able to recreate, the number of questions recreated for which the correct answer was recalled, the number of questions recreated for which at least one distractor answer was correctly recalled, and the ordinal scoring of question specificity. For this latter metric, recreated questions that captured the gist of the target question were considered low specificity, recreated questions that contained more accurate phrasing of the target question were considered medium specificity, and recreated questions that captured specific or exact phrasing of the target question, parts of the case vignette, and/or other details were considered high specificity.

In the analysis, we included the item difficulty of the examination question and its discrimination index to assess whether question difficulty impacted students’ ability to recreate the examination. Based on the data distribution, we tested for differences between generated vs non-generated items using a Mann-Whitney U test. Statistical significance was set at $p < .05$. This was a retrospective study deemed exempt by the University of North Carolina Institutional Review Board.

RESULTS

Within 48 hours of completing two midterm examinations, the class collectively recreated 90% of the questions and the respective correct answers (103/115, 90% for examination 2 and 89% for examination 3). On average,

students individually recreated 51% of these questions with the correct answer and at least one distractor. However, the percentage of questions for which distractors were recreated varied between the two examinations, with 31% recreated on the second examination and 82% recreated on the third examination. Of the 51% of questions with the correct answer in combination with at least one distractor, 10% had one distractor, 12% had two distractors, and 29% had all three distractors recreated correctly.

Cumulatively, 8% of the questions recreated were deemed low specificity, meaning the student captured a minimal amount of the general content or gist of the question but no specifics; 38% were deemed medium specificity, meaning the student captured the stem of the question, and the remaining 54% were deemed high specificity, meaning the student provided details about the question. Examples of low and high specificity questions are presented in Table 2.

When examining item difficulty, discrimination index, and question recreation, no notable trends were identified (Table 1). The median difficulty for questions not recreated vs recreated was 93% vs 89% ($p = .13$); the median discrimination index for not recreated vs recreated was .14 and .17, respectively ($p = .21$).

DISCUSSION

The goal of this study was to assess students' collective memory of a recently completed examination by having them recreate the questions from memory. The results demonstrate that soon after completing an examination, students can collectively recall a very large portion of the material.

The students were not able to recreate the entire examination. Interestingly, the 10% of questions that were not recreated seemed to be less difficult or challenging questions. There are several explanations for this phenomenon: encoding variability, attentional failure, and/or recollection failure. Generally, recall of information is enhanced when there are multiple routes to the target information rather than a singular route.^{19,20} Encoding variability theory suggests that the more cues a person has

to recall information, the better the recall.¹⁹ Therefore, an easy question, which may have less cues, may not elicit a strong enough recall. The second theory is attentional failure, which proposes that questions that do not garner sufficient attention are less memorable. For example, if the question is easy, it may not warrant the attentional processing, and as such, may not be memorable. The final explanation is recollection failure. When people successfully recollect information, it is because that recollection is associated with retrieval of specific contextual details about an event. Recollection failure occurs when contextual details are lacking, in part by increased familiarity. For example, if a question was discussed in class, it may be familiar, while "old" and older items in memory may be harder to recall.^{21,22}

In this study, the examination was recalled as a collaborative group, asynchronously. If a group of people can divide a memory task among themselves, the burden of remembering is distributed as each person uses the input of others as a memory aid. This is referred to as *transactive memory*.⁹ For this reason, when several people come together as a group to work on a task or brainstorm, they can perform better than they could as individuals. However, this advantage is contextual. For instance, one student's input could serve as a cue for another student and promote recall of other facts, and this would be consistent with the idea of *collaborative facilitation*.⁹ However, the reverse could also be true as collaborative brainstorming and recall often lead to *collaborative inhibition*.⁹

We did note some differences between the two examinations that may reflect the specificity of instructions. For the third examination, more specific instructions were provided to students, asking them to recreate question distractors. This resulted in an increased number of distractors listed (32% vs 81% of questions). Therefore, the instructions provided to students for recreating their collective memory may influence the outcome. This is consistent with the literature on eyewitness testimony in that how questions are asked leads to varying degrees of answer specificity.²³⁻²⁵

One strength of this study is that it was conducted in an authentic classroom environment. Additionally, students

Table 1. Summary of Question Specificity and Question Psychometrics for Items Recreated From Memory by Pharmacy Students Who Had Recently Completed an Examination

Specificity	Questions, No. (%)	Mean Question Difficulty (SD)	Mean Question Discrimination Index (SD)
Low	8 (7%)	.82 (.22)	.2 (.11)
Medium	39 (35%)	.78 (.22)	.21 (.16)
High	55 (49%)	.81 (.21)	.18 (.14)
Not recreated	11 (10%)	.91 (.08)	.13 (.09)

Table 2. Questions That Were Recreated From Memory by Pharmacy Students Who Had Recently Completed an Examination

Specificity	Recreated Question	Original
Low	Anxiety and worry neural circuits <ul style="list-style-type: none"> • CSTC loop^a 	ZT presents with anxiety symptoms of anxious worry, apprehensive expectations, and catastrophic thinking. Based on the pathophysiologic theories of anxiety, a dysfunction in which circuit is associated with these specific symptoms? <ul style="list-style-type: none"> • Amygdala-centered circuit • Mesocortical dopamine pathway • Nigrostriatal dopamine pathway • Cortico-striato thalamo-cortical (CSTC) circuit
Medium	What lab test result would you find in someone with CKD? <ul style="list-style-type: none"> • Low phosphate • Low calcium^a • High 1,25 OH vitamin D • Low PTH 	In CKD patients with mineral bone disorders, which of the following would you expect to see when looking at the lab values below? <ul style="list-style-type: none"> • low blood calcium • low blood parathyroid • high blood 1,25 diOH vitamin D • low blood phosphate
High	A Patient with a PMH of HTN, afib, hypotension, gout what type of thromboembolism is at risk and why? <ul style="list-style-type: none"> • Venous because a fib • Arterial because HTN • Venous because of age and gender • Arterial because of immobility for 3 weeks post-surgery^a 	LT is 48 year old obese man who is undergoing orthopedic surgery. Post-operative recovery is expected to take 3 weeks. LT has PMH of hypertension, atrial fibrillation, hypothyroidism, and gout. LT is at risk for what type of thromboembolism and why? LT is at risk for... <ul style="list-style-type: none"> • an arterial thromboembolism due to the 3 weeks of immobility in the post-operative recover period • a venous thromboembolism due to the diagnosis of atrial fibrillation • an arterial thromboembolism given the diagnosis of hypertension • a venous thromboembolism given age and gender

^a Correct response.

were unaware they were going to be asked to reproduce the examinations and therefore did not prepare or plan ahead to memorize the questions. Since these were first year, first semester Doctor of Pharmacy students, they were unfamiliar with the instructors and their question writing style. Finally, examinations from previous years had not been released and no practice examinations had been administered; therefore, it is unlikely students had additional exposure to the examination questions. A potential limitation of this study is that it is unknown exactly how many students were involved in recreating the examination. If only a few students participated, the accuracy of the cohort's overall ability to recreate these questions may have been underestimated. A final potential limitation was that students with higher academic ability, such as those in a professional program, may have had a higher ability to retain information and thus recreate examinations more easily than those with lower academic abilities.

An argument could be made that recalling questions from an examination and later sharing them with other students falls within the realm of academic dishonesty in the same sense that unauthorized use or attempted use of material, information, notes, study

aids, devices or communication in the learning environment is considered academic dishonesty.²⁶ However, students simply recalling an experience does not constitute academic dishonesty. The real concern is what such students do with the recalled information. Some might argue if they generate an examination and share it with future classes, that might constitute academic dishonesty. However, if a student generated an examination from memory and months later tutored a lower classman by providing helpful tips to be successful in class, is this cheating or peer tutoring? We know that peer and near peer tutoring is a powerful learning strategy.²⁷⁻³⁰ Additionally, sharing of experiences should deepen learning based on the theory of *shared in-group attention*.³¹ This theory suggests adults are more likely to engage in elaborative processing of material and make linkages to the broader knowledge structure when that object or experience is believed to be co-attended with one's social group.³¹ Thus, the sharing of examination experiences should promote learning.

Assuming students can collectively reproduce examinations from memory and that their natural tendency is to share information, which also may facilitate learning, we offer the following recommendations. Academicians

should reduce the high stakes nature of examinations and/or minimize the transactive nature of education (ie, the expectation for students to reproduce what faculty teach). In doing so, we may minimize some students' inclination to cheat and, hopefully, promote learning.^{31,33} In the learning process, we need to remember feedback is important. As such, we recommend that learning and feedback be the priority for assessments. If the focus is using assessments to promote learning, then the issue of assessment security is minimized as the goal is to *promote* learning rather than *evaluate* learning. Anderson and McDaniel found that if quizzes are returned so students can study them, there is minimal need for feedback. However, if quizzes are not returned, corrective feedback is necessary.¹⁶ Thus, simply returning students' assessments would reduce the need to provide corrective feedback. If assessments are to be used for summative judgements of learning that impact student progression through a curriculum, there should be a larger programmatic assessment plan. Best-practices suggest that this larger programmatic assessment plan should focus less on examinations and more on the whole student.³⁴ However, when high-stakes assessments are necessary (eg, licensure), then best practices for examination development should be followed.³⁵

CONCLUSION

By understanding the concept of collective memory, instructors' refusal or reluctance to return completed examinations to students or provide feedback to them because of concerns about examination security may be challenged. The results of this study should not be surprising as many students already practice recalling examinations. However, the findings may further encourage the use of assessments to aid in the learning process by either providing examinations back to students or offering feedback.

REFERENCES

1. Metcalfe J. Metacognitive judgments and control of study. *Curr Dir Psychol Sci*. 2009;18(3):159-63.
2. Son LK, Metcalfe J. Metacognitive and control strategies in study-time allocation. *J Exp Psychol Learn Mem Cogn*. 2000;26(1):204-21.
3. Thiede KW, Dunlosky J. Toward a general model of self-regulated study: an analysis of selection of items for study and self-paced study time. *J Exp Psychol Learn Mem Cogn*. 1999;25(4):1024-37.
4. Hsee CK, Ruan B. The Pandora Effect: the power and peril of curiosity. *Psychol Sci*. 2016;27(5):659-66.
5. Mullaney KM, Carpenter SK, Grotenhuis C, Buriianek S. Waiting for feedback helps if you want to know the answer: the role of curiosity in the delay-of-feedback benefit. *Mem Cogn*. 2014;42(8):1273-84.
6. Boud D. Feedback: ensuring that it leads to enhanced learning. *Clin Teach*. 2015;12(1):3-7.
7. Harrison CJ, Konings KD, Schuwirth L, Wass V, Vleuten CPMvd. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Health Sci Educ*. 2015;20(1):229-45.
8. van de Ridder JMM, van de Ridder JMM, Peters CMM, et al. Framing of feedback impacts student's satisfaction, self-efficacy and performance. *Adv Health Sci Educ*. 2015;20(3):803-16.
9. Hirst W, Echterhoff G. Remembering in conversations: The social sharing and reshaping of memories. *Ann Rev Psychol*. 2012;63(1):55-79.
10. Hirst W, Manier D. Towards a psychology of collective memory. *Memory*. 2008;16(3):183-200.
11. Hirst W, Yamashiro JK, Coman A. Collective memory from a psychological perspective. *Trend Cogn Sci*. 2018;22(5):438-51.
12. Weldon MS, Bellinger KD. Collective memory: collaborative and individual processes in remembering. *J Exp Psychol Learn Mem Cogn*. 1997;23(5):1160-75.
13. Brown W. Incidental Memory in a group of person. *Psychol Rev*. 1915;22(1):81-5.
14. Harber KD, Cohen DJ. The emotional broadcaster theory of social sharing. *J Lang Soc Psychol*. 2005;24(4):382-400.
15. Wenzel K, Reinhard M-A. Does the end justify the means? Learning tests lead to more negative evaluations and to more stress experiences. *Learn Motiv*. 2021;73: epub 101706.
16. Anderson FT, McDaniel MA. Restudying with the quiz in hand: when correct answer feedback is no better than minimal feedback. *J Appl Res Mem Cogn*. 2021:in press.
17. Anderson LW, Krathwohl DR. *A taxonomy for learning, teaching, and assessing : a revision of Bloom's taxonomy of educational objectives*. Complete ed. New York: Longman; 2001.
18. Hall EA, Spivey C, Kendrex H, Havrda DE. Effects of remote proctoring on composite exam performance Among Student Pharmacists. *Am J Pharm Educ*. 2021:8410.
19. Postman L, Knecht K. Encoding variability and retention. *Higher Educ* 1983;22(2):133-52.
20. Nairne JS. The myth of the encoding-retrieval match. *Memory*. 2002;10(5-6):389-95.
21. Brown NR, Buchanan L, Cabeza R. Estimating the frequency of nonevents: The role of recollection failure in false recognition. *Psychon Bull Rev*. 2000;7(4):684-91.
22. Koen JD, Aly M, Wang W-C, Yonelinas AP. Examining the causes of memory strength variability: recollection, attention failure, or encoding variability? *J Exp Psychol Learn Mem Cogn*. 2013;39(6):1726-41.
23. Ginot M, Py J, Colomb C. The differential effectiveness of the cognitive interview instructions for enhancing witnesses' memory of a familiar event. *Swiss J Psychol*. 2014;73(1):25-34.
24. Thorley C. Enhancing individual and collaborative eyewitness memory with category clustering recall. *Memory*. 2018;26(8):1128-39.
25. Verkamp F, Ginot M. Variations of the cognitive interview: Which one is the most effective in enhancing children's testimonies? *Appl Cogn Psychol*. 2010;24(9):1279-96.
26. McCabe DL, Trevino LK, Butterfield KD. Cheating in academic institutions: a decade of research. *Ethic Behav*. 2001;11(3):219-32.
27. Ginsburg-Block MD, Rohrbeck CA, Fantuzzo JW. A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *J Educ Psychol*. 2006;98(4):732-49.
28. Loda T, Erschens R, Loenneker H, et al. Cognitive and social congruence in peer-assisted learning - A scoping review. *PloS one*. 2019;14(9): 1-15e.
29. Tai J, Molloy E, Haines T, Canny B. Same-level peer-assisted learning in medical clinical placements: a narrative systematic review. *Med Educ*. 2016;50(4):469-84.

30. Williams B, Reddy P. Does peer-assisted learning improve academic performance? A scoping review. *Nurse Educ Today*. 2016;42:23-9.
31. Shteynberg G, Apfelbaum EP. The power of shared experience: simultaneous observation with similar others facilitates social learning. *Soc Psychol Personal Sci*. 2013;4(6):738-44.
32. McCabe D. Cheating: why students do it and how we can help them stop. *Am Educator*. 2001;25(4):38.
33. Pabian P. Why “cheating” research is wrong: New departures for the study of student copying in higher education. *Higher Educ*. 2015;69(5):809-21.
34. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve Tips for programmatic assessment. *Med Teach*. 2015;37(7):641-6.
35. Council NR. *High Stakes: Testing for Tracking, Promotion, and Graduation*. Washington, DC: The National Academies Press.; 1999.